

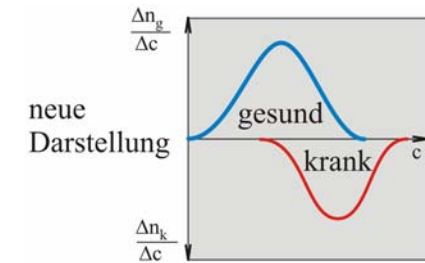
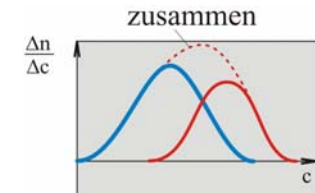
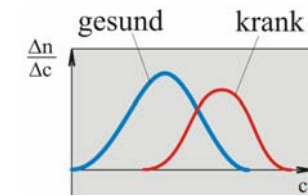
KAD 2013.11.27

## Überlappende Populationen

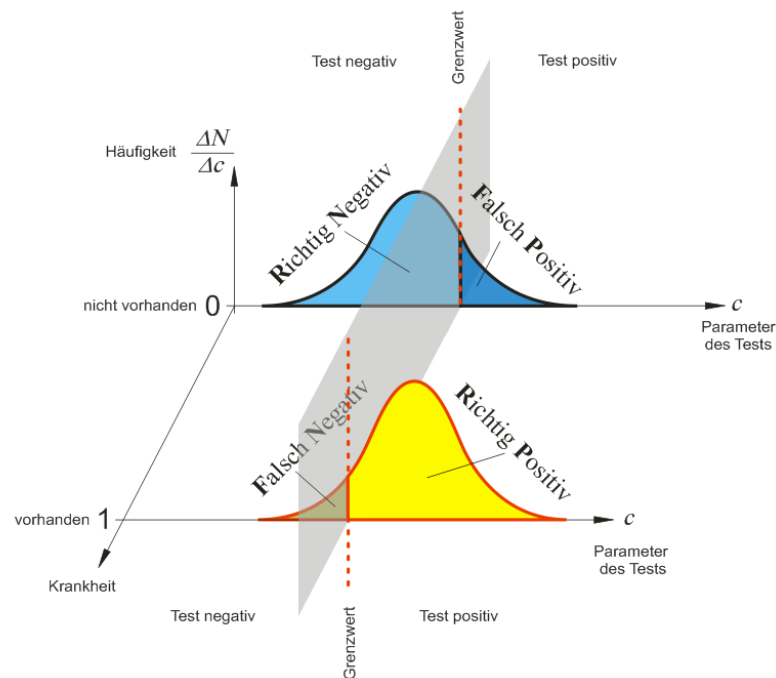
Annahme:

eine messbare Grösse (z.B. Konzentration) vergrößert sich in der kranken Population

die Veränderung ist wichtig, nicht die Vergrößerung



2

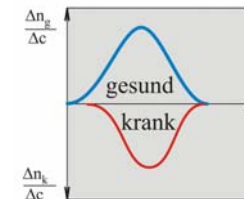


3

## Das Mass der Überlappung

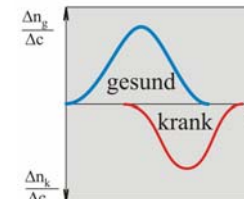
totale Überlappung

nutzlose Methode



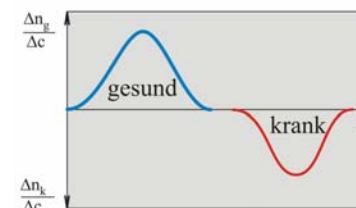
teilweise Überlappung

reelle Methoden



perfekte Separation

perfekte Methode

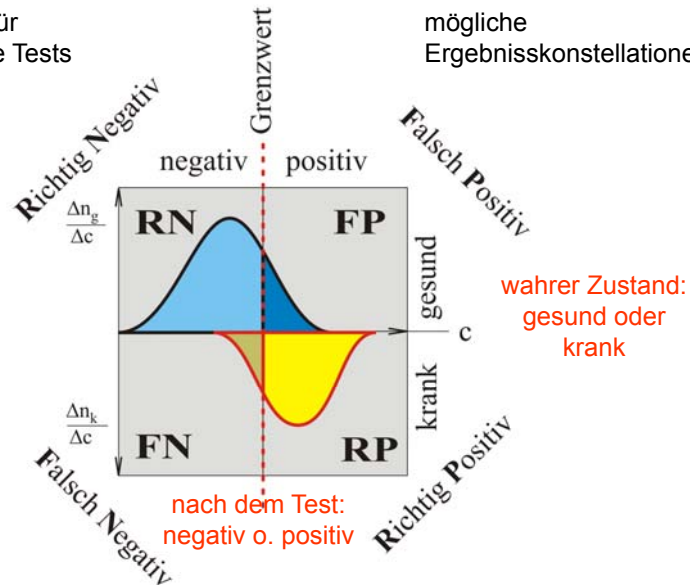


4

Basismatrix für  
diagnostische Tests

## Wahrheitsmatrix

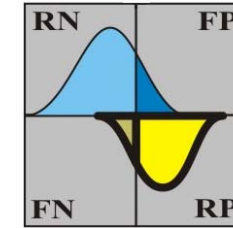
mögliche  
Ergebniskonstellationen



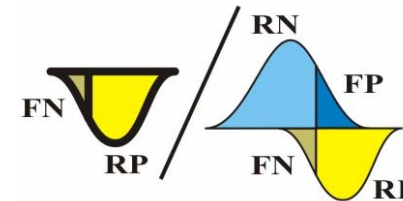
5

## Prävalenz

$$(w = \frac{de - sp}{se - sp})$$



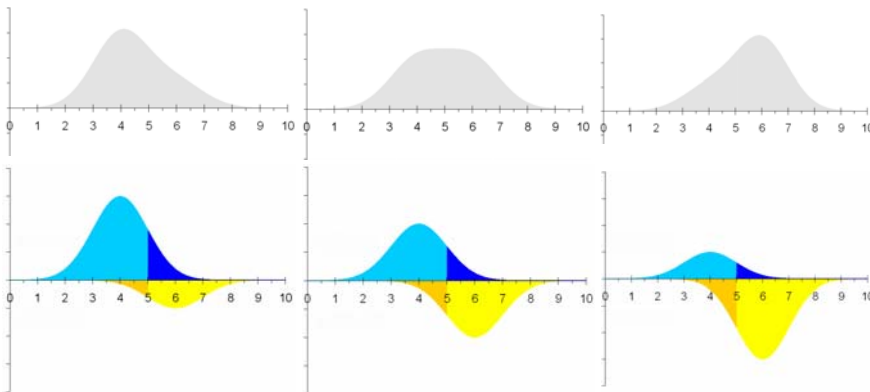
**Prävalenz:** Häufigkeit einer Krankheit in einer Population, Krankheitshäufigkeit, Wahrscheinlichkeit vor dem Test (Vortestwahrscheinlichkeit, a-priori-Wahrscheinlichkeit)



$$w = \frac{RP + FN}{RP + RN + FN + FP}$$

$$w = \frac{\text{alle Kranken}}{\text{alle Untersuchten}}$$

6



w = 25%

w = 50%

w = 75%

7

Die Zuverlässigkeit diagnostischer Tests wird mit den **Kennwerten** (Validätsparameter) beschrieben.

Sensitivität

Spezifität

Relevanz

Segreganz

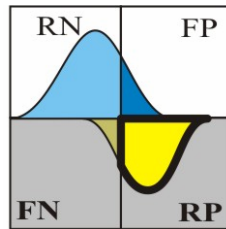
Jeder Test sollte an einem internationalen Standard geeicht werden, und es sollte eine **Referenzmethode** (Goldstandard) zur Erfassung des tatsächlichen Zustandes des Patienten verfügbar sein.



8

## Diagnostische Sensitivität

(se)



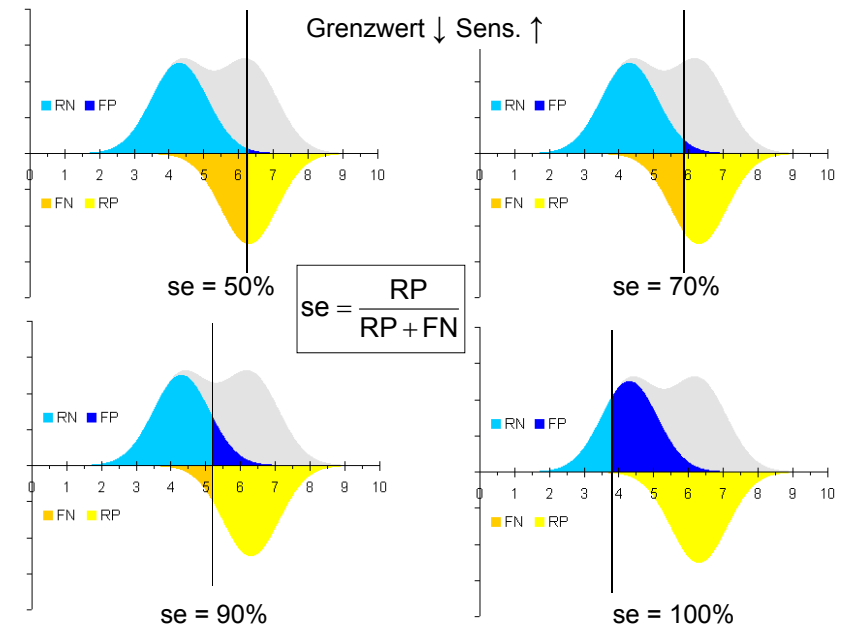
Empfindlichkeit. Wahrscheinlichkeit, einen Kranken als positiv zu erkennen



$$se = \frac{\text{richtig positiv}}{\text{krank}} = \frac{RP}{RP + FN}$$

Tests mit hoher Sensitivität sind bei der **Frühdagnostik** (screening) von Krankheiten erwünscht, und wenn es darauf ankommt, dass möglichst wenig Kranke unentdeckt bleiben.

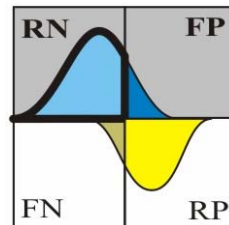
9



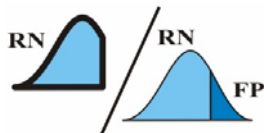
10

## Diagnostische Spezifität

(sp)



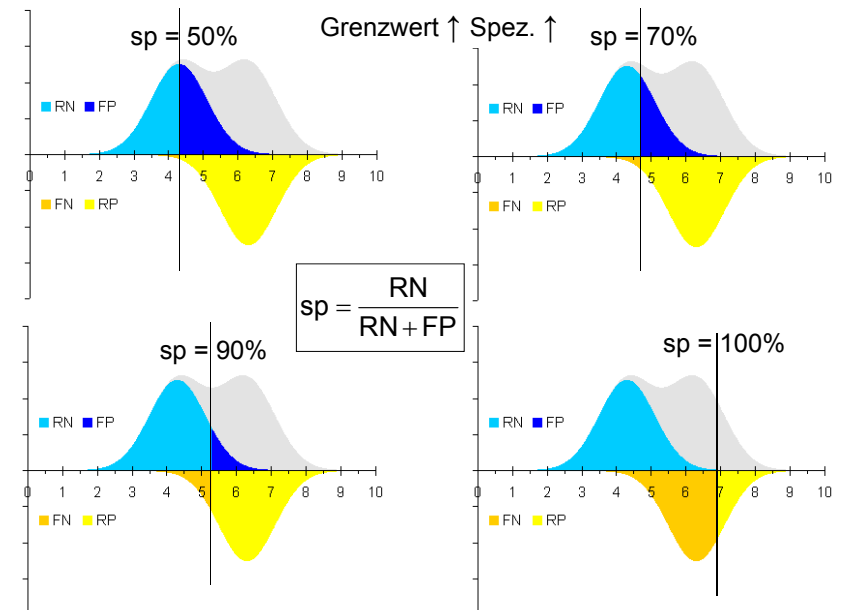
Wahrscheinlichkeit, einen Gesunden als negativ zu erkennen



$$sp = \frac{\text{richtig negativ}}{\text{gesund}} = \frac{RN}{RN + FP}$$

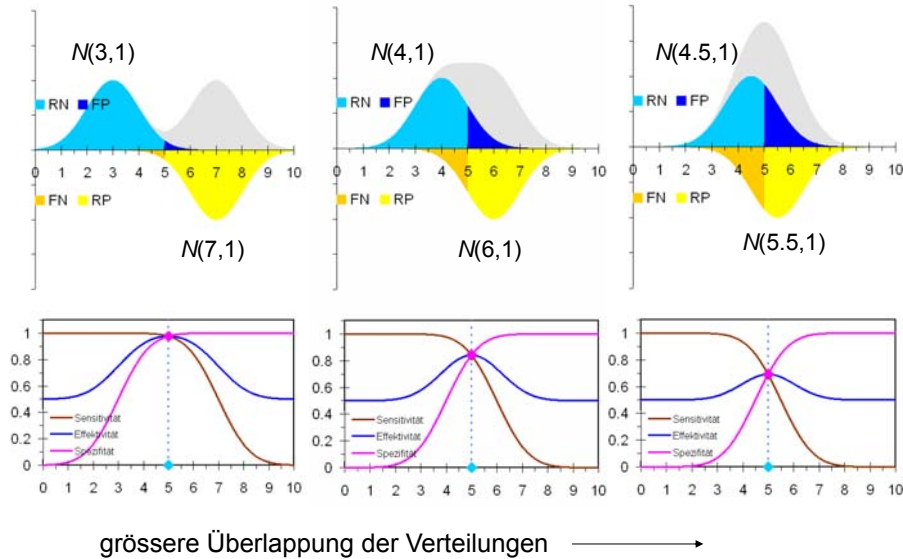
Tests mit hoher Spezifität sind als **Bestätigungstests** erwünscht und in allen Situationen, in denen eine falsch-positive Diagnose fatale Folgen hätte.

11



12

## Sensitivität u. Spezifität: gegenläufige Eigenschaften von Testen

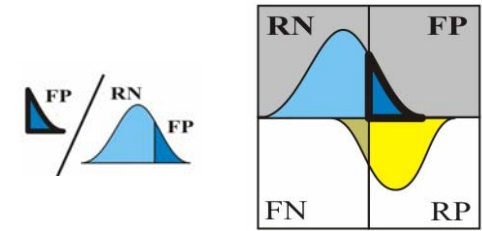


13

## Diagnostische Falschpositivrate

$$1 - sp = \frac{FP}{RN + FP}$$

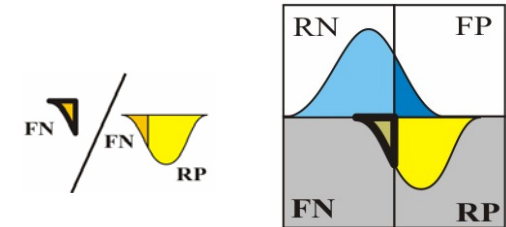
(vgl. Fehler 1. Art)



## Diagnostische Falschnegativrate

$$1 - se = \frac{FN}{FN + RP}$$

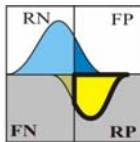
(vgl. Fehler 2. Art)



14

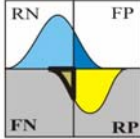
horizontale Raten hängen von der Prävalenz nicht ab

Sensitivität  
(se)



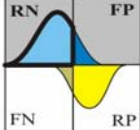
$$se = \frac{RP}{RP + FN}$$

Falschnegativrate  
(1-se)



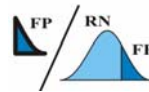
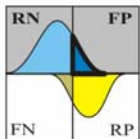
$$1 - se = \frac{FN}{FN + RP}$$

Spezifität  
(sp)



$$sp = \frac{RN}{RN + FP}$$

Falschpositivrate  
(1-sp)



$$1 - sp = \frac{FP}{RN + FP}$$

15

Vorhersagewerten (prädiktive Werte, vertikale Raten)

Wahrscheinlichkeiten nach dem Test, Nachtestwahrscheinlichkeiten, a-posteriori-Wahrscheinlichkeiten

## Diagnostische Relevanz

positiv prädiktiver Wert,  
positiver Vorhersagewert,  
positive predictive value

$$PPV = \frac{se \cdot w}{se \cdot w + (1 - sp) \cdot (1 - w)}$$

Wahrscheinlichkeit eines Test-Positiven, krank zu sein



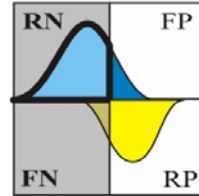
$$PPV = \frac{\text{richtig positiv}}{\text{positiv}} = \frac{RP}{RP + FP}$$

16

## Diagnostische Segreganz

negativ prädiktiver Wert,  
negativer Vorhersagewert,  
negative predictive value

$$NPV = \frac{sp \cdot (1 - w)}{sp \cdot (1 - w) + (1 - se) \cdot w}$$



Wahrscheinlichkeit eines Test-Negativen, gesund zu sein

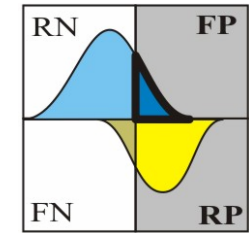


$$PVN = \frac{\text{richtig negativ}}{\text{negativ}} = \frac{RN}{RN + FN}$$

17

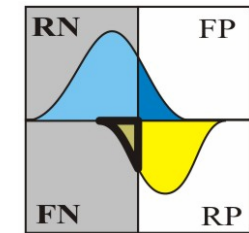
## Falschalarm(rate)

$$1 - PPV = \frac{FP}{FP + RP}$$



## falsche Beruhigung(srate)

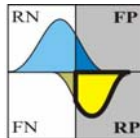
$$1 - NPV = \frac{FN}{FN + RN}$$



18

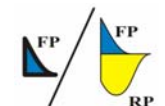
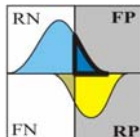
vertikale Raten hängen von der Prävalenz ab

Relevanz  
(PPV)



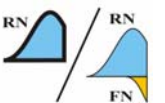
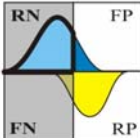
$$PPV = \frac{RP}{FP + RP}$$

Falschalarmrate  
(1-PPV)



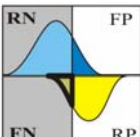
$$1 - PPV = \frac{FP}{FP + RP}$$

Segreganz  
(NPV)



$$NPV = \frac{RN}{RN + FN}$$

falsche Beruhigung  
(1-NPV)



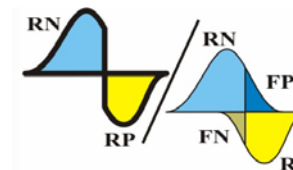
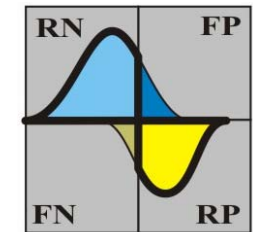
$$1 - NPV = \frac{FN}{RN + FN}$$

19

## Diagnostische Effektivität

richtige Klassifikationsrate

$$de = se \cdot w + sp \cdot (1 - w)$$



$$de = \frac{\text{richtig positiv} + \text{richtig negativ}}{\text{gesund} + \text{krank}}$$

$$de = \frac{RP + RN}{RN + FP + FN + RP}$$

oft: Grenzwert ist so gewählt, dass Effektivität maximal ist

20

## Effekt der Prävalenz

PVN = 90%

Beispiel A: **w = 50%**

sp = 90%

		Test	
		negativ	positiv
Gold-standard	gesund	90	10
	krank	10	90

se = 90%

(de = 90%)

PVP = 90%

PVN = 99%

Beispiel B: **w = 10%**

sp = 90%

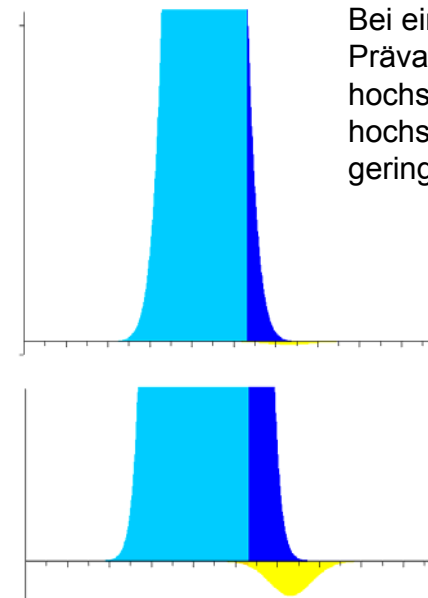
		Test	
		negativ	positiv
Gold-standard	gesund	810	90
	krank	10	90

se = 90%

(de = 90%)

PVP = 50%

21



Bei einer sehr kleineren Prävalenz können die hochsensitive und gleichzeitig hochspezifische Tests sehr geringe Relevanz haben.

Prävalenz = 0.1 %

Sensitivität = 98 %

Spezifität = 98 %

↓  
Relevanz = 4 %

22

## Übersichtstabelle

← bedingte wahrscheinlichkeit (Bayes)

Sensitivität	se	$\frac{RP}{RP+FN}$	$p(P K)$	Testpositiven zw. den Kranken	Richtigpositivrate, Empfindlichkeit
Spezifität	sp	$\frac{RN}{RN+FP}$	$p(N G)$	Testnegativen zw. den Gesunden	Richtignegativrate
Falschnegativrate	1-se	$\frac{FN}{RP+FN}$	$p(N K)$	Testnegativen zw. den Kranken	
Falschpositivrate	1-sp	$\frac{FP}{RN+FP}$	$p(P G)$	Testpositiven zw. den Gesunden	
Relevanz; positiver prädiktiver Wert	PVP	$\frac{RP}{RP+FP}$	$p(K P)$	Kranken zw. den Testpositiven	positiver Vorhersagewert
Segreganz; negativer prädiktiver Wert	PVN	$\frac{RN}{RN+FN}$	$p(G N)$	Gesunden zw. den Testnegativen	negativer Vorhersagewert
Falschalarmrate	1-PVP	$\frac{FP}{RP+FP}$	$p(G P)$	Gesunden zw. den Testpositiven	Fehlalarmrate
falsche Beruhigungsrate	1-PVN	$\frac{FN}{RN+FN}$	$p(K N)$	Kranken zw. den Testnegativen	

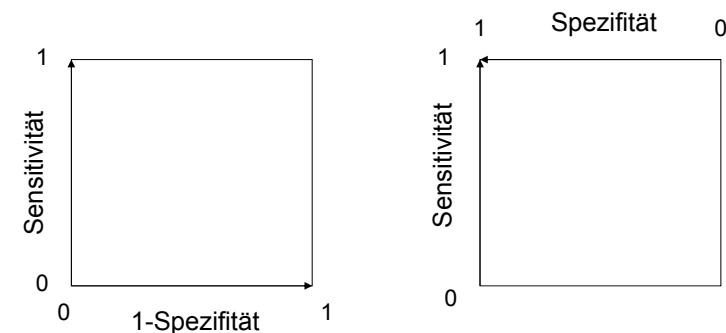
23

## Vergleichung verschiedener diagnostischer Methode. ROC Kurven

ROC: receiver-operator (operating) characteristic

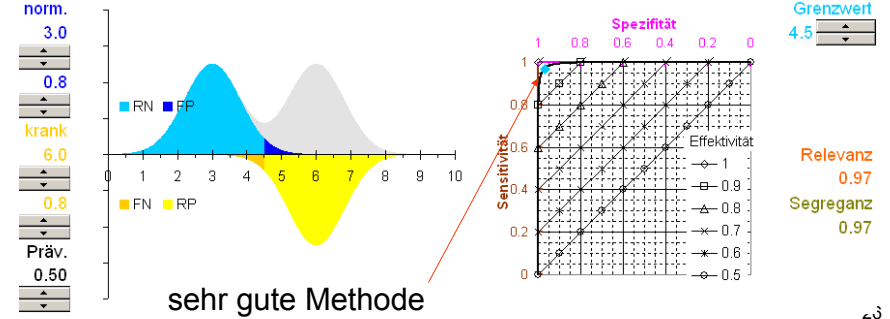
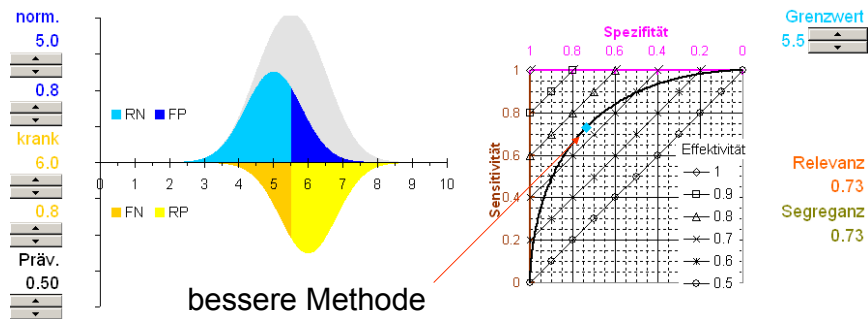
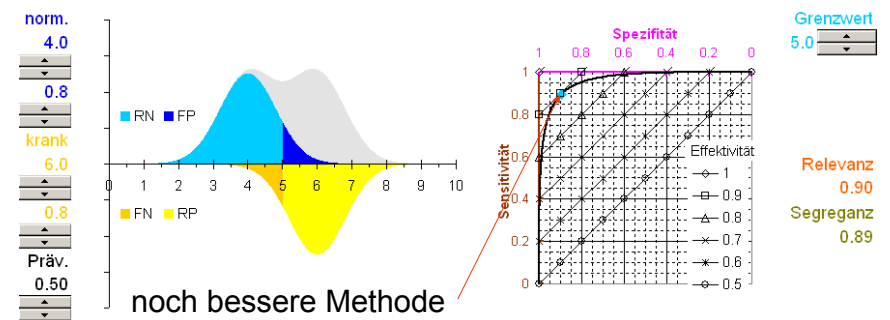
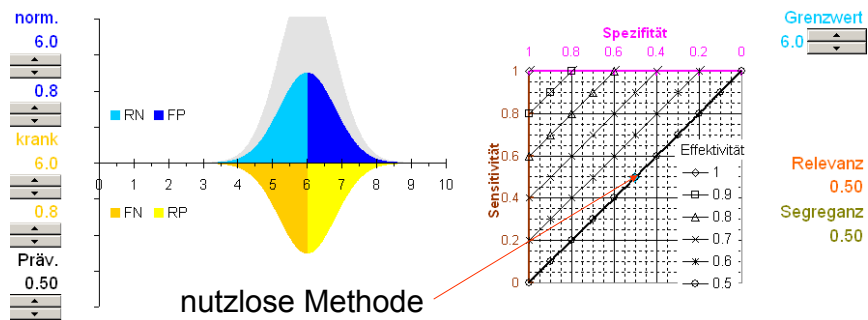
ca. 1950: erste ROC Analyse (receiver: Radar Empfänger)

ca. 1970: die erste medizinische Anwendungen

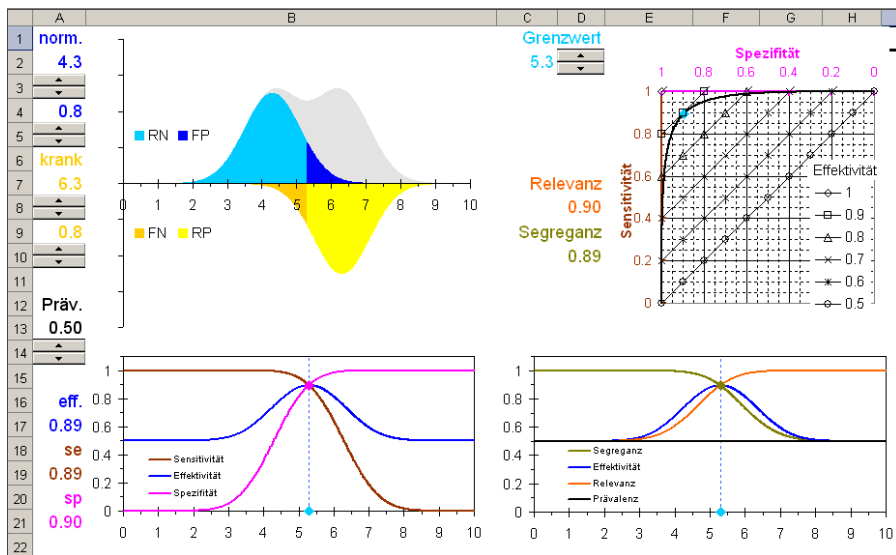


24



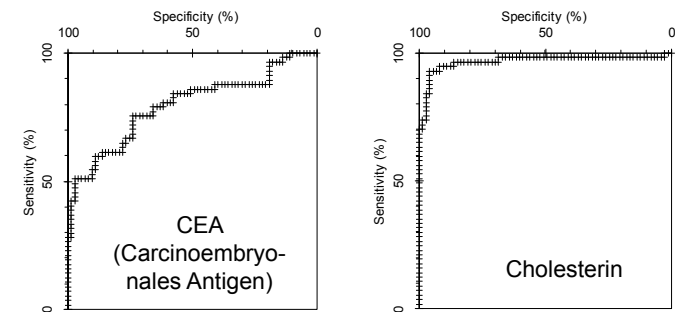


## ROC Analyse



## Beispiel: Tumormarker im Bauchwasser (Ascites)

Die Erhöhung von CEA (und/oder Cholesterin) Konzentration im Bauchwasser kann mit Karzinose in Zusammenhang bringen.



Welche Methode ist besser? Wie kann man den optimalen Grenzwert auswählen?

Gulyás M, Kaposi AD, Elek G, Szollár LG, Hjerpe A, Value of carcinoembryonic antigen (CEA) and cholesterol assays of ascitic fluid in cases of inconclusive cytology, J Clinical Pathology 2001 (54) 831-835

$$de = se \cdot w + sp \cdot (1 - w)$$

$$\frac{de}{1 - w} = \frac{w}{1 - w} se + (sp - 1) + 1$$

$$(1 - sp) + \frac{de}{1 - w} - 1 = \frac{w}{1 - w} se$$

$$se = \left( \frac{1 - w}{w} \right) (1 - sp) + \left( \frac{1}{w} de + \frac{w - 1}{w} \right)$$

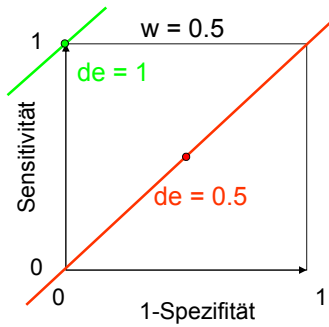
Steigung

Achsenabschnitt

wenn  $w = 0.5$ :  $se = 1 \cdot (1 - sp) + 2 \cdot de - 1$

Die Punkte, die gleiche diagnostische Effektivität haben, sind auf den Geraden mit einer Steigung von 1.

Wenn  $de = 0.5$  ist, dann beträgt der Achsenabschnitt 0.



$$se = \left( \frac{1 - w}{w} \right) (1 - sp) + \left( \frac{1}{w} de + \frac{w - 1}{w} \right)$$

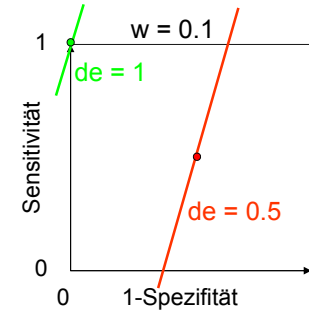
Steigung

Achsenabschnitt

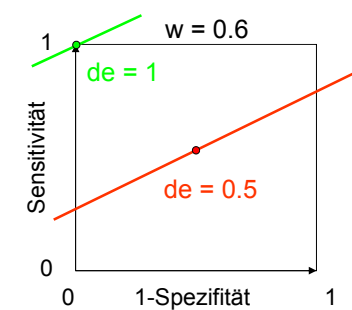
wenn  $w < 0.5$ : Die Steigung der Geraden mit gleicher diagnostischen Effektivität ist grösser als 1.

wenn  $w > 0.5$ : Die Steigung der Geraden mit gleicher diagnostischen Effektivität ist kleiner als 1.

z.B.  $w = 0.1$ , die Steigung: 9



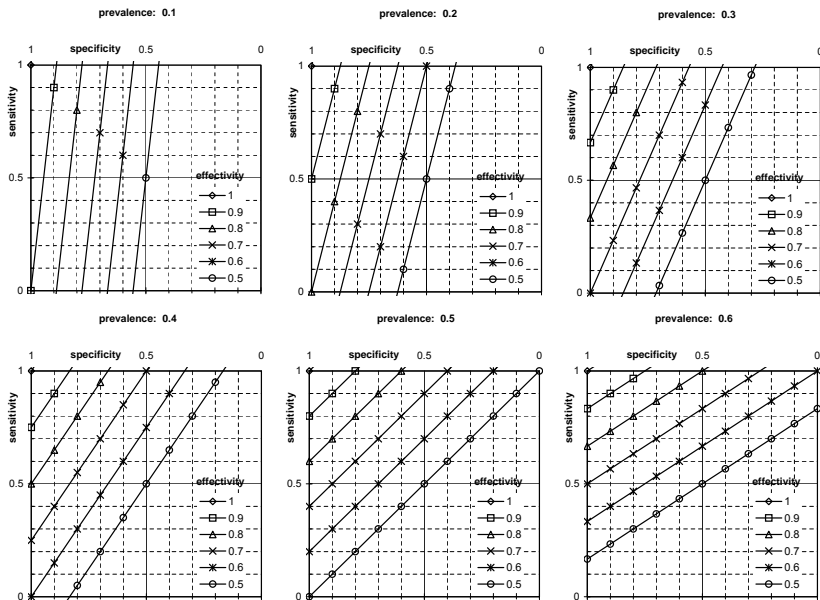
z.B.  $w = 0.6$ , Steigung: 0.66



29

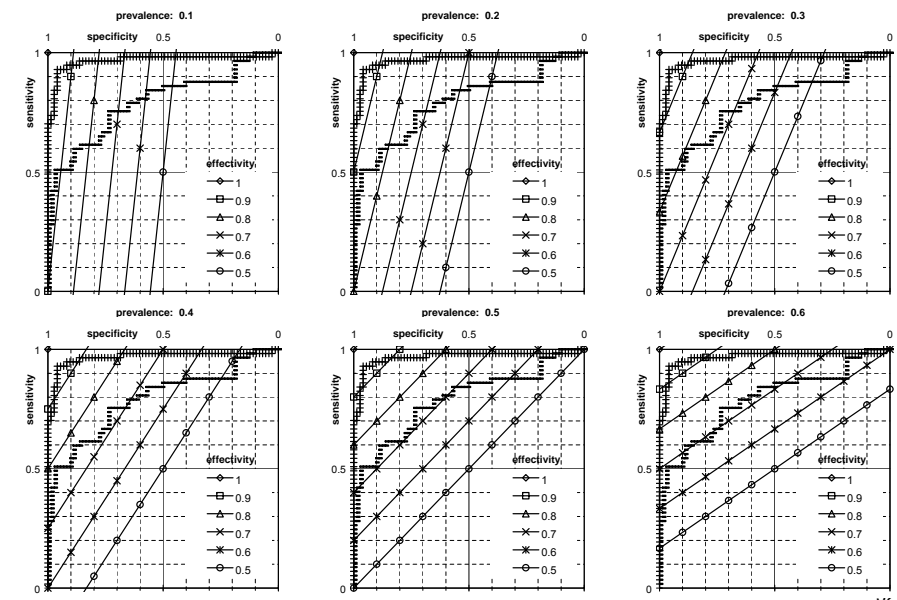
30

## Isoeffektive Kurven auf ROC



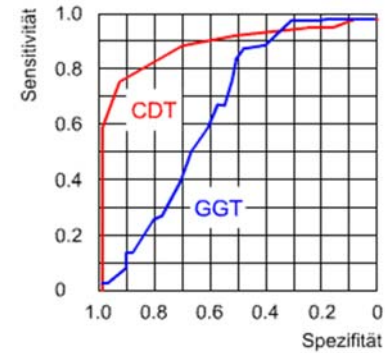
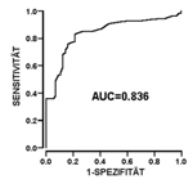
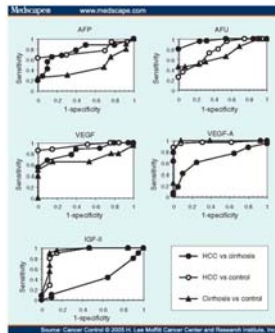
31

## Ascites (+ Cholesterin, – CEA)





## Weitere Beispiele



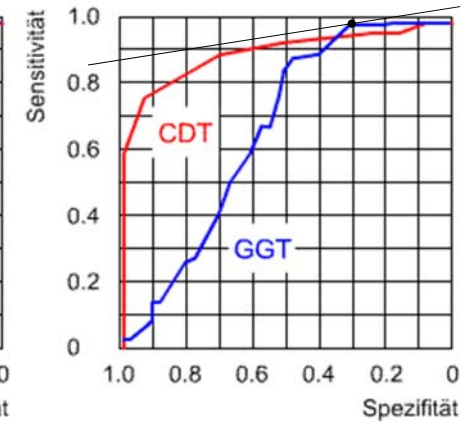
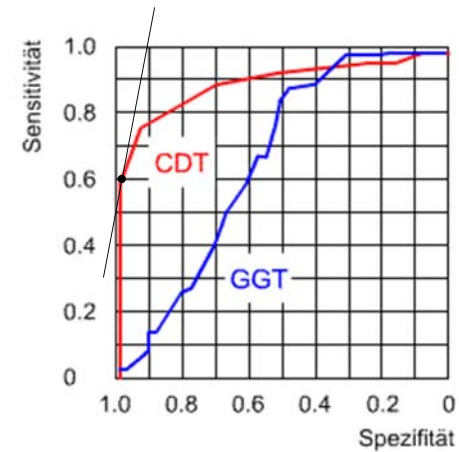
ROC für CDT (carbohydrate deficient transferrin) und GGT (gamma-Glutamyltransferase) in Bezug auf Alkoholismus. Da CDT praktisch immer auf der linken, oberen Seite der GGT liegt, ist CDT ein wesentlich besser Test für Alkoholkonsum als GGT

33

## Beispiel: maximalisieren wir die diagnostische Effektivität!

bei einem kleineren Prävalenzwert ist die CDT Methode besser

bei einem höheren Prävalenzwert ist die GGT Methode besser



34