

Biostatistics Seminar 2019

医学統計セミナー

- 臨床検査値の評価 -

下川敏雄

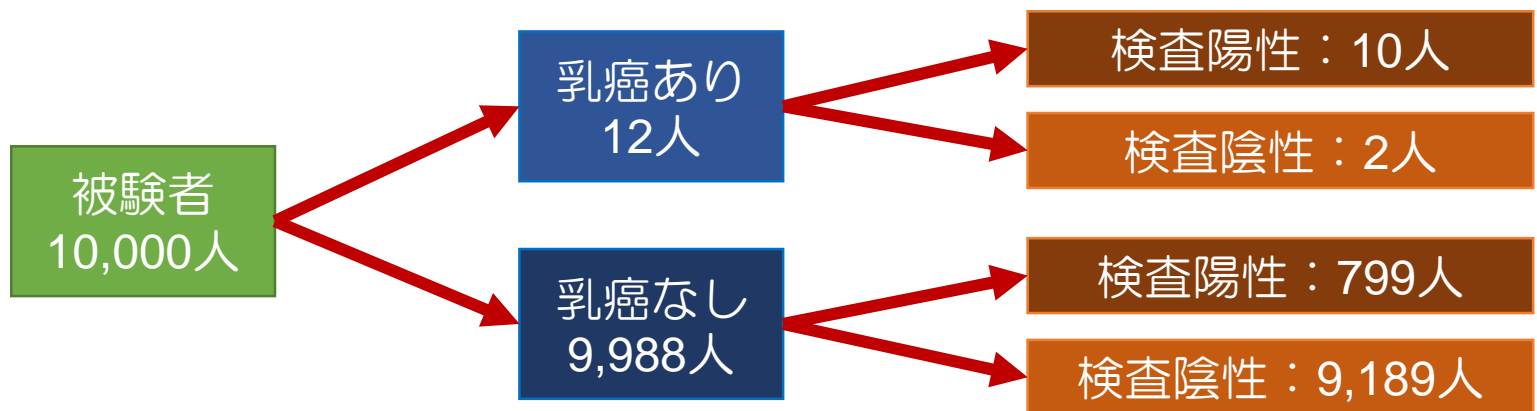
和歌山県立医科大学附属病院 臨床研究センター

例示：マンモグラフィ検査による乳癌検査の仮想例(新谷, 2015)

乳癌に対するマンモグラフィ検査の予測確度を評価するために、10,000人の検査結果と乳癌の有無を調査した。

	乳癌あり	乳癌なし	合計
検査陽性	10	799	809
検査陰性	2	9,189	9,191
合計	12	9,988	10,000

乳癌とマンモグラフィ検査の時間的な流れ



感度・特異度の留意点

感度(Sensitivity)

感度 = $TP/(TP+FN)$: 疾患有の被験者を陽性と正しく判断する確率

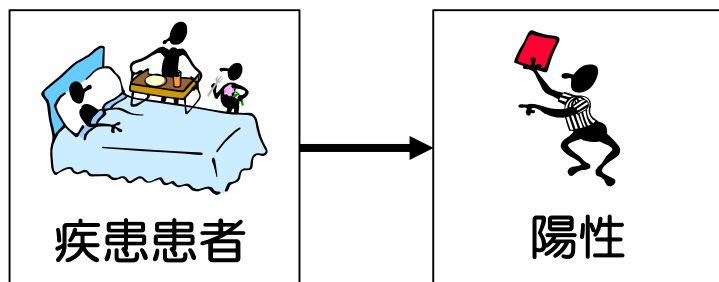
特異度(Specificity)

特異度 = $TN/(TN+FP)$: 疾患無の被験者を陰性と正しく判断する確率

利点：当該疾患の有病率(prevalence)に影響されず診断性能を評価できる。

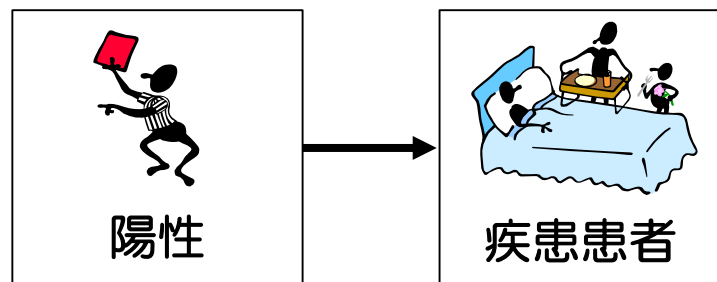
注意点：「感度 ≠ 疾患の確率」であることに注意する。

感度



感度は、結果がわかっている状況のなかで評価している。

臨床的有用性



陽性と診断されたときに、本当に疾患だった確率が重要になる。

感度・特異度

	乳癌あり	乳癌なし	合計
検査陽性	10 (TP)	799 (FP)	809
検査陰性	2 (FN)	9,189 (TN)	9,191
合計	12	9,988	10,000

	疾患あり	疾患なし
陽性	感度 真陽性 (TP: True Positive)	特異度 偽陽性 (FP: False Positive)
陰性	偽陰性 (FN: False Negative)	真陰性 (TN: True Negative)

感度 (Sensitivity) : 疾患ありの被験者を陽性と正しく診断する確率

$$\text{感度} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

事例の場合には、TP10例、FN2名なので
感度 = $\frac{10}{10+2} = 0.833$ (83.3%)

特異度 (Specificity) : 疾患なしの被験者を陰性と正しく診断する確率

$$\text{特異度} = \frac{\text{TN}}{\text{TN} + \text{FP}}$$

事例の場合には、TN799例、FP9,189名なので
特異度 = $\frac{9,189}{799+9,189} = 0.920$ (92.0%)

陽性的中率(陽性予測値)・陰性的中率(陰性予測値)

	疾患あり	疾患なし
陽性	真陽性 (TP)	偽陽性 (FP)
陰性	偽陰性 (FN)	真陰性 (TN)

陽性的中率 (TP / (TP + FP))
陰性的中率 (TN / (TN + FN))

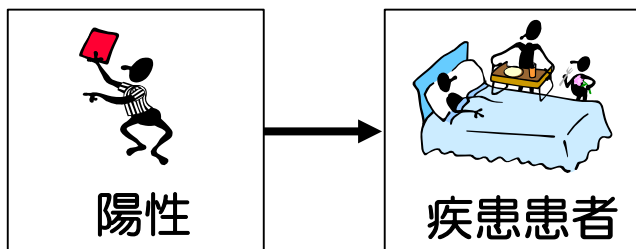
疾患の有無には有病率が関係する

陽性的中率(positive predictive value)

陽性的中率 = $TP / (TP + FP)$: 陽性の被験者のなかで疾患ありの確率

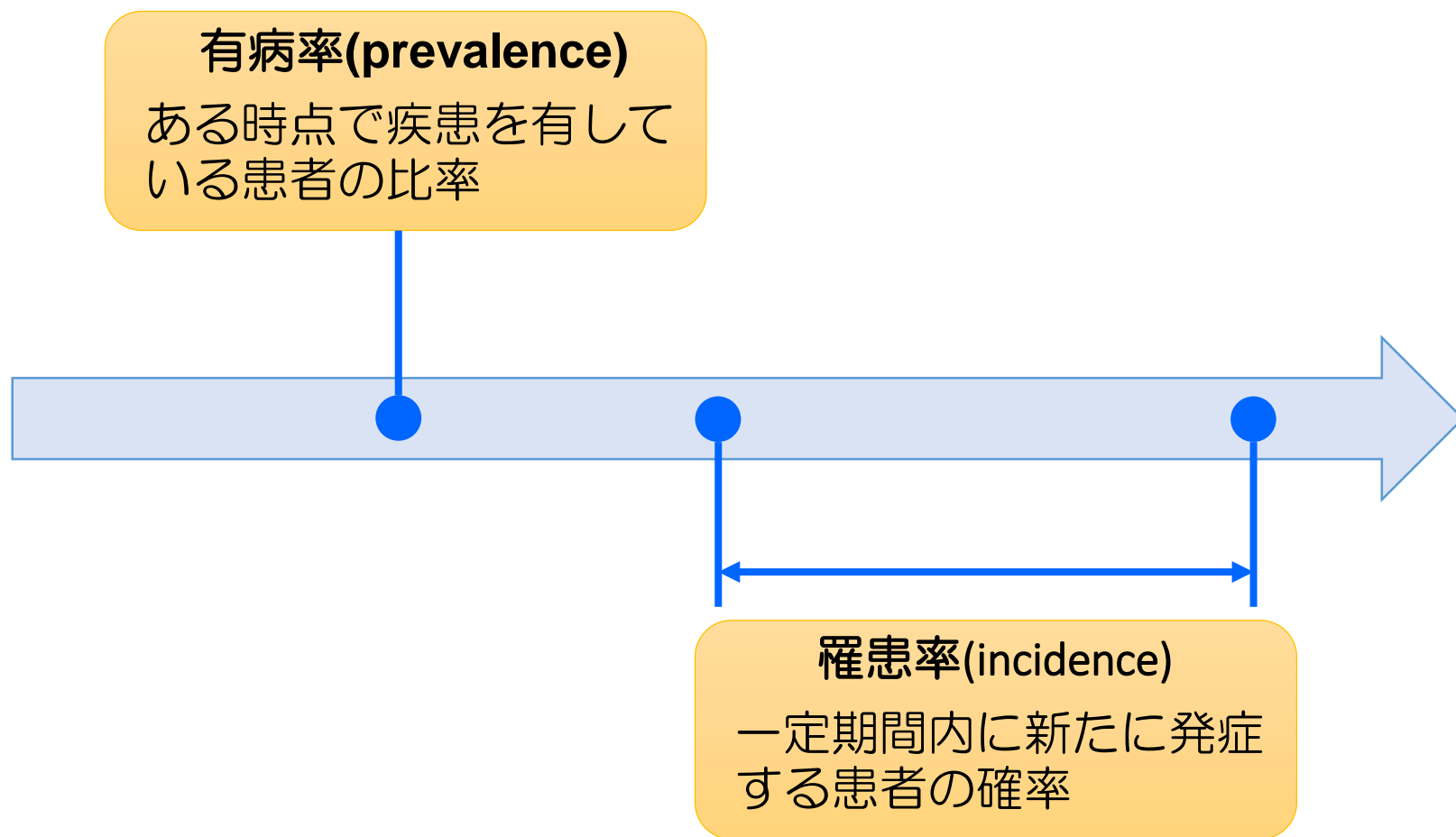
陰性的中率(negative predictive value)

陰性的中率 = $TN / (TN + FN)$: 陰性の被験者のなかで疾患なしの確率



利点：臨床的有用性は陽性的中率で評価できる。
注意点：これらの指標は、有病率の影響を受ける。

余談：有病率(prevalence)と罹患率(incidence)



陽性的中率と陰性的中率の計算

	乳癌あり	乳癌なし	合計
検査陽性	10 (TP)	799 (FP)	809
検査陰性	2 (FN)	9,189 (TN)	9,191
合計	12	9,988	10,000

	疾患あり	疾患なし
陽性	真陽性 (TP)	偽陽性 (FP)
陰性	偽陰性 (FN)	真陰性 (TN)

陽性的中率 (TP / (TP + FP))
陰性的中率 (TN / (FN + TN))

陽性的中率：陽性の被験者のなかで疾患ありの確率

$$\text{陽性的中率} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

事例の場合には、TP10例、FN2名なので
陽性的中率 = $\frac{10}{10 + 799} = 0.012$ (1.2%)

陰性的中率：陰性の被験者のなかで疾患なしの確率

$$\text{陰性的中率} = \frac{\text{TN}}{\text{FN} + \text{TN}}$$

事例の場合には、TN799例、FP9,189名なので
陰性的中率 = $\frac{9,189}{2 + 9,189} = 1.000$ (100.0%)

有病率と4種類の指標の関係

	疾患あり	疾患なし	
陽性	真陽性 (TP: True Positive)	偽陽性 (FP: False Positive)	陽性的中率 = $\frac{TP}{TP+FP}$
陰性	偽陰性 (FN: False Negative)	真陰性 (TN: True Negative)	陰性的中率 = $\frac{TN}{TN+FN}$

感度 = $\frac{TP}{TP+FN}$ 特異度 = $\frac{TN}{TN+FP}$

陽性的中率(陰性的中率)の定義

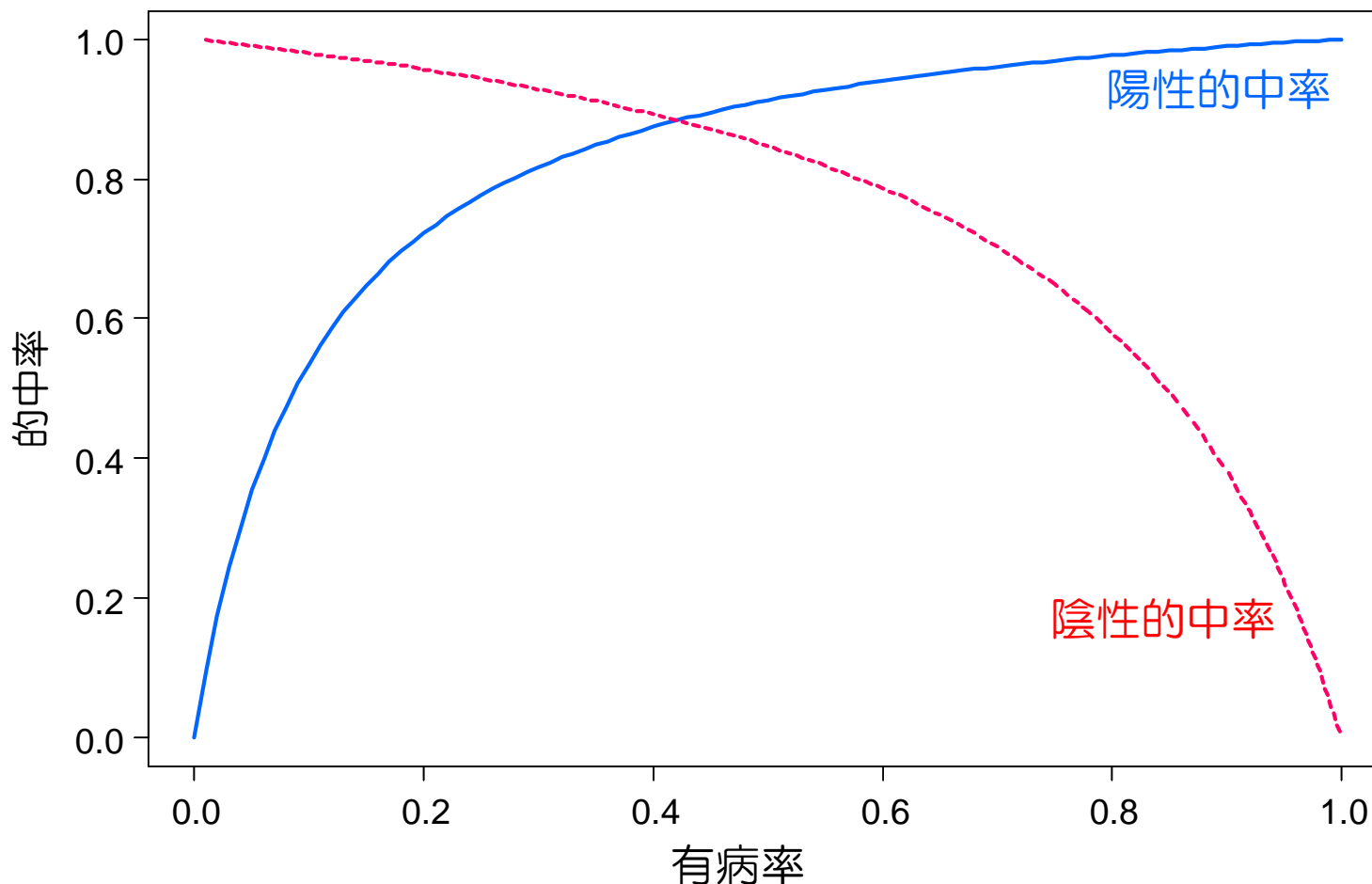
感度・特異度と陽性的中率・陰性的中率の関係は、Bayesの定理を用いて計算できる。

$$\text{陽性的中率} = \frac{\text{有病率} \times \text{感度}}{(\text{有病率} \times \text{感度}) + (1 - \text{有病率}) \times (1 - \text{特異度})}$$

「真の有病率 = データから計算した有病率」のとき、先ほどの計算式になる。

有病率と陽性的中率・陰性的中率の関係

感度 = 0.833 (83.3%)
特異度 = 0.920 (92.0%)



有病率が高くなるほど陽性的中率が高くなり、陰性的中率が低くなる。例えば、

- 有病率が0.030(3.0%)のとき：陽性的中率は0.244(24.4%)
- 有病率が0.005(0.5%)のとき：陽性的中率は0.050(5.0%)

である。

診断精度の評価尺度：尤度比

	疾患あり	疾患なし
陽性	真陽性 (TP: True Positive)	偽陽性 (FP: False Positive)
陰性	偽陰性 (FN: False Negative)	真陰性 (TN: True Negative)

感度 = $\frac{TP}{TP+FN}$

特異度 = $\frac{TN}{TN+FP}$

陽性尤度比(likelihood ratio of a positive result)

$$\text{陽性尤度比} = \text{感度} / (1 - \text{特異度})$$

「疾患有が疾患無よりも何倍陽性になりやすいか」を表す指標

疾患有を陽性と診断した場合と疾患無を陽性と診断した場合の比で表されており、大きいほど確定診断に優れるといえる(一般に尤度比といった場合には陽性尤度比を表す。)

陰性尤度比(likelihood ratio of a negative result)

$$\text{陰性尤度比} = (1 - \text{感度}) / \text{特異度}$$

「疾患有が疾患無よりも何倍陰性になりやすいか」を表す指標

疾患有を陰性と診断した場合と疾患無を陰性と診断した場合の比で表される指標である(あまり用いられない)。

尤度比の計算

	乳癌あり	乳癌なし	合計
検査陽性	10	799	809
検査陰性	2	9,189	9,191
合計	12	9,988	10,000

$$\text{感度} = \frac{10}{10+2} = 0.833 \text{ (83.3\%)}$$

$$\text{特異度} = \frac{9,189}{799+9,189} = 0.920 \text{ (92.0\%)}$$

陽性尤度比(likelihood ratio of a positive result)

$$\text{陽性尤度比} = \frac{\text{感度}}{1 - \text{特異度}} = \frac{0.833}{1 - 0.920} = 10.4$$

陰性尤度比(likelihood ratio of a negative result)

$$\text{陰性尤度比} = \frac{1 - \text{感度}}{\text{特異度}} = \frac{1 - 0.833}{0.920} = 0.18$$

二つの定性検査の類似性の評価：Kappa係数

2人の病理医による非小細胞肺癌の組織標本75枚の組織学的分類の結果(Gardis, 2009)

		病理医 A		
		Grade II	Grade III	合計
病理医 B	Grade II	41	3	3
	Grade III	4	27	27
	合計	45	30	75

Kappa係数の計算式

観測度数

病理医 A

病理医 B

	Grade II	Grade III	合計
Grade II	41	3	44
Grade III	4	27	31
合計	45	30	75

期待度数

病理医 A

病理医 B

	Grade II	Grade III	合計
Grade II	$\frac{45 \times 44}{75} = 26.4$	$\frac{30 \times 44}{75} = 17.6$	44
Grade III	$\frac{45 \times 31}{75} = 18.6$	$\frac{30 \times 31}{75} = 12.4$	31
合計	45	30	75

$$\text{測定者の全一致率} = \frac{41 + 27}{75} = 0.907$$

$$\text{偶然による一致率} = \frac{26.4 + 12.4}{75} = 0.517$$

測定者間的一致率が偶然によって期待される一致率よりどれくらい大きいか



Kappa係数=

測定者の全一致率 - 偶然による一致率

1 - 偶然による一致率



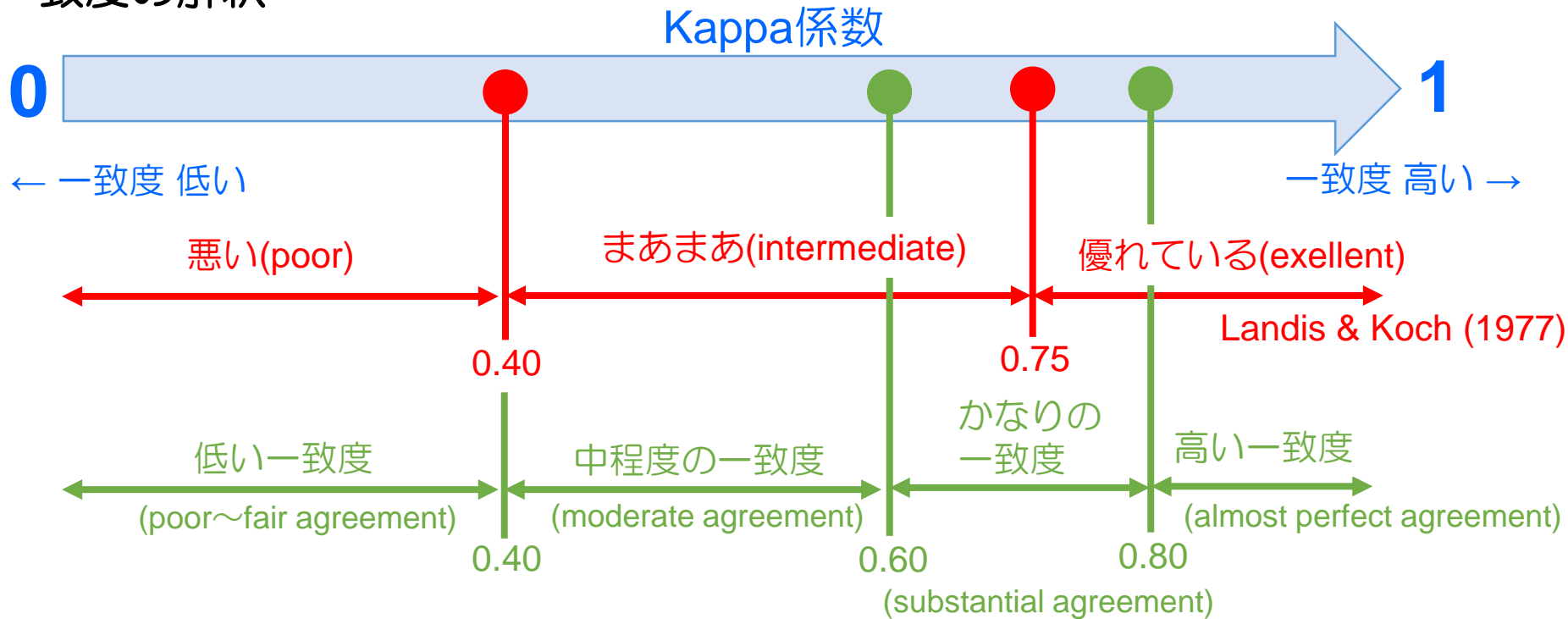
2人の測定者的一致率が偶然による一致を含めずに最大限取り得る値はいくらか

Kappa係数の計算とその結果

$$\text{Kappa係数} = \frac{\text{測定者の全一致率} - \text{偶然による一致率}}{1 - \text{偶然による一致率}}$$

$$\text{Kappa係数} = \frac{0.907 - 0.517}{1.000 - 0.517} = 0.810$$

一致度の解釈



事例では高い一致度が見られている

余談：対応があるクロス集計表

65歳以上の高齢者を対象に、転倒予防訓練と運動機能の低下の有無に関する研究が実施された。この研究では、200人の被験者に対して、転倒予防訓練前に運動機能検査を行い、3カ月の転倒予防訓練後に同様の検査を実施している(柳川他,2011)。

		歩行訓練後		合計
		低下あり	低下なし	
歩行訓練前	低下あり	40 (20.0%)	60 (30.0%)	100
	低下なし	30 (15.0%)	70 (35.0%)	100
合計		70	130	200

- (1) 訓練前，訓練後ともに運動機能が低下している割合は20%
- (2) 訓練前は運動機能が低下していたが，訓練後に低下が認められなかった割合は30%
- (3) 訓練前は運動機能の低下が認められなかったが，訓練後に低下していた割合は15%
- (4) 訓練前，訓練後ともに運動機能の低下が認められなかった割合は35%

一貫性を見るということと違いを見るということ

一貫性を見る場合

例(2種類の定性検査の一貫性評価)

2つの検査を同じ被験者に行ったばあいに、検査結果が一致するか否かを評価する研究

		検査B	
		陽性	陰性
検査A	陽性	(a)	(b)
	陰性	(c)	(d)

(a) 検査Aおよび検査Bともに陽性、(d) 検査Aおよび検査Bともに陰性の度数(被験者数)が診断結果が検査A、検査Bで診断結果が異なる(b)(c)よりも大きくなることを期待される

一致したか否かを検討する
(Kappa係数)

変化(違い)を見る場合

例(手術による不安感に関する調査)

医師からの手術に関する説明前に、手術に対する不安感(あり、なし)を質問したうえで、説明後に同じ質問を行う研究

		説明後	
		あり	なし
説明前	あり	(a)	(b)
	なし	(c)	(d)

説明前に不安ありだった患者が説明後に不安なしに変化することが期待される。

変化したか否かを評価する
(対応がある場合のオッズ比)

McNemar検定

		歩行訓練後		合計
		低下あり	低下なし	
歩行訓練前	低下あり	40 無	60 訓練後良	100
	低下なし	30 訓練前良	70 無	100
合計		70	130	200

訓練前に低下ありで訓練後に低下なしの割合 (■のセルの割合)

$$\text{割合} = \frac{60}{200} = 0.300 (30.0\%)$$

訓練前に低下なしで訓練後に低下ありの割合 (■のセルの割合)

$$\text{割合} = \frac{30}{200} = 0.150 (15.0\%)$$



二つの割合を比較することで評価するのがMcNemar検定

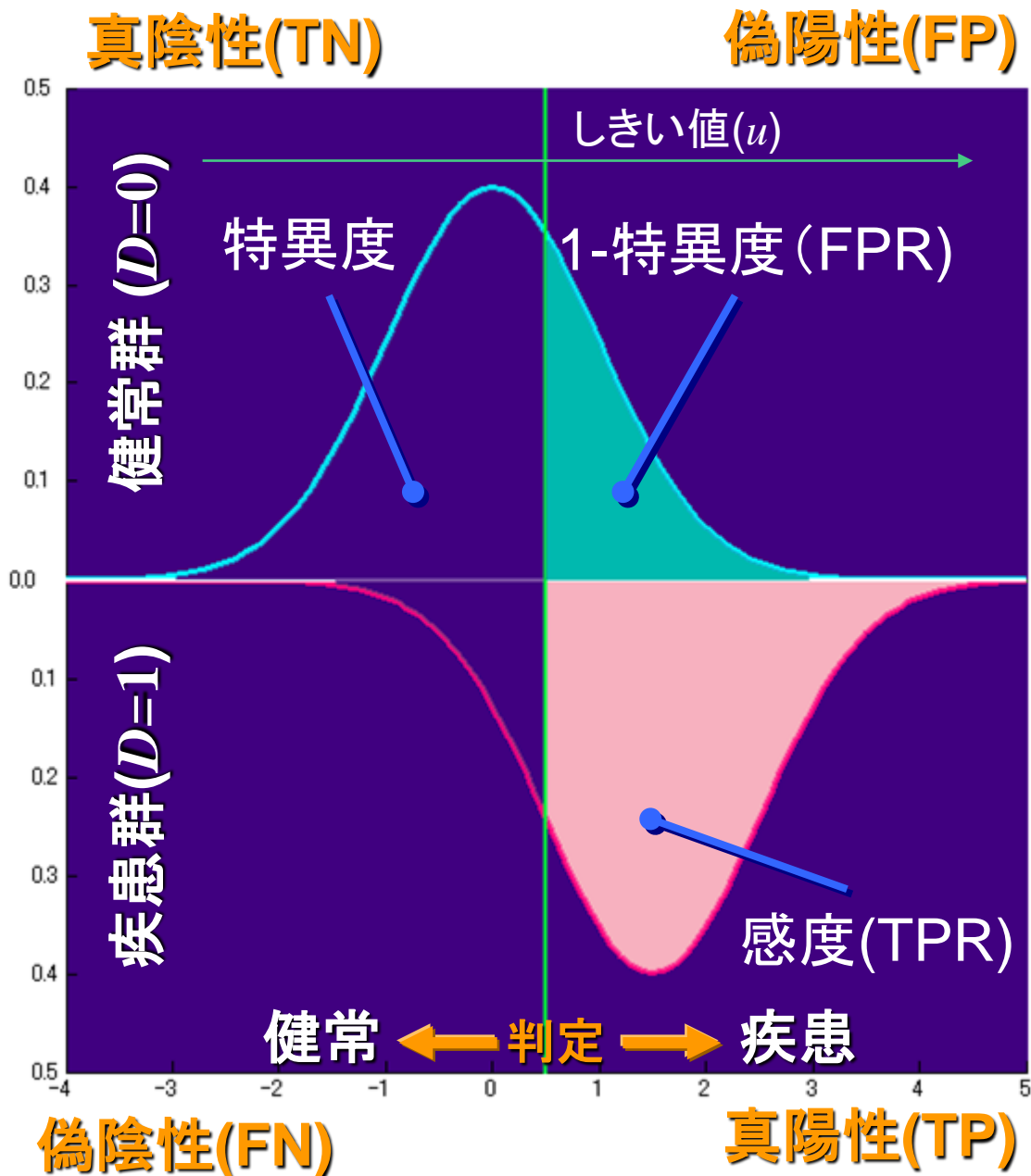
その結果、**p値は0.0016**なので、有意だった。つまり、歩行訓練前後で運動機能の低下に変化が認められた。

定量検査の評価：ROC曲線

頭部外傷症の重篤度を識別するために、CK-BB(クレアチン・キナーゼBB)が有効か否かを判定している。ここに、重篤度は、重度および非重度の2値とする。Zhou *et al.*(2002)は、CK-BBが成人での重篤度の判定だけでなく、未成人においても利用可能であるか否かを検討している。

重症群					非重症群		
140	740	543	490	523	136	60	46
1087	126	913	156	76	286	17	
230	153	230	356	303	281	27	
183	283	463	350	353	23	126	
1256	90	60	323	206	200	100	
700	303	509	1560		146	253	
16	193	576	120		220	70	
800	76	671	216		96	40	
253	1370	80	443		100	6	

定量検査値における感度・特異度



		グループ	
		疾患	健常
検査	陽性 ($\geq u$)	TP	FP
	陰性 ($< u$)	FN	TN

真陽性(TP)

疾患群において、カットオフ値(u)以上となる。

偽陽性(FP)

疾患群において、カットオフ値(u)未満となる。

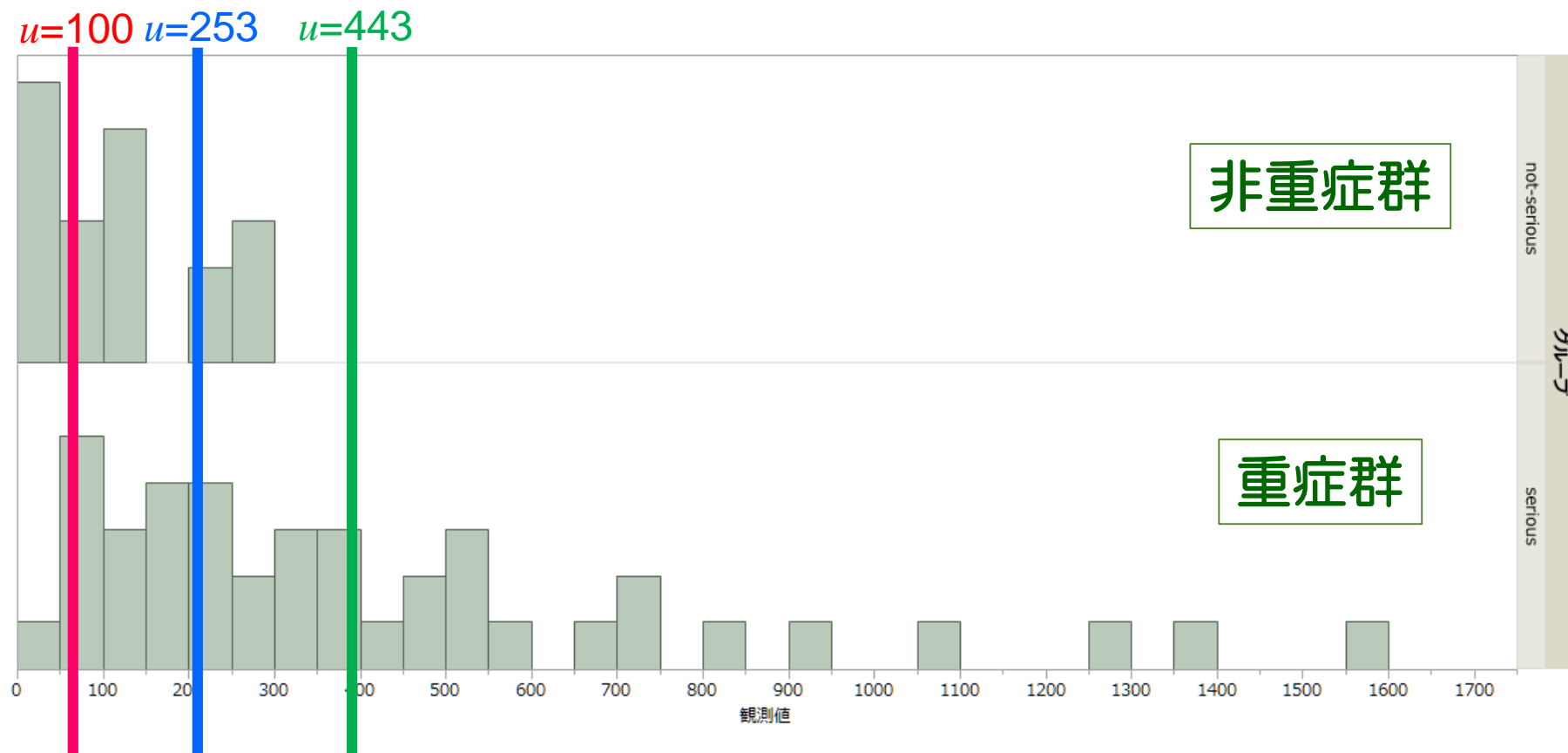
真陰性(TN)

非疾患群において、カットオフ値(u)以上となる。

偽陰性(FN)

非疾患群において、カットオフ値(u)未満となる。

カットオフ値と感度・特異度の関係 (テキストと事例が異なります)



非重症群

重症群

not-serious

serious

グループ

観測値

		グループ	
		重症	非重症
検査	陽性	35	10
	陰性	6	9

感度 = 0.854
1-特異度 = 0.526

		グループ	
		重症	非重症
検査	陽性	24	3
	陰性	17	16

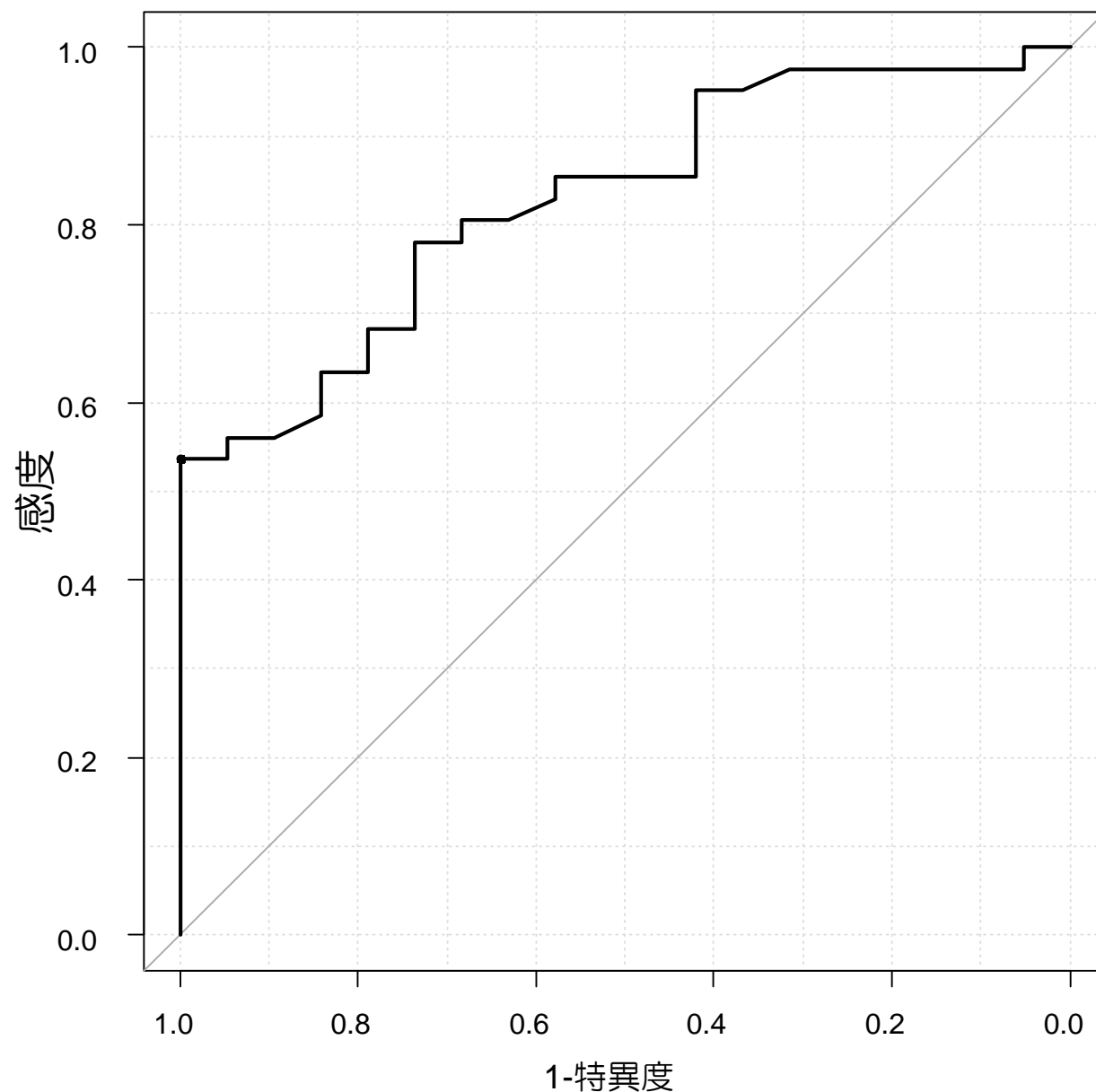
感度 = 0.774
1-特異度 = 0.158

		グループ	
		重症	非重症
検査	陽性	16	0
	陰性	25	19

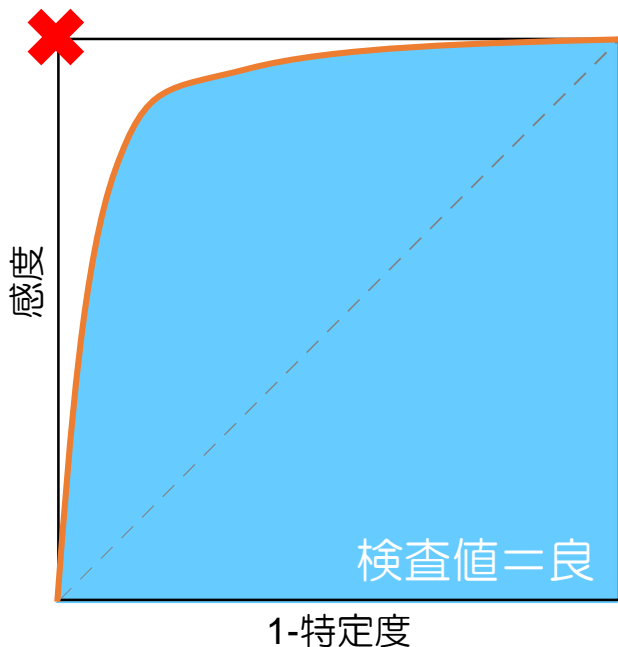
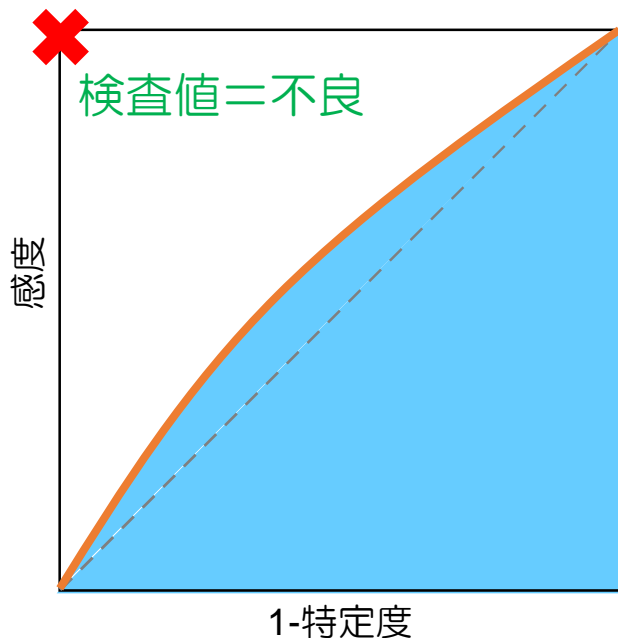
感度 = 0.390
1-特異度 = 0.000

ROC曲線の描写(テキストと事例が異なります)

	X軸	Y軸
観測値	1-特異度	感度
1560	0.000	0.024
1370	0.000	0.049
1256	0.000	0.073
1087	0.000	0.098
913	0.000	0.122
800	0.000	0.146
740	0.000	0.171
700	0.000	0.195
671	0.000	0.220
576	0.000	0.244
543	0.000	0.268
523	0.000	0.293
509	0.000	0.317
⋮	⋮	⋮
70	0.632	0.951
60	0.684	0.976
46	0.737	0.976
40	0.790	0.976
27	0.842	0.976
23	0.895	0.976
17	0.947	0.976
16	0.947	1.000
6	1.000	1.000
6	1.000	1.000



ROC曲線の解釈



ROC曲線

カットオフ値 u を小さいほうから大きいほうに動かしたときの、

(1-特異度, 感度)

をプロットしたものがROC曲線である。

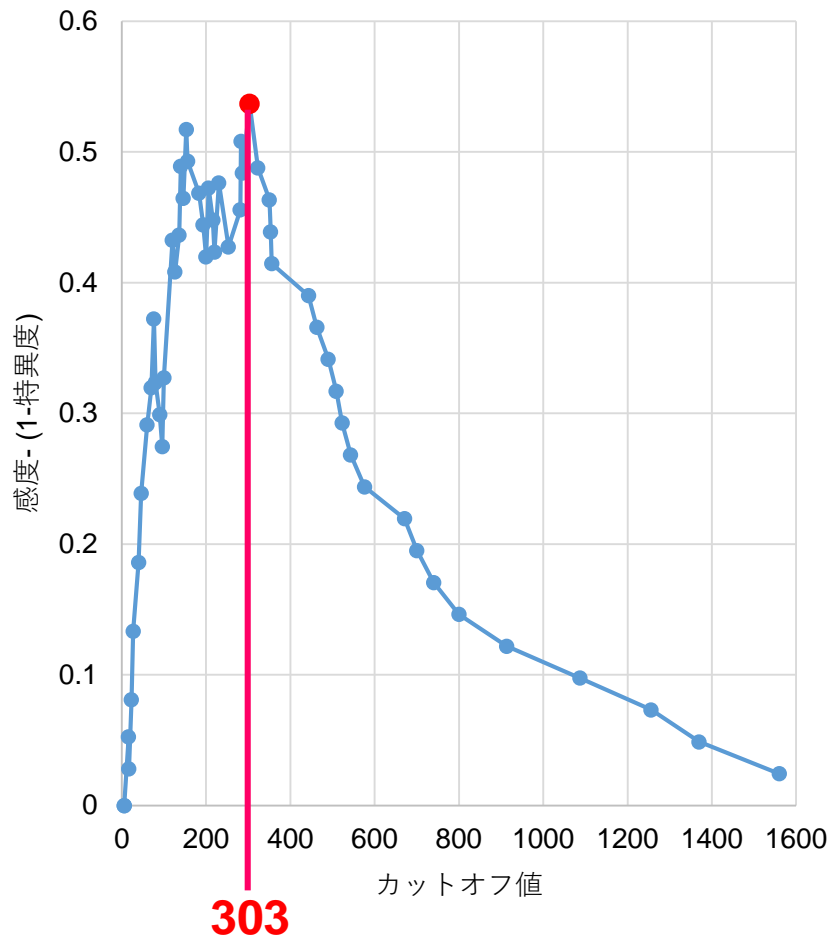
ROC曲線は、45度の直線に対して扇型の形状を示す。座標(0,1)[**X**で表された場所]に近くなるほど良好な診断であるといえる。

曲線化面積AUC

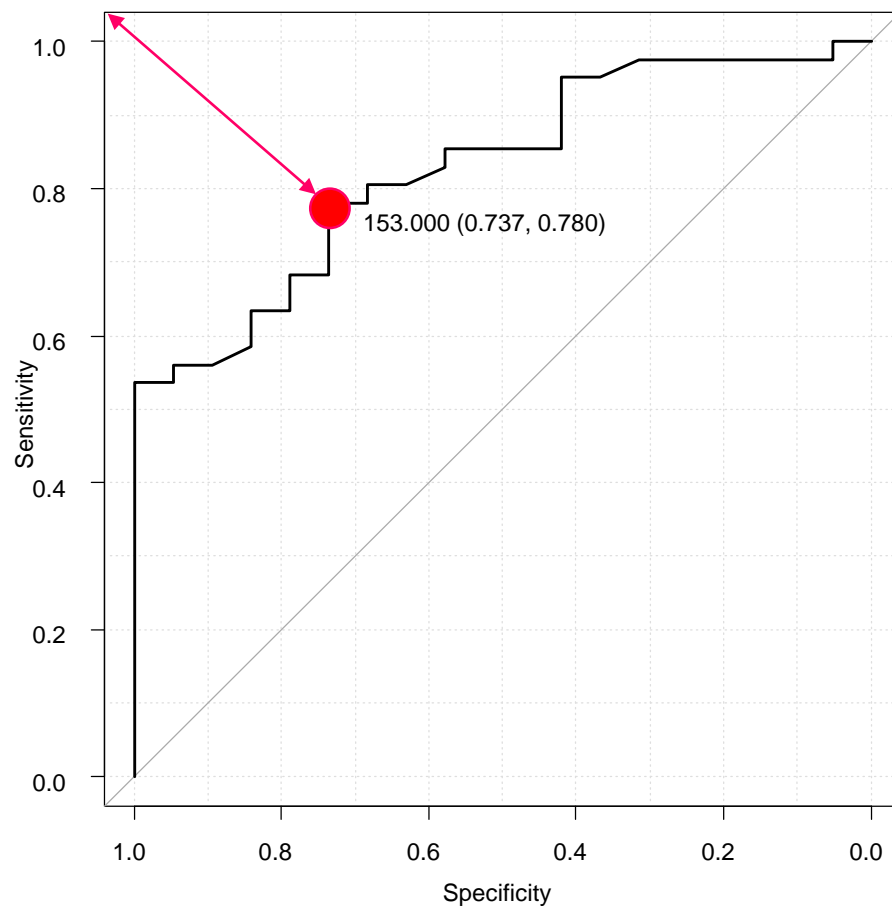
ROC曲線の曲線以下の面積である(■の面積)。複数の検査を比較したり、診断能を数値化するのに用いられる。

最適なカットオフ値の選定

感度 - (1 - 特異度)が最大のカットオフ値
(感度 + 特異度)でも同じ意味



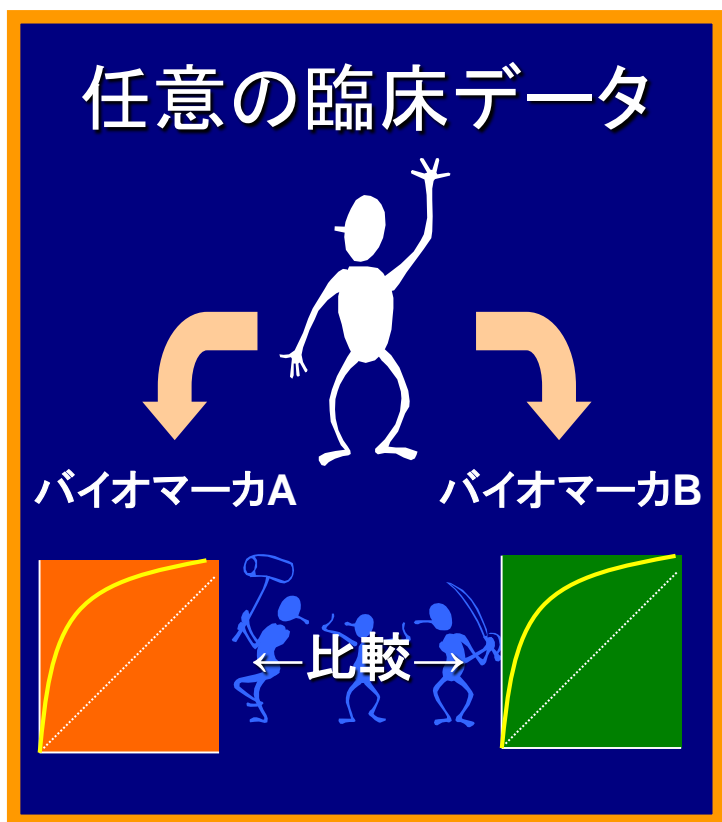
ROC曲線の中で、最も座標(0,1)に近いときのカットオフ値



ROC曲線における比較：対応がある場合とない場合

対応がある場合

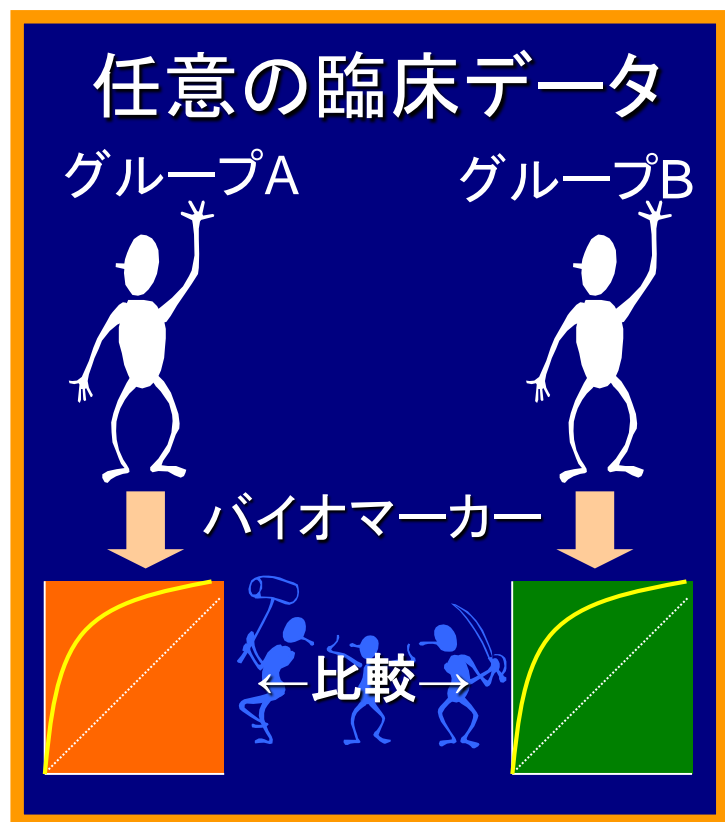
複数の検査を同一被験者を実施し、検査の診断能を比較する場合



実臨床では殆どがこちら側

対応がない場合

同一検査を実施したときに、被験者層によって診断能に違いがないか否かを評価する場合

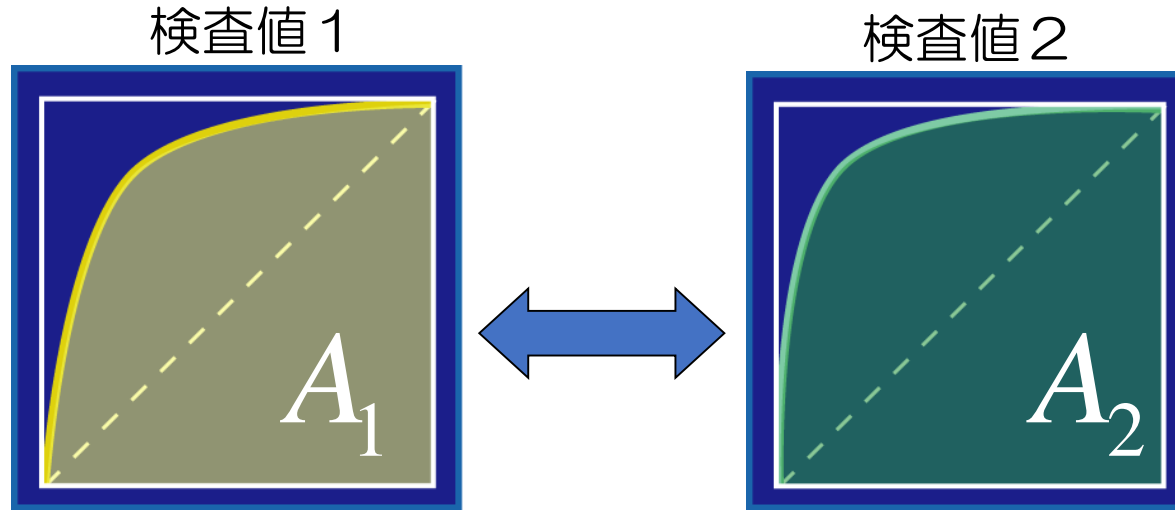


事例検討：仮想データ

これは、ある疾患に対する2種類の検査方法に関するデータである。二つの検査方法の診断能に違いがあるか (対応のあるデータ)

疾患の有無	検査方法1	検査方法2	疾患の有無	検査方法1	検査方法2
あり	59	95	なし	63	93
あり	61	123	なし	64	99
あり	55	74	なし	65	119
あり	66	145	なし	64	92
あり	52	64	なし	68	112
なし	60	84	なし	64	99
なし	61	128	なし	69	113
なし	51	79	あり	62	92
あり	60	112	あり	64	112
あり	61	107	なし	67	128
あり	56	67	なし	65	111
なし	65	98	なし	66	105
なし	63	105	なし	62	104
なし	58	95	なし	66	106
なし	59	79	あり	65	112
あり	61	81	あり	60	115
あり	62	91	なし	68	128
あり	65	142	あり	62	116
あり	63	84	なし	68	134
あり	62	85	なし	70	172

ROC曲線の曲線下面積に対する検定(Delong, 1988)



帰無仮説：2つの検査値のROC曲線の曲線下面積は等しい($H_0 : A_1 = A_2$)

対立仮説：2つの検査値のROC曲線の曲線下面積は等しくない($H_1 : A_1 \neq A_2$)

ROC曲線の曲線下面積を比較することで、2個の検査値の診断能に関する評価ができる。

その他に、

- それぞれのカットオフ値における比較 (カットオフ値での診断能)
- ロジスティック回帰における尤度比検定

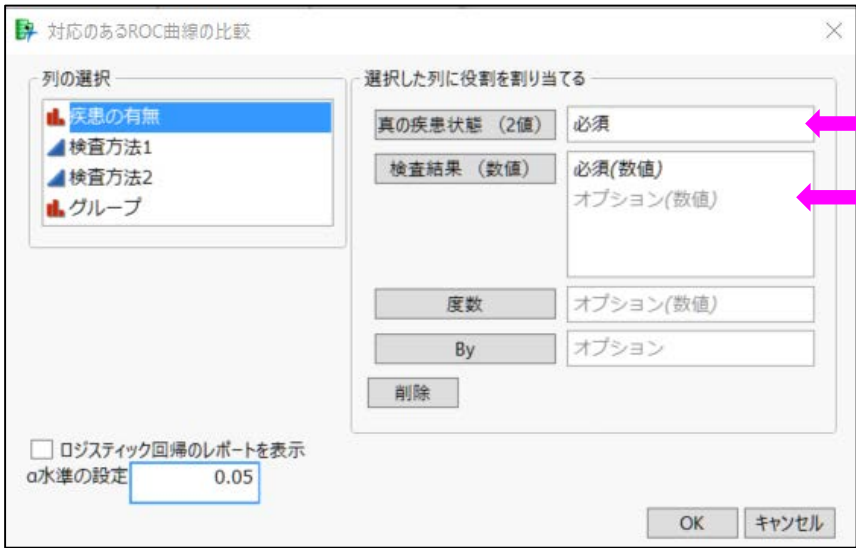
などが考えられるが、いずれも検査値の診断能というわけではなく、前者はカットオフ値に依存し、後者は診断能ではなく当てはまりの良さを評価している。

JMPによる対応のあるデータでのROC曲線のAUCの比較

ROC曲線のAUCを比較するには、JMPの外部スクリプトを利用する必要がある。外部スクリプトは、臨床研究センターで管理しているため、要望がある場合は、連絡をお願いします。

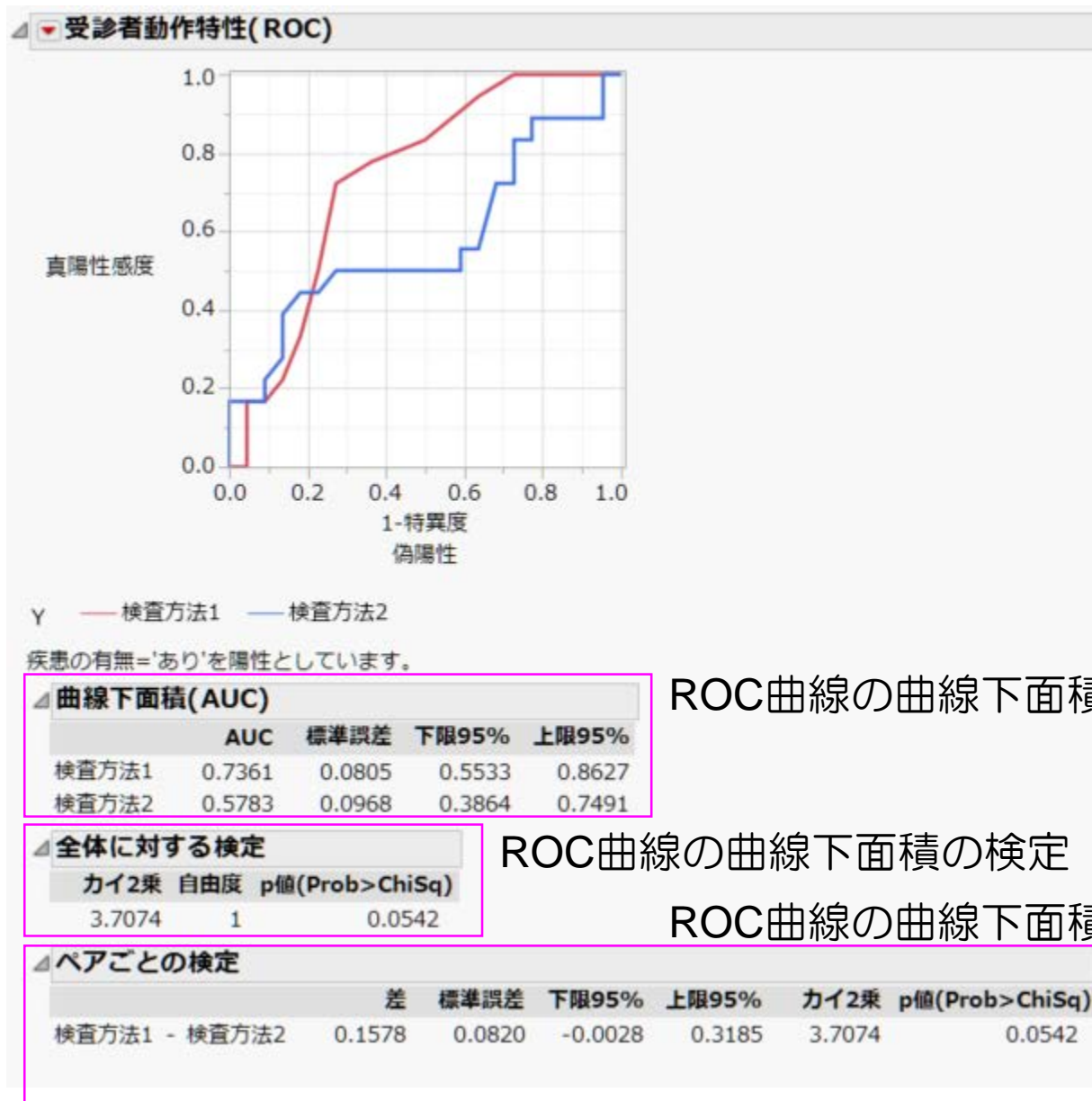
JMPの手順

- STEP.1** データファイルを読み込む
- STEP.2** 「ファイル」→「対応のあるROC曲線の比較」を選択し、「開く」の横の▼を押して実行を選ぶ。



疾患の有無を表す変数
検査値(2個以上)

JMPでの結果



ROC曲線の曲線下面積及び95%CI

ROC曲線の曲線下面積の検定

ROC曲線の曲線下面積の差の95%CI

Thank you for your kind attention

