

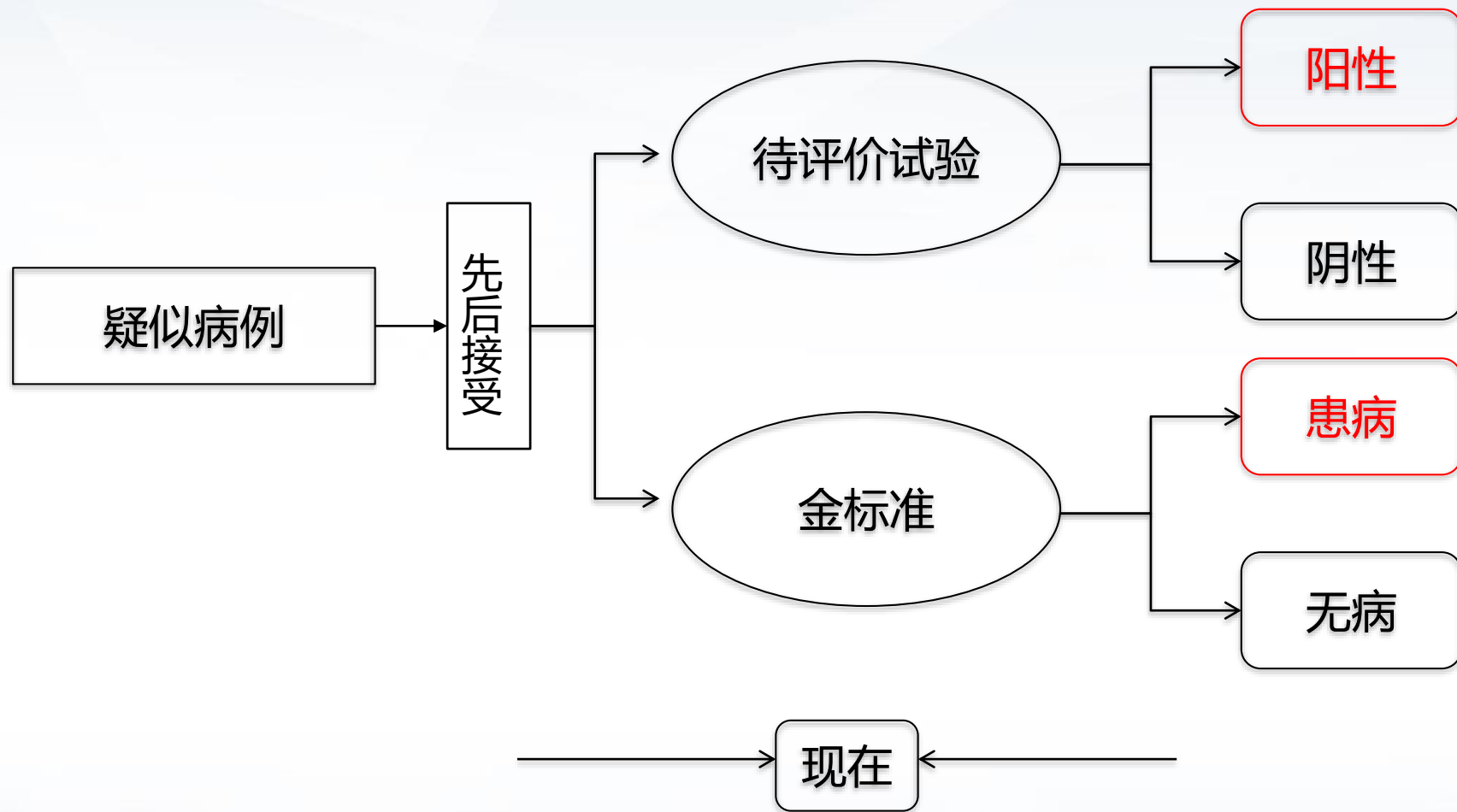


诊断试验的ROC分析

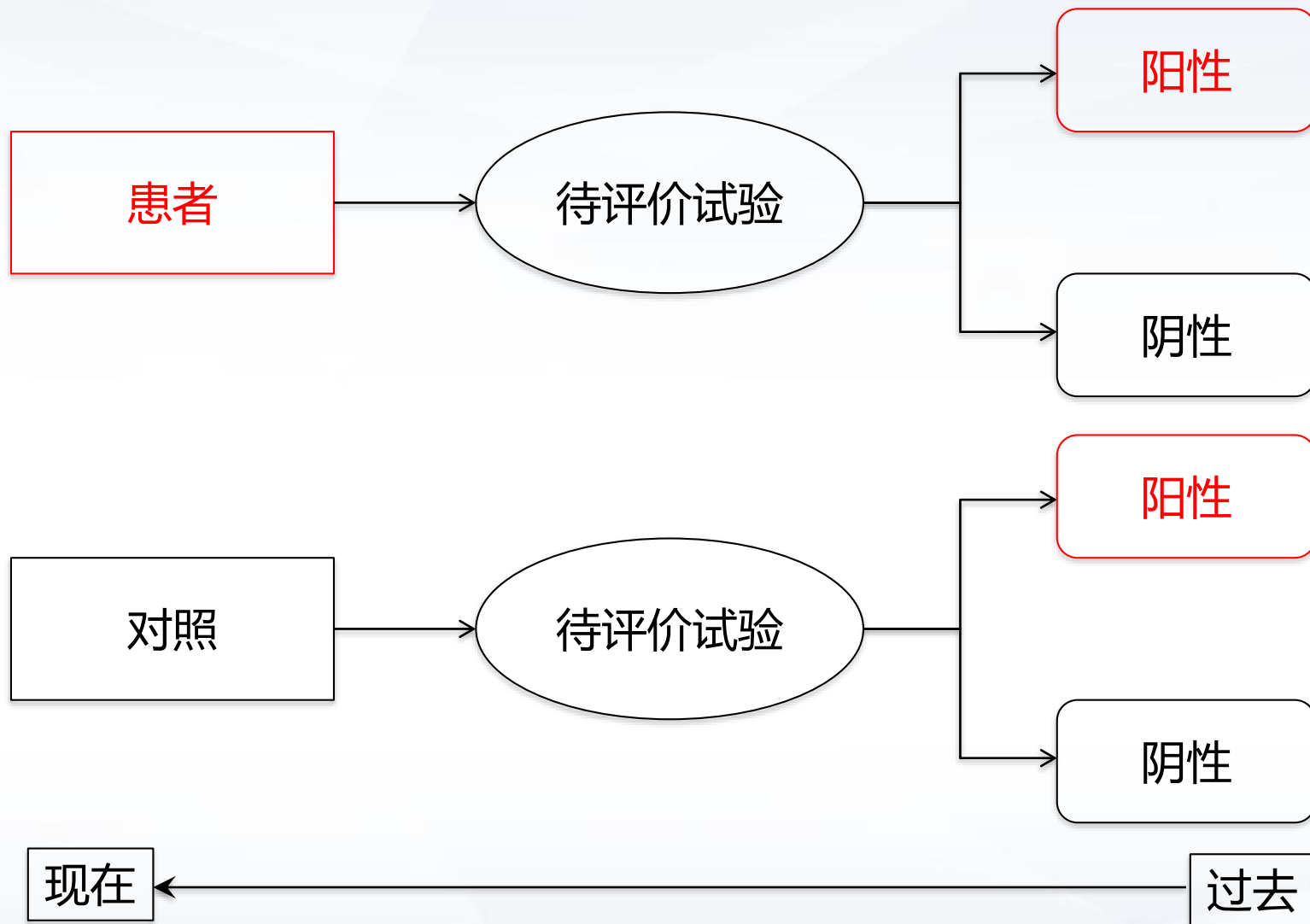
复旦大学附属肿瘤医院 周支瑞



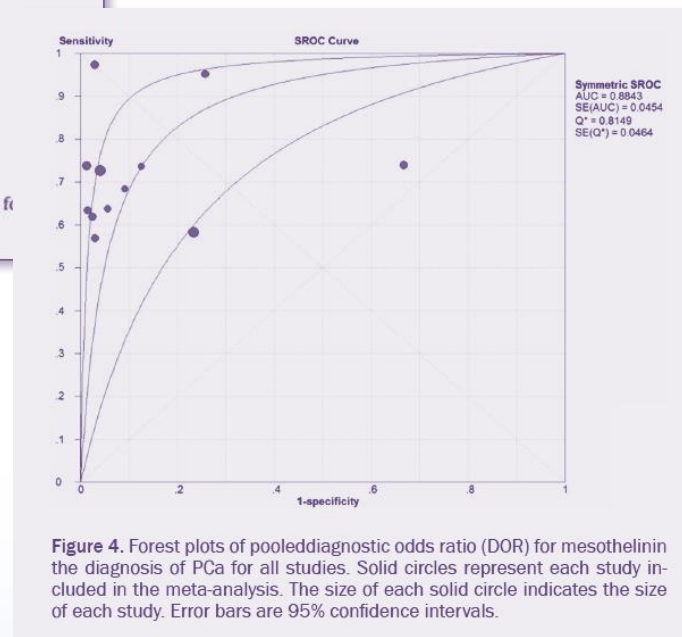
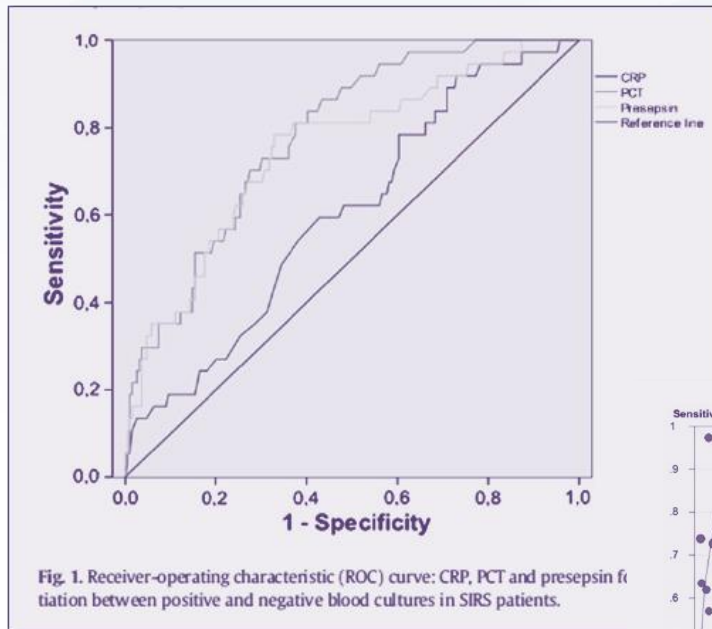
诊断试验设计原理 -- 横断面设计/单门设计



诊断试验设计原理 -- 病例对照设计/双门设计



- 患病率
- 敏感性
- 特异性
- 阳性预测值
- 阴性预测值
- 阳性似然比
- 阴性似然比
- 诊断比数比
- 曲线下面积

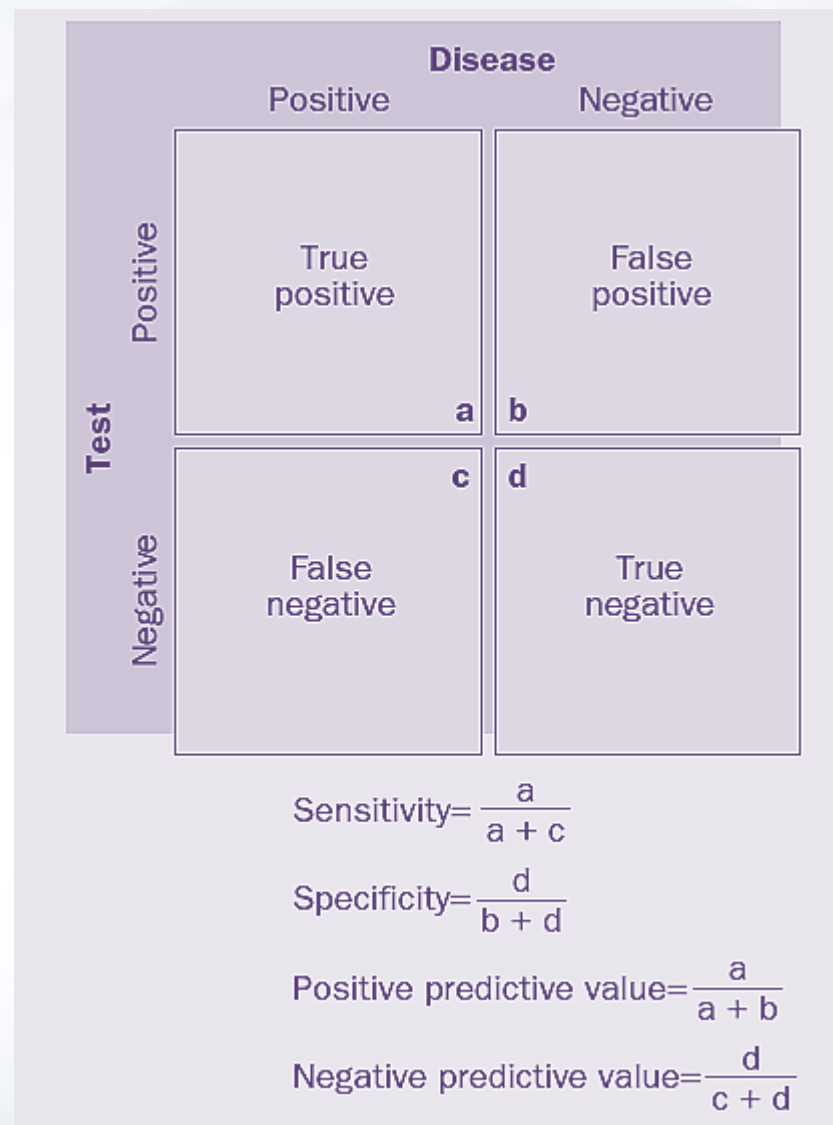


诊断试验的统计指标基本概念

表X. 配对四格表排列

		参考标准诊断结果		
		患病	无病	合计
待评价检 查的结果	阳性	a (真阳性)	b (假阳性)	a+b
	阴性	c (假阴性)	d (真阴性)	c+d
合计		a+c	b+d	N

- 敏感性 = $a/(a+c)$
- 特异性 = $d/(b+d)$
- 阳性预测值(PPV) = $a/(a+b)$
- 阴性预测值NPV = $d/(c+d)$
- 阳性似然比(PLR) = $\frac{a/(a+c)}{b/(b+d)}$ 真阳性与假阳性之比
- 阴性似然比(NLR) = $\frac{d/(b+d)}{c/(a+c)}$ 真阴性与假阴性之比
- 准确率 = $(a+d)/(N)$



例21-1 采用 ECG 对具有急性持久胸痛的 700 名患者进行诊断，经证实有 520 例出现了心肌梗塞，其余 180 例没有出现（见表 21-2），试计算 ECG 诊断试验的几个常用评价指标。

表 21-2

ECG 诊断试验的结果

ECG 诊断结果	心肌梗塞		合 计
	出 现	不 出 现	
阳性	416(TP)	9(FP)	425
阴性	104(FN)	171(TN)	275
合计	520	180	700(N)

(1) 一致百分率。

病例正确诊断为阳性与对照正确诊断为阴性的例数之和占总例数的百分率。计算公式为：

$$\text{一致百分率} = \frac{TP + TN}{N} \times 100\%$$

其标准误为：

$$SE_{\text{一致百分率}} = \sqrt{(TP + TN)(FP + FN) / N^3}$$

$$\text{本例一致百分率} = \frac{416 + 171}{700} \times 100\% = 83.86\%$$

$$SE_{\text{一致百分率}} = \sqrt{(416 + 171)(9 + 104) / 700^3} = 1.39\%$$

(2) 灵敏度。

实际患病且被诊断为阳性的概率就是灵敏度，也称为真阳性率，即：

$$Sen = \frac{TP}{TP + FN} = TPR$$

其标准误为：

$$SE_{Sen} = \sqrt{TP \times FN / (TP + FN)^3} = \sqrt{Sen(1 - Sen) / (TP + FN)}$$

本例 $Sen=416/520=0.8$ ，即真阳性率 $TPR=0.8$ ，80%出现心肌梗塞患者被 ECG 诊断为阳性。其标准误为：

$$SE_{Sen} = \sqrt{Sen(1 - Sen) / (TP + FN)} = \sqrt{0.8(1 - 0.8) / 520} = 1.75\%$$

(3) 特异度。

实际未患病且被诊断为阴性的概率就是特异度，即：

$$Spe = TN / (TN + FP)$$

其标准误为：

$$SE_{Spe} = \sqrt{FP \times TN / (FP + TN)^3} = \sqrt{Spe(1 - Spe) / (FP + TN)}$$

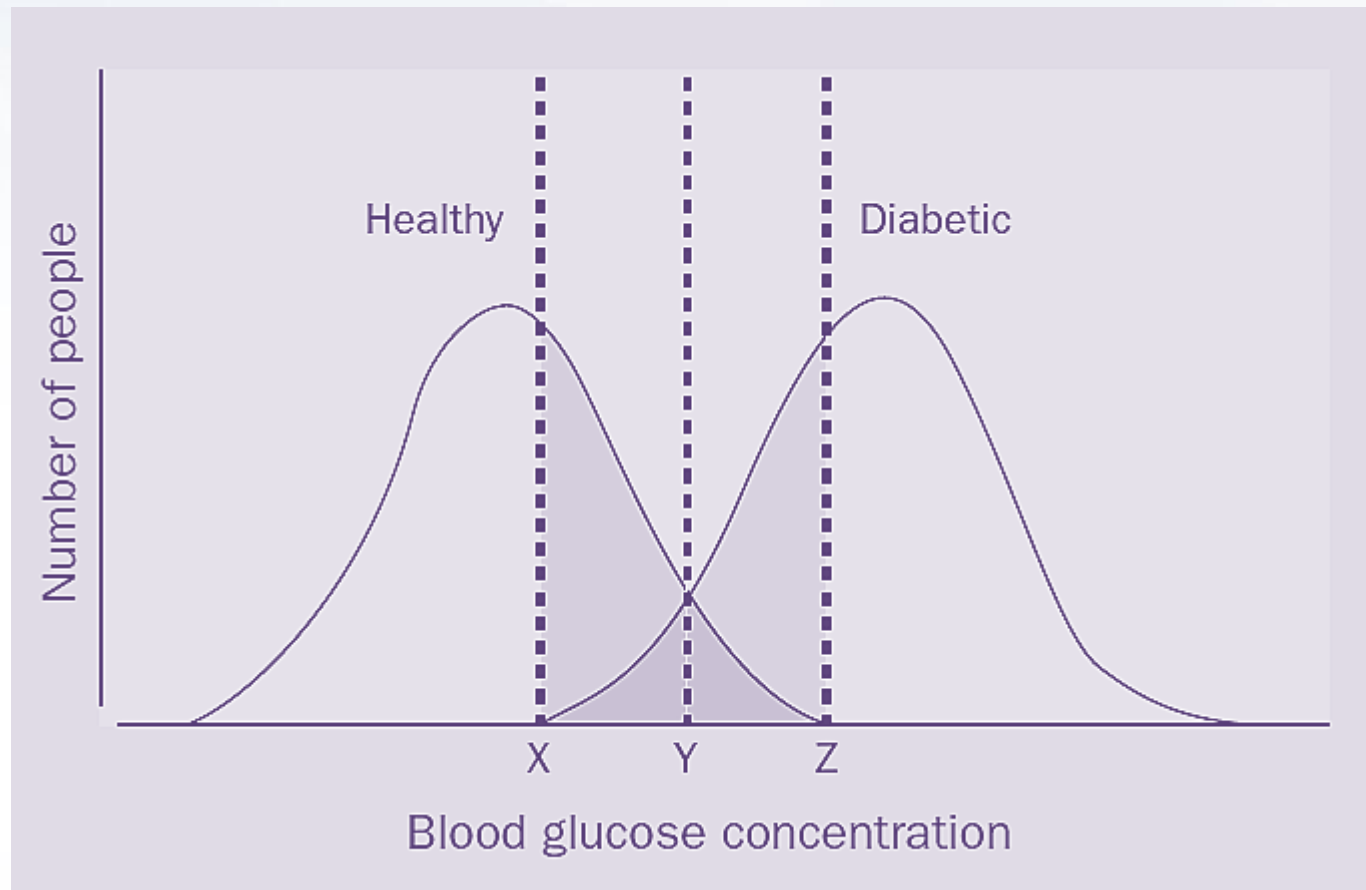
本例 $Spe = 171/180 = 0.95$ ，即 95% 未出现心肌梗塞患者诊断结果为阴性。其标准误为：

$$SE_{Spe} = \sqrt{0.95(1 - 0.95) / 180} = 1.62\%$$

该指标只与对照组有关，反映了诊断试验排除非病例的能力。

由以上公式可以导出漏诊率 $\beta = 1 - Sen = FN / (TP + FN)$ ，漏诊率也就是假阴性率；也可导出误诊率 $\alpha = 1 - Spe = FP / (FP + TN)$ ，误诊率也就是假阳性率。

本例漏诊率 $\beta = 1 - sen = 1 - 0.8 = 0.2$ ，误诊率 $\alpha = 1 - spe = 1 - 0.95 = 0.05$ ，即假阳性率 $FPR = 0.05$ 。



- 减少漏诊必然会增加误诊
- 减少误诊必然会增加漏诊
- 提高敏感性必然降低特异性
- 提高特异性必然降低敏感性

(4) Youden 指数

真阳性率与假阳性率之差就是 Youden 指数，即：

$$J = Sen + Spe - 1 = TPR - FPR$$

其标准误为：

$$\begin{aligned} SE_J &= \sqrt{TP \times FN / (TP + FN)^3 + FP \times TN / (FP + TN)^3} \\ &= \sqrt{Sen(1 - Sen) / (TP + FN) + Spe(1 - Spe) / (FP + TN)} \end{aligned}$$

本例 $J = 0.8 - 0.05 = 0.75$ ，即 Youden 指数为 0.75。其标准误为：

$$SE_J = \sqrt{0.8(1 - 0.8) / 520 + 0.95(1 - 0.95) / 180} = 0.0239$$

(5) 阳性似然比

真阳性率与假阳性率之比，即灵敏度与误诊率之比就是阳性似然比 (LR_+)。

$$LR_+ = TPR / FPR = Sen / (1 - Spe)$$

本例 $LR_+ = 0.8 / 0.05 = 16$ ，即阳性似然比为 16。

LR_+ 的取值范围为 $(0, \infty)$ ，其值越大，检测方法证实疾病的能力越强。

LR_+ 的标准误涉及到对数变换，这里不予给出。以下几个指标的标准误计算也较复杂，也不予给出。

配对四格表 -- 计算阴性似然比

(6) 阴性似然比

假阴性率与真阴性率之比，即漏诊率与特异度之比就是阴性似然比（ LR_- ）。

$$LR_- = (1 - TPR) / (1 - FPR) = (1 - Sen) / Spe$$

本例 $LR_- = (1 - 0.8) / 0.95 = 0.2105$ ，即阴性似然比为 0.2105。

LR_- 的取值范围为 $(0, \infty)$ ，其值越小，检测方法排除疾病的能力越好。

(7) 阳性预报值

通常情况下，当要对某疾病作出诊断时，并不知道金标准的结果，只知道诊断试验是阳性或阴性。而临床医生更想知道的是：当诊断试验为阳性时，受试者真正有病的概率有多大，为阴性时又有多大把握能排除此病。这就需要引入阳性预报值（ PV_+ ）与阴性预报值（ PV_- ）的概念。

诊断试验结果为阳性时，受试者实际为病例的概率就是阳性预报值，即：

$$PV_+ = TP / (TP + FP) = \frac{SenP_0}{SenP_0 + (1 - Spe)(1 - P_0)}$$

由式中可以看出，当灵敏度与特异度为常数时，增加患病率，将降低 $(1 - Spe)(1 - P_0)$ ，增加 $SenP_0$ 的值，从而整个分母的值减少，阳性预报值增加。

(8) 阴性预报值

诊断试验结果为阴性时，受试者实际为非病例的概率就是阴性预报值（ PV_- ），即：

$$PV_- = TN / (TN + FN) = \frac{SpeP_0}{SpeP_0 + (1 - Sen)(1 - P_0)}$$

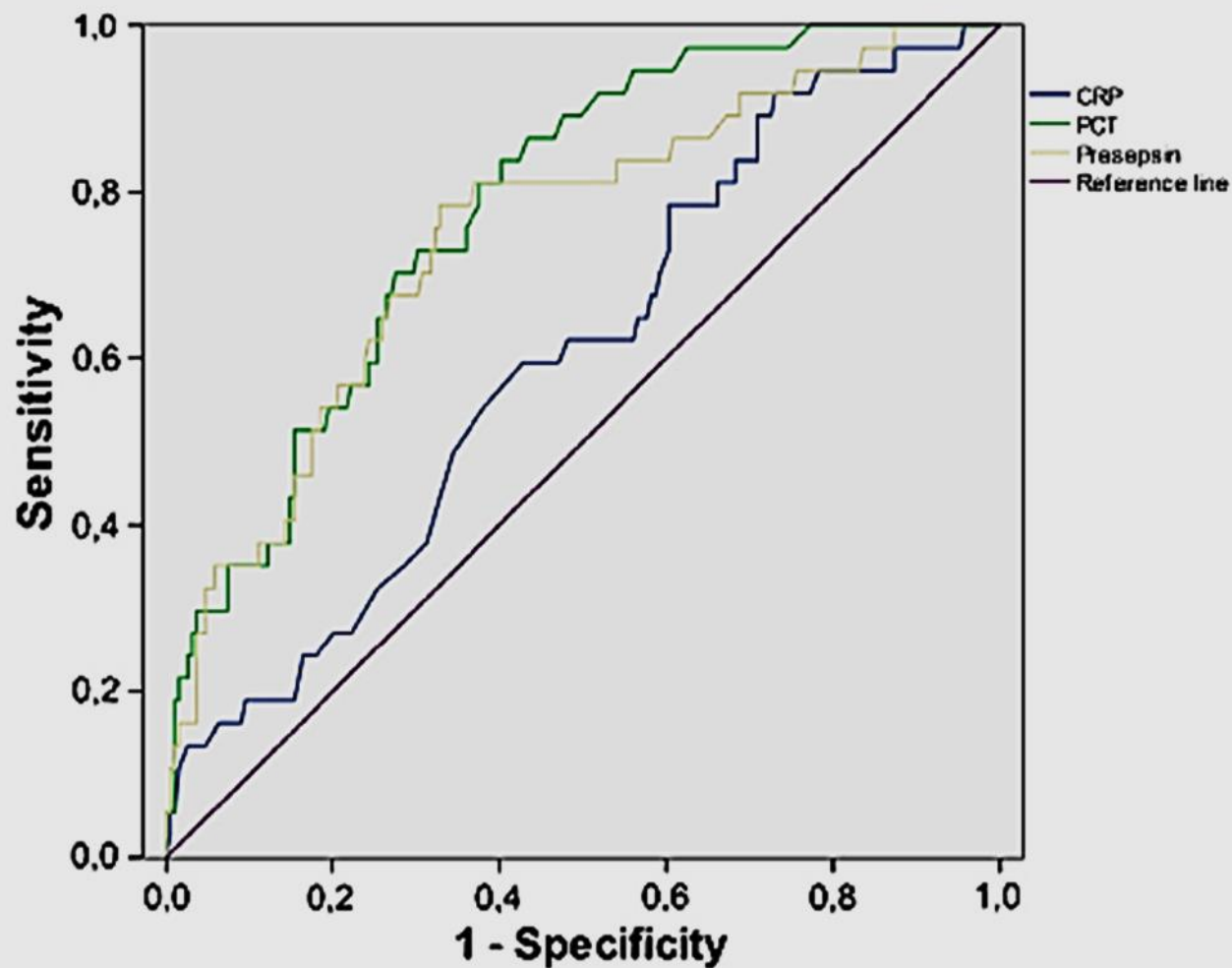
式中符号的意义与前面相同。当灵敏度与特异度为常数时，增加患病率将降低阴性预报值。

PV_+ 和 PV_- 的取值范围均在（0，1）之间。对于相同的患病率，其值越接近 1，检测方法的诊断价值越高。

【举例】AFP和DCP诊断肝癌

- 当AFP的诊断界值设定为15ng/ml时，诊断肝癌的敏感性为70%，特异性为80%。
- 当DCP的诊断界值设定为20ng/ml时，诊断肝癌的敏感性为80%，特异性为70%。
- 问题：AFP和DCP诊断肝癌的能力孰优孰劣呢？

受试者工作特征曲线 (ROC)



- 曲线本质是不同界值下敏感性和特异性的组合；ROC曲线下面积(AUC) 介于0.5-1.0之间，AUC的大小反映试验诊断疾病的能力
- ROC反映多个诊断界点下诊断能力的总和；便于比较两个试验的诊断能力；确定最佳诊断界值。

ROC分析 案例 1

例 21-2 假设某诊断试验的病例组和对照组分别有 25 和 20 例受试者，其检测结果见表 21-3。试进行 ROC 分析，确定合适的临界点。

表 21-3

病例组与对照组的检测值

病例组 ($n_1=25$)	X_1	6.5 13.5 12.8 6.2 13.9 14.7 9.5 9.0 6.9 16.8 13.3 10.8 12.2 14.9 13.7 12.8 5.3 11.8 12.4 7.6 13.3 11.9 11.2 12.3 12.7
对照组 ($n_2=20$)	X_2	8.5 6.4 4.6 1.7 9.7 5.3 4.9 5.7 3.8 6.5 6.3 5.4 3.3 4.7 8.6 6.3 5.9 4.8 4.5 5.2

诊断试验 案例1 代码

```
> example21_2 <- read.table ("example21_2.csv", header=TRUE, sep=",")
> example21_2
> attach(example21_2)
> summary(example21_2)
> str(example21_2)
> install.packages("ROCR")
> library(ROCR)
> pred <- prediction(example21_2$value, example21_2$group)
> pred
> perf <- performance(pred,"tpr","fpr")
> plot(perf)
> perf1 <- performance(pred, "prec", "rec")
> plot(perf1)
> perf2 <- performance(pred, "sens", "spec")
> plot(perf2)
> auc <- performance(pred,"auc")
> auc
```

例21-3 假设某研究欲比较HPV-DNA与HPV E6/E7 mRNA对中重度子宫颈上皮不典型增生的诊断价值。试进行ROC分析。

诊断试验 案例2 代码

```
> example21_3 <- read.table ("example21_3.csv", header=TRUE, sep=",")
> attach(example21_3)
> summary(example21_3)
> str(example21_3)
> install.packages("ROCR")
> library(ROCR)
> pred1 <- prediction(example21_3$mRNA, example21_3$oncology)
> pred1
> pred2 <- prediction(example21_3$dna, example21_3$oncology)
> pred2
> perf1 <- performance(pred1,"tpr","fpr")
> perf2 <- performance(pred2,"tpr","fpr")
> auc1 <- performance(pred1,"auc")
> auc1
> auc2 <- performance(pred2,"auc")
> auc2
> plot(perf1, col="blue")
> plot(perf2, col="red", add=TRUE)
```

祝各位月月有数据，年年有文章！



Thanks!

