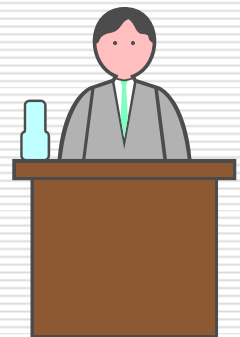


臨床研究デザインと医用統計の実践法入門①

岩手医科大学 消化器・肝臓内科

遠藤龍人

ryuendo@iwate-med.ac.jp





今日お話すること

- ◎ 臨床研究の現状と課題
 - 医学研究論文の信頼性を高めることを目指す新しい取り組み
 - STARDイニシアチブ(診断研究の報告)
- ◎ リサーチクエスチョンから臨床研究デザインへ
- ◎ 医学検査の有用性を評価する時のポイント
 - 信頼性評価: 診断診察データの一致率とkappa値
 - 妥当性評価: 感度、特異度、ROC曲線
- ◎ サンプルサイズの設定
 - 感度分析

日本における臨床研究の現状と課題

- ② 基礎研究に比べて臨床研究の貢献度が少ない
 - エビデンスレベルの高い論文、引用回数の多い論文が少ない。
- ② 研究者へのインセンティブの低さ
 - 時間、人手がかかる
 - 論文数を稼げない、大学の中で評価されない
- ② 研究費獲得が困難
- ② システム(教育、実践)の問題
 - 臨床疫学、医療統計学、EBMの教育が手薄
 - 研究者をサポートする体制の欠如

臨床研究志向のMD研究者は “絶滅危惧種”の恐れ

New Physician-Investigators Receiving National Institutes of Health Research Project Grants A Historical Perspective on the “Endangered Species”

Howard B. Dickler, MD

Di Fang, PhD

Stephen J. Heinig, MA

Elizabeth Johnson, MS

David Korn, MD

DECLINES IN THE NUMBER OF physician-scientist applicants and recipients of National Institutes of Health (NIH) research and training awards in the 1970s generated concerns that physician clinical investigators would become an “endangered species” if trends continued unaltered.¹ In succeeding decades, as unprecedented scientific and technological progress in biomedicine and related disciplines moved discoveries ever closer to clinical relevance, concerns continued about the adequacy of the physician-scientist workforce. Consistently expressed concerns include changing market forces adversely affecting the environment for clinical scholarship in medical schools and teaching hospitals, and clinical research faring less well than laboratory research in the NIH’s well-established review and award processes for grants.

Since the 1970s, eminent panels and authors have examined the plight of phy-

Context Although concerns have persisted for decades about the production of new physician clinical scientists and their success in receiving and sustaining research supported by the National Institutes of Health (NIH), no comprehensive analysis documents the experiences of first-time investigators with an MD over a long period.

Objective To ascertain the perseverance and comparative success of physician-scientists competing for NIH research (R01) grants awarded over 40 years.

Design, Setting, and Participants A longitudinal, comparative study of all first-time applicants and recipients of NIH R01 grants between 1964 and 2004 stratified by the principal investigators’ major degrees (MD, PhD, or MD and PhD) and their proposed involvement in research of humans or human tissues.

Main Outcome Measures Number of first- and second-time NIH R01 grant applicants and recipients by academic degree and by research type (clinical vs nonclinical).

Results The annual number of first-time investigators with an MD only as NIH R01 grant applicants remained remarkably stable over 4 decades (41-year mean of 707 [range, 537-983] applicants). Among first-time applicants, those with an MD consistently had less success in obtaining funding (mean annual percentage [MAP], 28%) than either investigators with a PhD (MAP, 31%; $P = .03$ vs MD only) or both an MD and a PhD (MAP, 34%; $P < .001$ vs MD only and $P = .002$ vs PhD only). Among investigators who obtained a first R01 grant, those with an MD were consistently less likely (MAP, 70%) than those with a PhD (MAP, 73%; $P = .04$ vs MD only) or those with an MD and a PhD (MAP, 78%; $P < .001$ vs MD only and $P = .007$ vs PhD only) to obtain a subsequent R01 grant. First-time applicants with an MD were much more likely to propose clinical research (MAP, 67%) than applicants with an MD and a PhD (MAP, 43%) and applicants with a PhD only (39%). First-time applicants with an MD only who proposed clinical research were funded at lower rates than their MD-only counterparts proposing nonclinical research (23% vs 29%, respectively; $P < .001$).

Conclusions From 1964-2004, the number of physician-investigators applying for first R01 grants showed little net change. Physician-investigators consistently experienced higher rates of attrition and failure, even after receiving a first R01 grant, and those proposing clinical research were less successful in obtaining funding than physicians proposing nonclinical research.

JAMA. 2007;297:2496-2501

www.jama.com

臨床研究についての誤解

- ② ヒトのサンプルを使った研究のことである
- ② アイディア勝負である
- ② 終わってから統計家に相談すれば良い
 - サンプルサイズ(n)が大きければ良い
 - “有意差 * : $P < 0.05$ の呪縛”
 - まるで”死人を蘇らせる”ような神頼み的存在？
- ② RCT以外のスタディは意味がない

7 Wrong Ways in Clinical Research (7つの御法度)

- ② データをとってから研究デザインを考える(泥縄的)
- ② リサーチクエスチョンが明確・具体的でない
- ② 対象が不明確. 抽出法, 参入・除外基準を設定せず
- ② 主要なアウトカム変数を設定しない. 変数の吟味なし
- ② 変数の測定方法の信頼性と妥当性を検討しない
- ② 解析計画を事前に作成しない. サンプルサイズ, 検出力, effect sizeを事前に設定しない.
- ② 結果の解釈: 統計的有意差のみで, 臨床的・社会的に意味がある差かどうかを検討せず.

Opposite!

Design first, Analysis last

臨床研究の報告の仕方に関する現状

- ① 報告すべき内容と実際の文献で公開されている内容に大きな隔たりがある
- ② 読者、査読者、編集者がその妥当性を判断できない
- ③ 必要項目がもれなく記載されるような指針が必要

医学研究論文の信頼性を高めることを目指す新しい取り組み

- ④ 生物医学雑誌への統一投稿規定(国際医学雑誌編集者委員会 ICMJE)
- ④ CONSORT声明(RCTの報告)
- ④ QUOROM声明(RCTに関するメタ・アナリシスの報告)
- ④ STARDイニシアチブ(診断研究の報告)
- ④ オタワ声明、ICMJE提案(臨床試験の国際的登録)
- ④ STROBE声明(観察研究の報告)

(Master of Clinical Research: MCR)

アドミッションポリシー

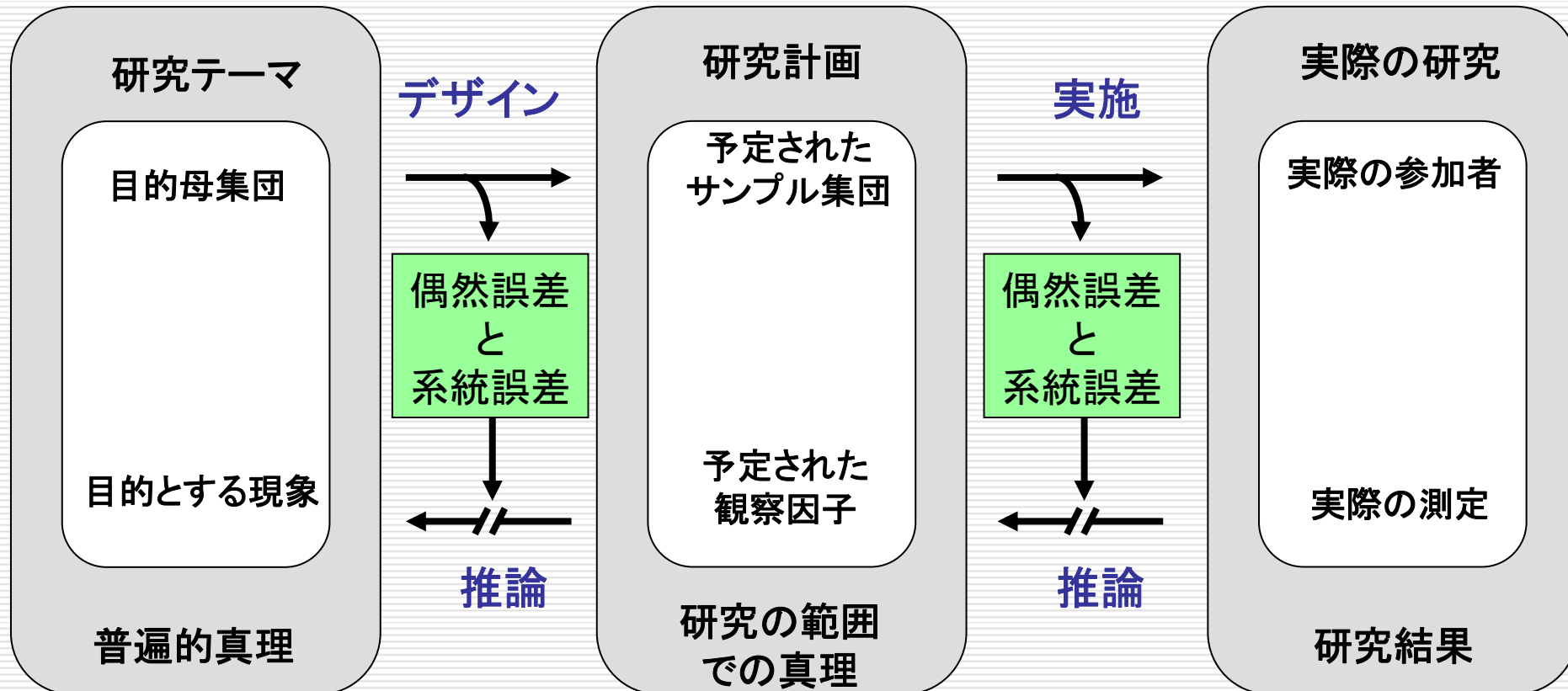
- 医師・歯科医師を対象とした一年制のコース
- ヒトや集団を単位とした臨床研究を行う研究者の育成

学習目標

- 臨床研究の基礎理論、知識、実践スキルを修得する
- 臨床上の疑問に基づいた臨床研究を計画し、プロトコール作成、実施、マネジメント基本的解析処理、解釈、論文作成などの一連の作業を独力行えるようにする

研究のダイナミズム (研究デザインが大切な理由)

Hulley S



研究テーマに対する正しい解答が得られるかどうかは、研究デザインや実施の段階で、推論の妨げとなる誤差の混入をいかにうまくコントロールするかにかかっている。

判断の誤り

「ない傾向をあるとする誤り」 = α エラー (P値、有意水準)

「ある傾向をないとする誤り」 = β エラー

真の世界

差なし

差あり

研究者の結論

差なし

(帰無仮説を受理)

正しい判断

誤った判断

β エラー

(第二種の過誤)

誤った判断

α エラー

(第一種の過誤)

正しい判断

検出力 ($1 - \beta$)

差あり

(帰無仮説を棄却)

研究プロトコル開発 (臨床研究実施計画書)

臨床経験を含む関連データの系統的レビュー



最善のアイディアと妥当な仮説設定



プロトコル作成

- 選択基準(診断基準・除外基準)
- エンドポイントの妥当性・信頼性・客観性
- 適切な統計的デザイン(サンプルサイズ・対照)
- 危険回避義務(リスクマネジメント・中止基準)
- 中間解析、評価

プロトコルの3つの基本要素

どんな**患者**に対して

どんな**検査**
治療を

どうやって**評価**するのか

(〇〇と比較して)

要求されること

- ・明確であること
- ・互いに整合性があること

RQの構造化から研究デザインへ

リサーチ・クエスチョンの構造化(定式化)

Patient	: 誰に(<u>対象</u>)
Exposure	: 何をすると(<u>介入・要因</u>)
Comparison	: 何と <u>比較</u> して
Outcome	: どうなる? (<u>効果</u>)

カテゴリー

診断法・要因

治療

予後



研究デザイン

横断研究

症例対照研究

コホート研究、RCT

臨床研究の種類

1. 記述疫学研究
(病気頻度、分布、診療パターン、自然歴)
2. 要因と害(or 益)との関連を分析する研究
3. 治療・予防の有効性・安全性の評価
4. 診断法の評価

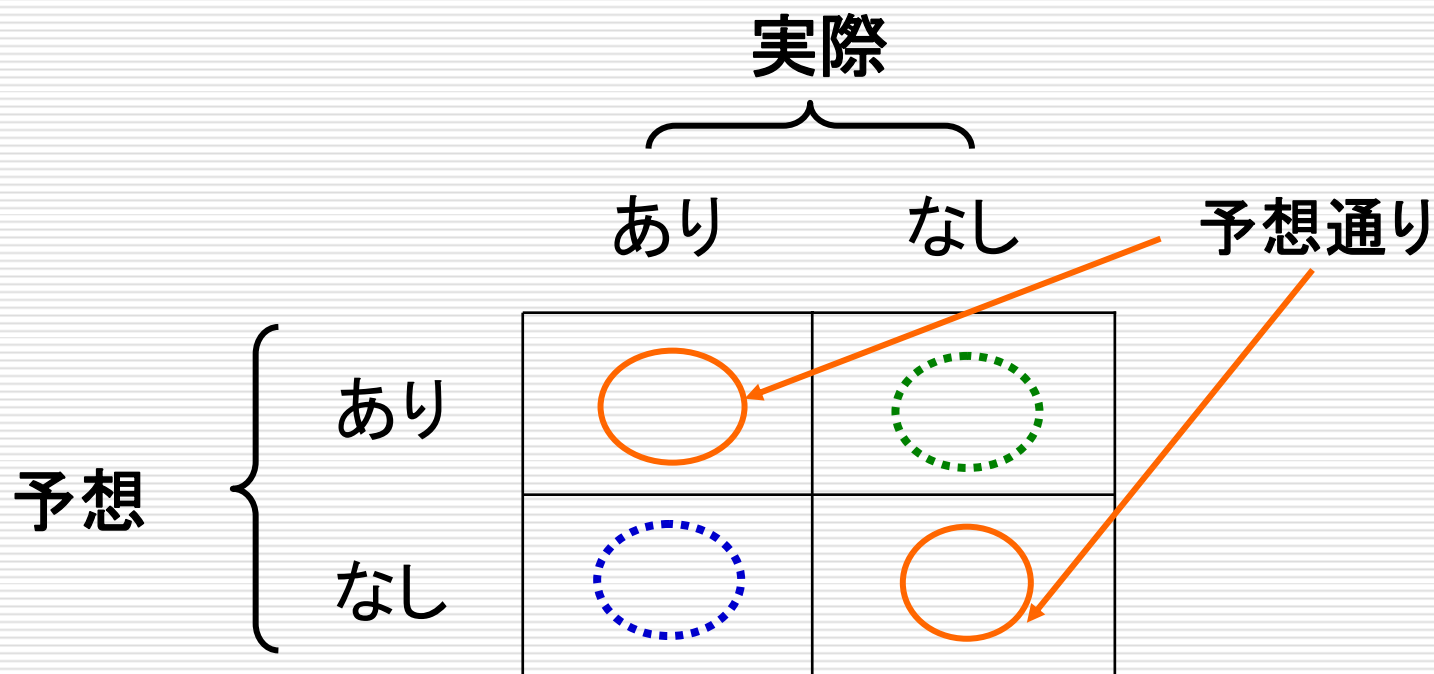
現象には4通りのとらえ方がある

- ㊦ 実際にそうであり、そのように見える
- ㊦ 実際にそうではなく、そのようには見えない
- ㊦ 実際はそうでないのに、そのように見える
- ㊦ 実際にはそうであるのに、そのようには見えない

これらすべての事象を視野に入れ、
適格に見極めるのが賢者の仕事である

エピクテトス、紀元前2世紀

現象の4つのとらえ方



医学検査の有用性を評価する時のポイント

- ④ 検査の精度 precision (再現性 reproducibility)
- ④ 検査の真度 (正確性) accuracy (妥当性 validity)
- ④ 臨床判断に及ぼす効果
- ④ 臨床的アウトカムに及ぼす効果

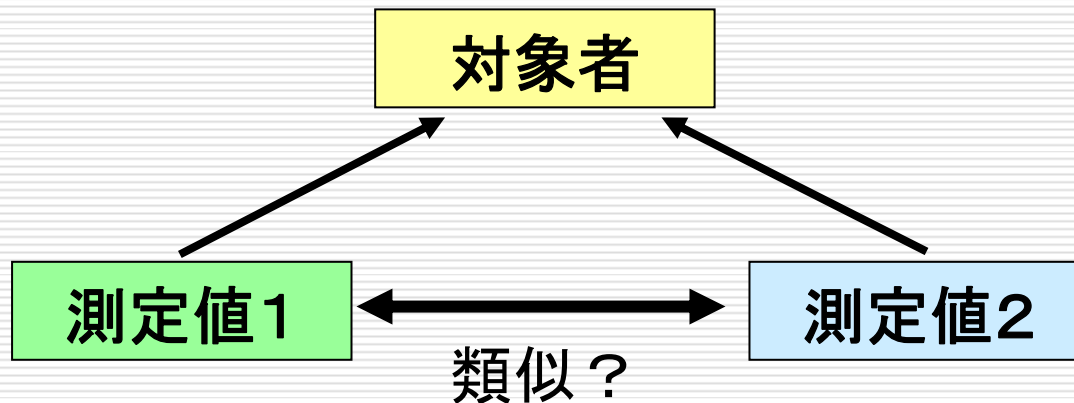
一般的には・・・、
既存の診断に比べて正しく診断できるか、より安全か、安価かという
点から、アウトカムへの効果を間接的に評価する

医学検査の有用性を評価する時のポイントと研究デザイン

ポイント	研究デザイン
検査の再現性	観察者内・観察者間変動および施設間変動の検討
検査の真度	横断研究、ケースコントロール研究、コホート研究。結果を <u>ゴールドスタンダードと比較</u>
臨床判断に及ぼす効果	実効研究 diagnostic yield study、検査前後の臨床判断に及ぼす影響に関する研究
臨床的アウトカムに及ぼす効果	ランダム化臨床試験、コホート研究、ケースコントロール研究。予測因子は検査の実施testingで、アウトカムは死亡、罹病、罹病・治療に伴うコスト

信頼性 Reliability とは？

- 当該測定・評価法を用いて、1人の対象者（または対象物）に対して繰り返し測定・評価を行った場合の、**測定値（評価値）の類似性の程度**
- ≡「再現性」: 同一評価者の場合に用いる
(異なる評価者の場合には使用しない)



信頼性 Reliability とは

<同じモノを測っているか？>

測定値が偶然に左右されず安定である度合い

- 定度(精度)に関する概念
- 測定結果が安定である度合い
- 研究では、ゴールドスタンダードを必要としない

「検査の再現性」を検討するための研究デザインと統計学的指標

研究デザイン

観察者(評価者)内・観察者(評価者)間変動および施設間変動の検討

統計学的指標

一致率、 κ 係数、変動係数、差の平均値と分散
(相関係数は避けること)

信頼性 Reliability

- 測定者間変動 (Inter-rater reliability)
 - 複数の審査員が同じことを判定して一致を見る
 - Kappa statisticsやintra-class correlation
 - 通常はこちらのやり方をする

- 測定者内変動 (Intra-rater reliability)
 - 同一の審査員が同じ判定を繰り返す
 - あまり使われない

Kappa stastics

カテゴリーデータ変数の場合の解析



CTとMRIは

どの程度結果が一致していると言えるか？

		MRI		
		(+)	(-)	
CT	(+)	10	2	12
	(-)	10	18	28
		20	20	40

カテゴリー変数解析法

		MRI		
		(+)	(-)	
CT	(+)	10	2	12
	(-)	10	18	28
		20	20	40

⇒ **一致率**: 40検査中、28例(70%)が一致している

問題: 全く偶然(ばらばらだったら)何%の確率で一致するか?

2者の判断の一致度

		MRI		
		(+)	(-)	
CT	(+)	10	2	12
	(-)	10	18	28
		20	20	40

- MRIが(+)となるのは0.5、CTが(+)となるのは0.3
∴ (+) = (+)で確率的に一致する確率は $0.5 \times 0.3 = 0.15$
- MRIが(-)となるのは0.5、CTが(-)となるのは0.7
∴ (-) = (-)で確率的に一致する確率は $0.5 \times 0.7 = 0.35$
- ◆両方を足すと、偶然の一致は0.5(50%)で起こる

Kappa統計量

偶然の一致を考慮した上での一致度の目安

$$\text{Kappa} = \frac{P(\text{実際に一致}) - P(\text{偶然で一致})}{1 - P(\text{偶然で一致})}$$

この例に当てはめると・・・

$$\text{Kappa} = \frac{0.7 - 0.5}{1 - 0.5} = 0.4$$

※殆どは0～1であるが、マイナスの値になることもある
-1 (完全不一致)、0 (一致度が偶然)、+1 (完全一致)

このCTとMRIの例では、一致度はいまひとつ

Kappaの評価

あくまでも目安(なるべく使わない方が・・・)

< 0.4	低一致率 (poor)
$0.4 \sim 0.6$	中程度の一一致率 (mild)
$0.6 \sim 0.8$	かなりの一一致率 (moderate)
> 0.8	高度の一一致率 (excellent)

Kappaにも弱点がある

- 分布に大きく影響される
 - 偏りが大きいと小さい値になる
- 一致度が高いのに低い値になることがある
- たくさんのレベルを組み替えると、そのカット値に値が影響される

Kappaの弱点

どちらも90%一致

		検査1	
		+	-
検査2	+	450	50
	-	50	450

Kappa=0.80

		検査1	
		+	-
検査2	+	20	90
	-	10	880

Kappa=0.25

- どちらも90%一致なのに、右のKappaは小さい
- 偏りに左右される



Kappaのまとめ

- あくまでも目安
- 分布の異なる群を、「こちらは一致が良い、悪い」と比べるのには使えない
(使わない方が無難です)
- 一致度と一緒に報告するのが安全

Bland-Altman plot

連続変数の場合の解析

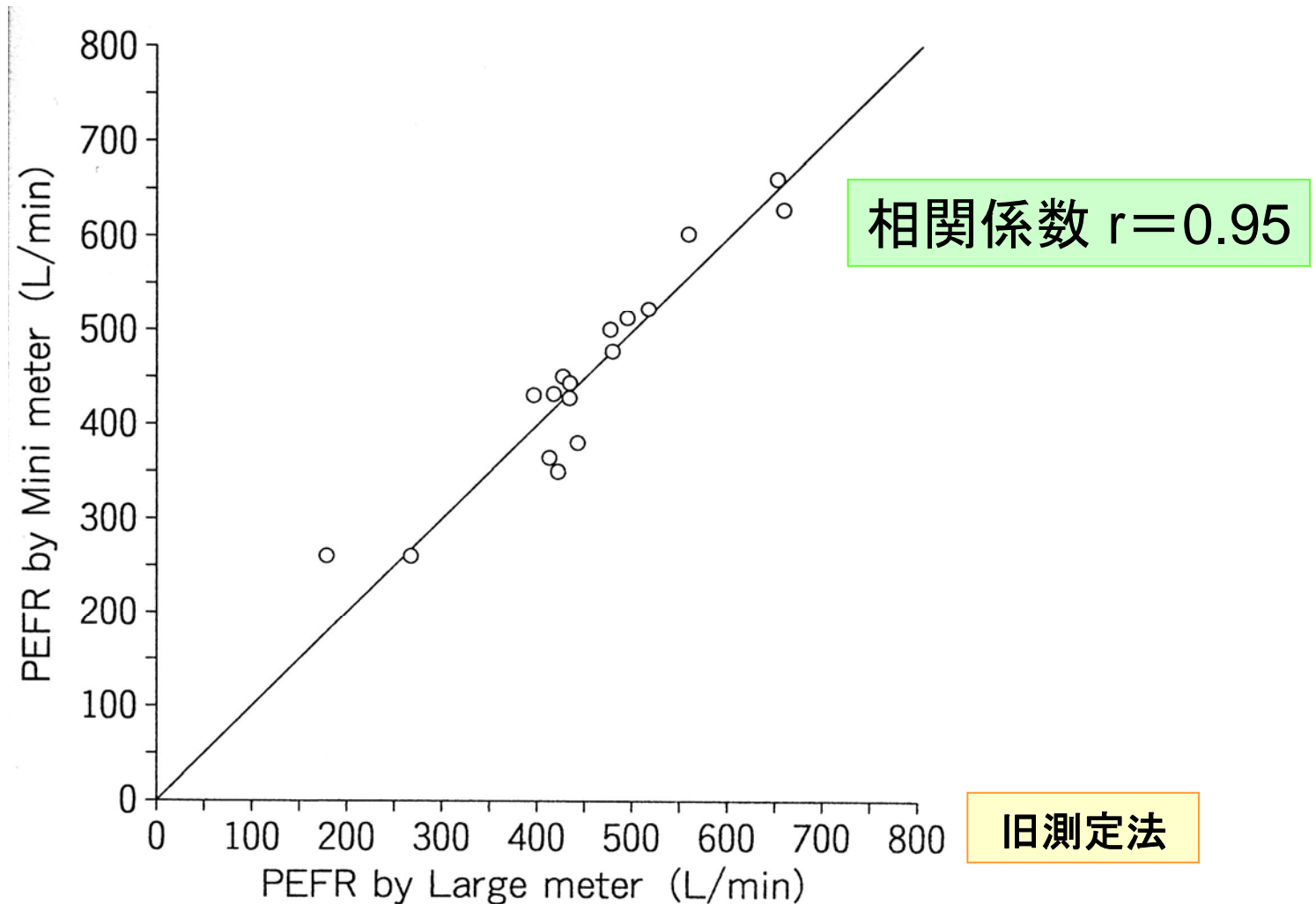
ピットフォール

■ 例

新しい簡易型装置と従来の標準型装置の両方で呼気流量を測定したところ、相関係数 $r=0.95$ と高くかつ有意な相関が得られたので、簡易型測定装置は十分に有用であると報告した

2法の測定値の分布図

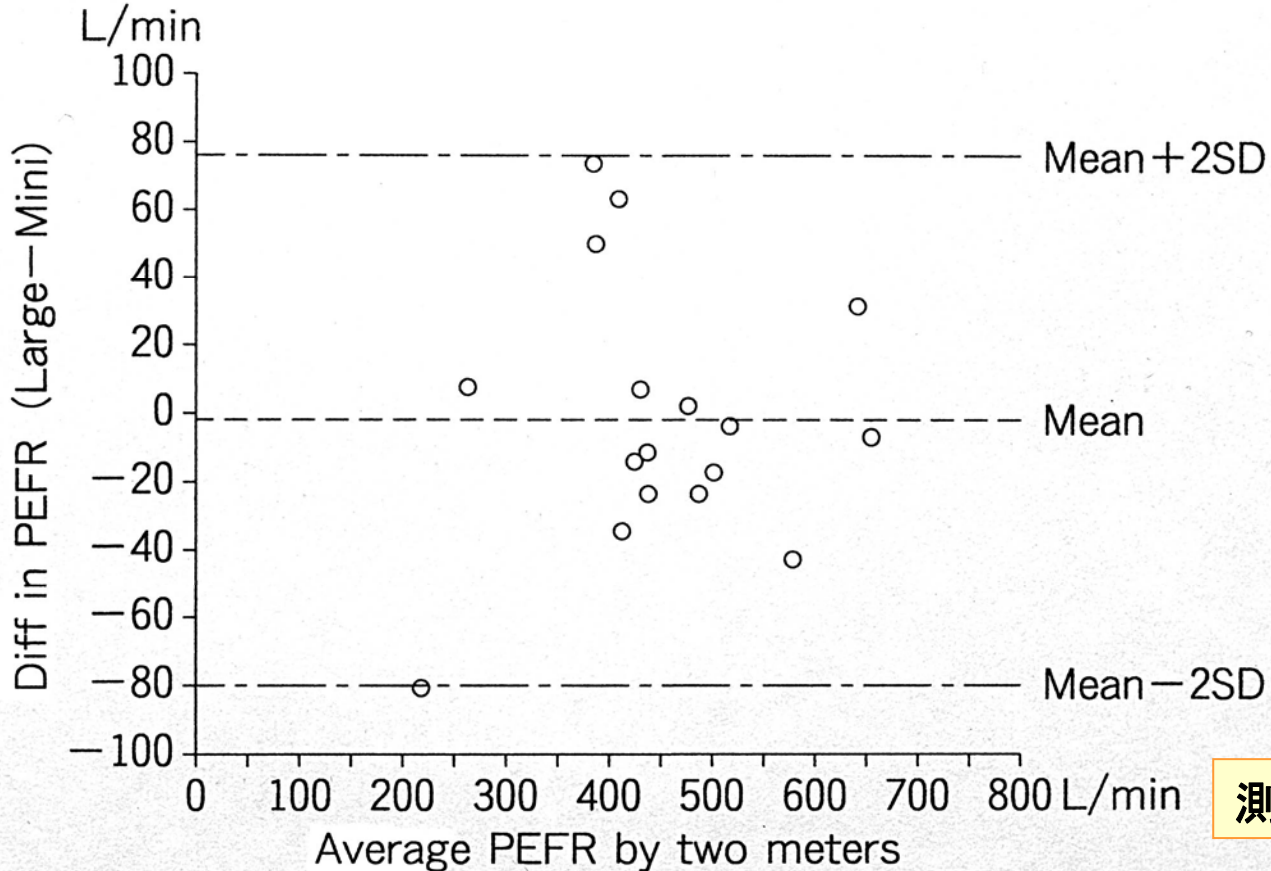
新測定法



旧測定法

Bland-Altman plot

2つの検査法間あるいは機器間の一致度



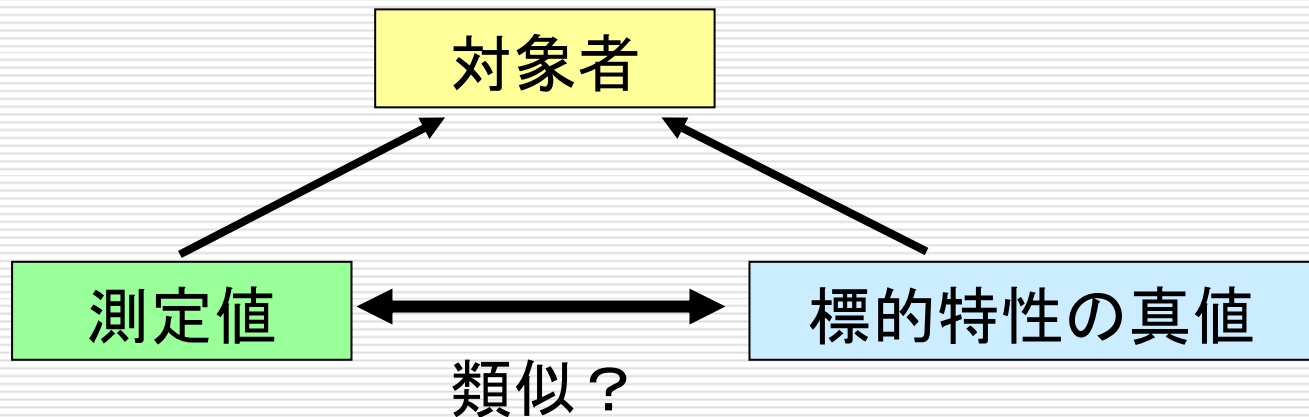
測定値間の差

測定間の平均

・測定値間の差と平均値の分布を示す

妥当性とは？

- 当該測定・評価法による測定・評価値と測ろうとしているもの(標的特性)との間の類似性
- 測りたいものを測っているかどうか？



妥当性 Validity とは

＜測りたいものを測っているか？＞
測定値が真の値と一致する度合い

- 「その検査がどの程度正しい結果を与えるか？」
- ゴールドスタンダードが必要

妥当性の評価

- 既に真の値を測定できることがわかっている確立した方法で測定された結果 = 標準値 (至適基準) Gold Standard と比較することで評価する

検査法の妥当性の例

検査法	診断	Gold Standard
カフ血圧計	血圧	動脈内カテーテル
運動負荷心電図	冠状動脈狭窄症	心カテーテル検査

「検査の真度」を検討するための研究デザイン と統計学的指標

研究デザイン

横断研究、ケースコントロール研究、コホート研究。
結果をゴールドスタンダードと比較する

統計学的指標

感度、特異度、陽性・陰性予測力、ROC曲線、尤度比

「診断検査研究」のデザインにおける留意点

- 確証バイアス verification bias (精査バイアス work-up bias)
 - 所見によってゴールドスタンダード検査を受ける人を決める
⇒ゴールドスタンダード検査を受ける基準の中に検証対象の検査や所見を含まないように、連続サンプルで評価する
- 二重ゴールドスタンダードバイアス double gold standard bias
 - 検査が陽性と陰性の場合とで異なるゴールドスタンダードを適用する(ゴールドスタンダードが侵襲的検査の時など)
 - ゴールドスタンダード間の一致性が問題になる

「検査の真度」を検討するための研究デザインの留意点

■ アウトカム評価

- 疾患の存在: ゴールドスタンダードの測定者のマスク化(盲験化)
- 予後判定研究: 治療決定権のない医師による判断

■ サンプリング

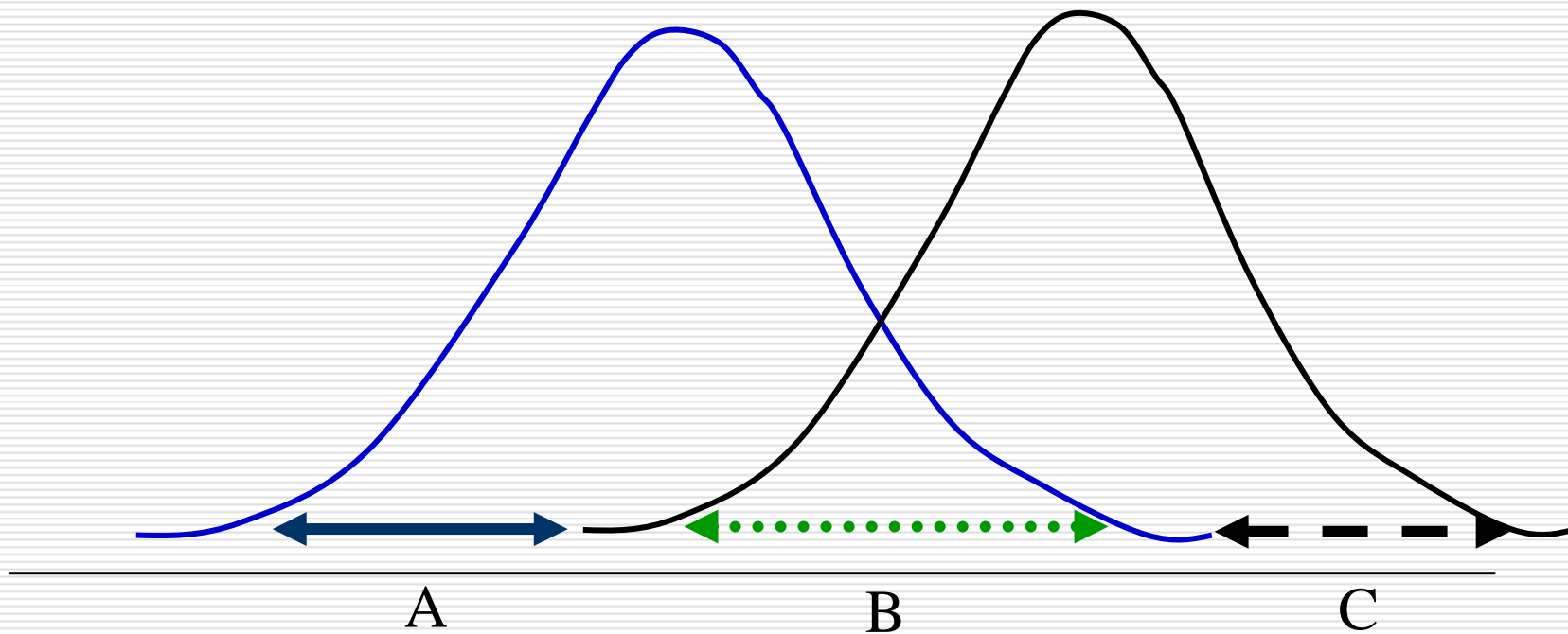
- 診断が比較的容易な重篤患者は避ける
- コントロールには、対象とする疾患と共通する症状を有している健常人を採用する

被験者の分布

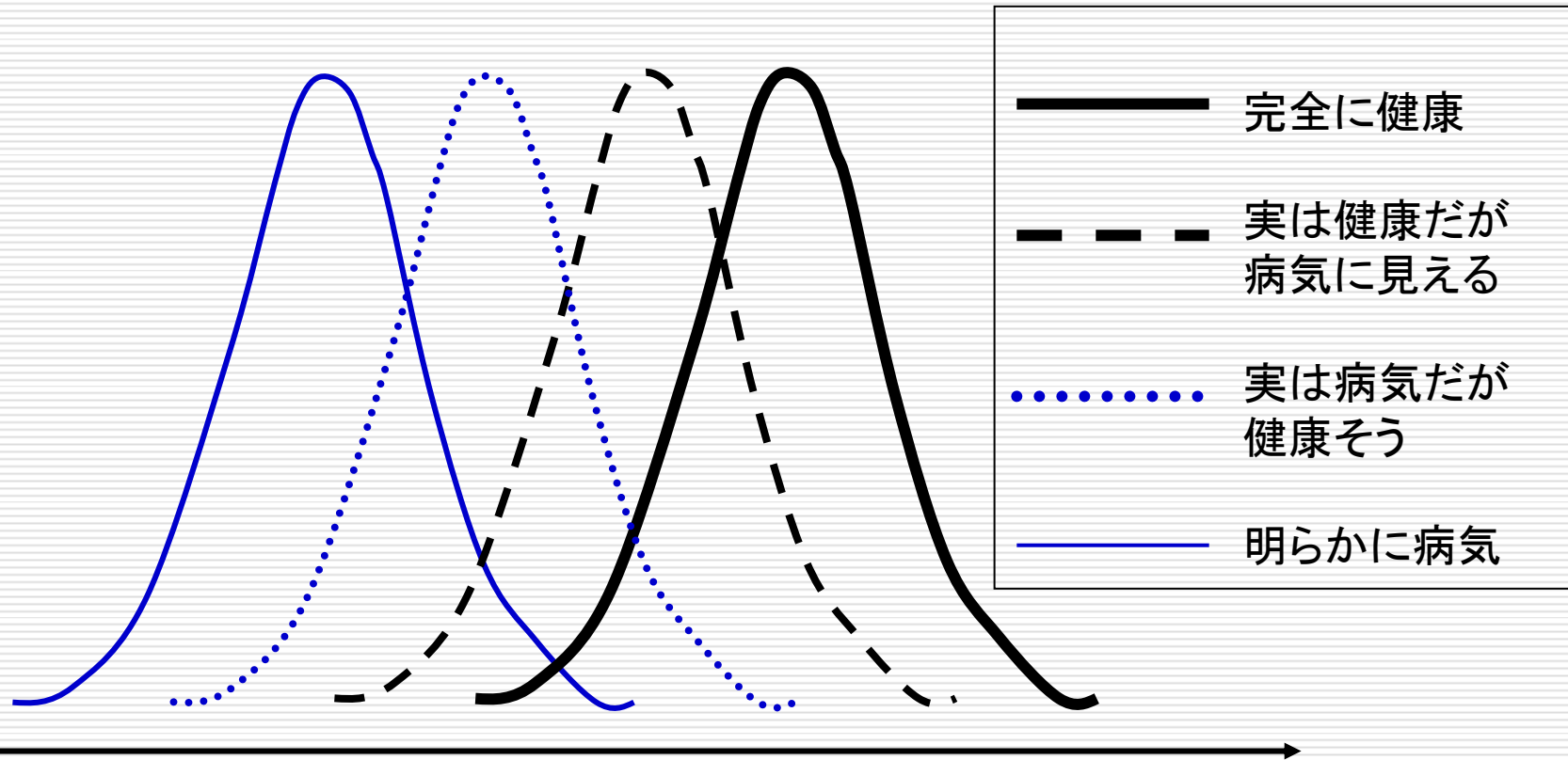
(検討対象によるバイアスはないのか?)

疾患なし(正常)

疾患あり(異常)



範囲バイアス spectrum bias



正当な評価のためには、スクリーニングの適用となる連続症例を対象とする

感度

- 疾患を有する者 ($a+c$) のうち、疾患を有すると診断される者 (検査陽性) (a) の割合

		疾 患	
		+	-
検 査	+	a	b
	-	c	d

$$\text{感度} = \frac{a}{a+c} = \frac{\text{真の陽性者数}}{\text{疾患を有する者の数}}$$

特異度

- 疾患がない者 ($b+d$) のうち、疾患を有さないと診断される者 (検査陰性) (d) の割合

		疾 患	
		+	-
検 査	+	a	b
	-	c	d

$$\text{特異度} = \frac{d}{b+d} = \frac{\text{真の陰性者数}}{\text{疾患の無い者の数}}$$

感度と特異度の例

		疾患		合計
		+	-	
検査	+	80	100	180
	-	20	800	820
合計		100	900	1000

$$\text{感度} = \frac{80}{100} = 80\%$$

$$\text{特異度} = \frac{800}{900} = 89\%$$



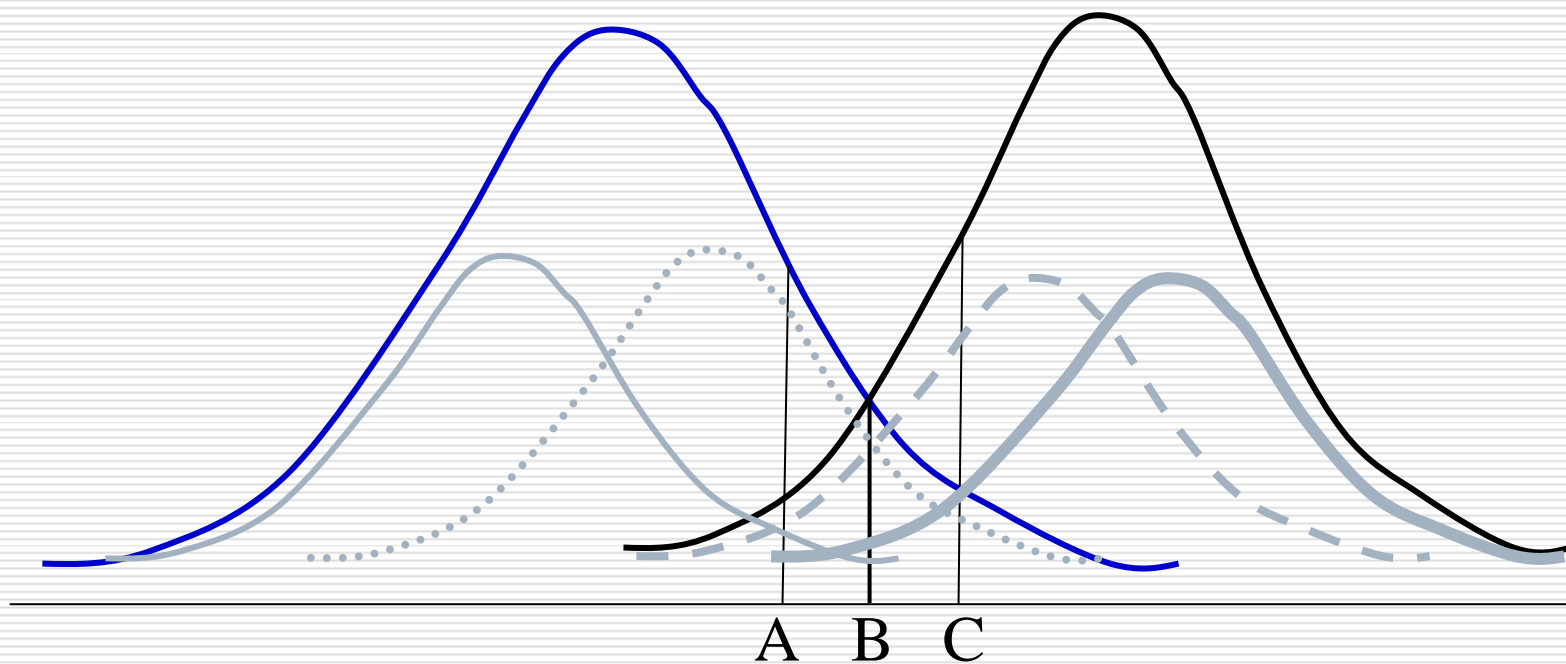
感度・特異度 まとめ

- ◎ 感度・特異度＝検査の「固有の性能」
 - ✓ 何らかのGold Standardを基準に算定される
 - ✓ 連続量の場合は、カットオフ値をずらせば変化する
 - ✓ 範囲バイアス spectrum bias に注意
- ◎ 感度＝病気を持っている人を拾い上げる性能
高いほど、「見落とし」(偽陰性)が少ない
- ◎ 特異度＝「だめ押し」
高いほど、「見過ぎ」(偽陽性)が少ない
- ◎ ! 注意:いくら高くても確定診断根拠としては不十分

被験者の分布

疾患なし(正常)

疾患あり(異常)



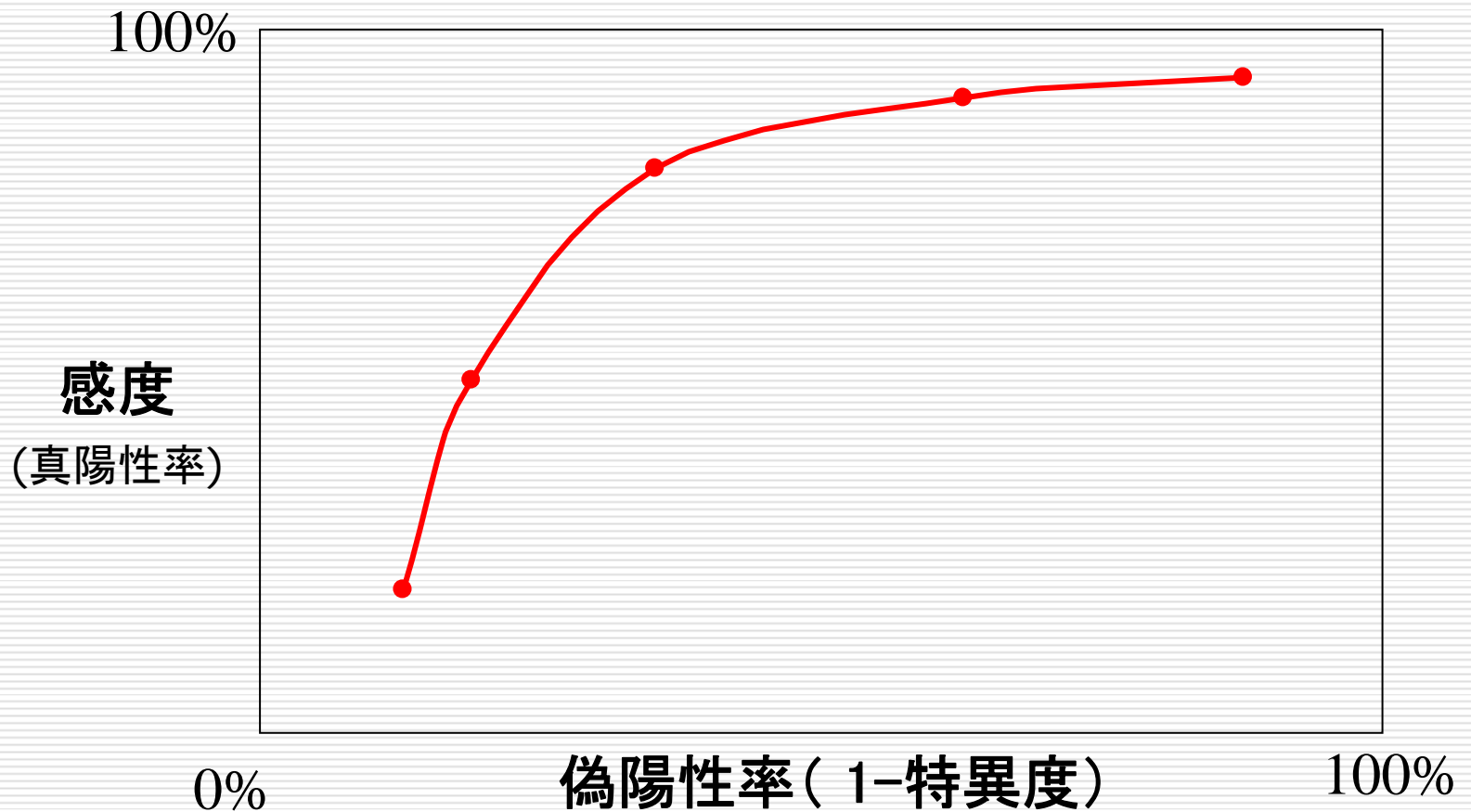
理想の検査とは？

- 感度が高い検査
- 特異度が高い検査
- これらの度合いを一度に示したい
- カットオフ値にとらわれずに、検査の性能を示す方法は？

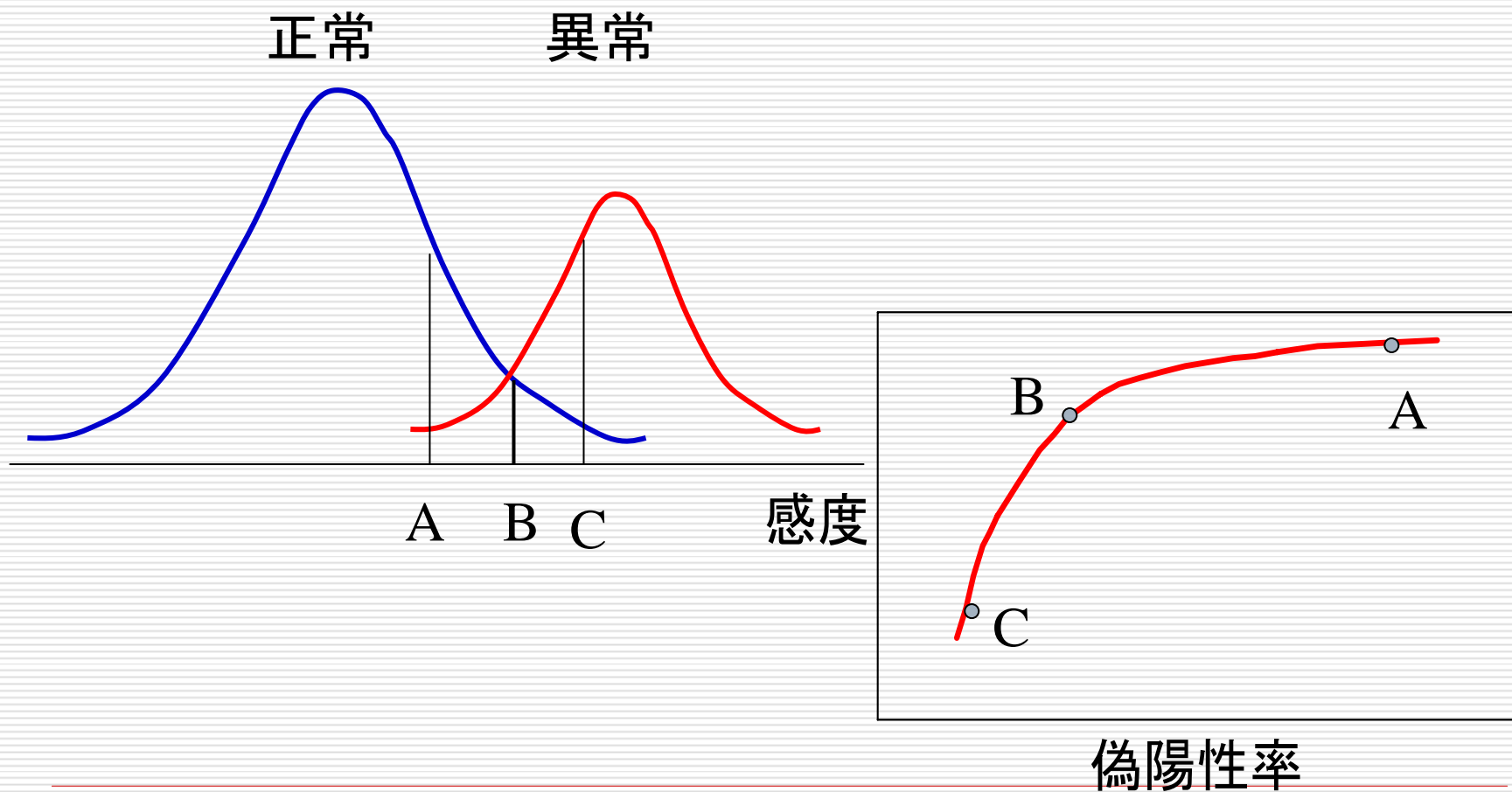
ROC曲線 (Receiver-operating characteristic curve)
受信者操作特性曲線

ROC曲線

ROC (Receiver-operating characteristic curve)



ROC曲線と被験者の分布との関係



ある予測式のROC曲線とAUC

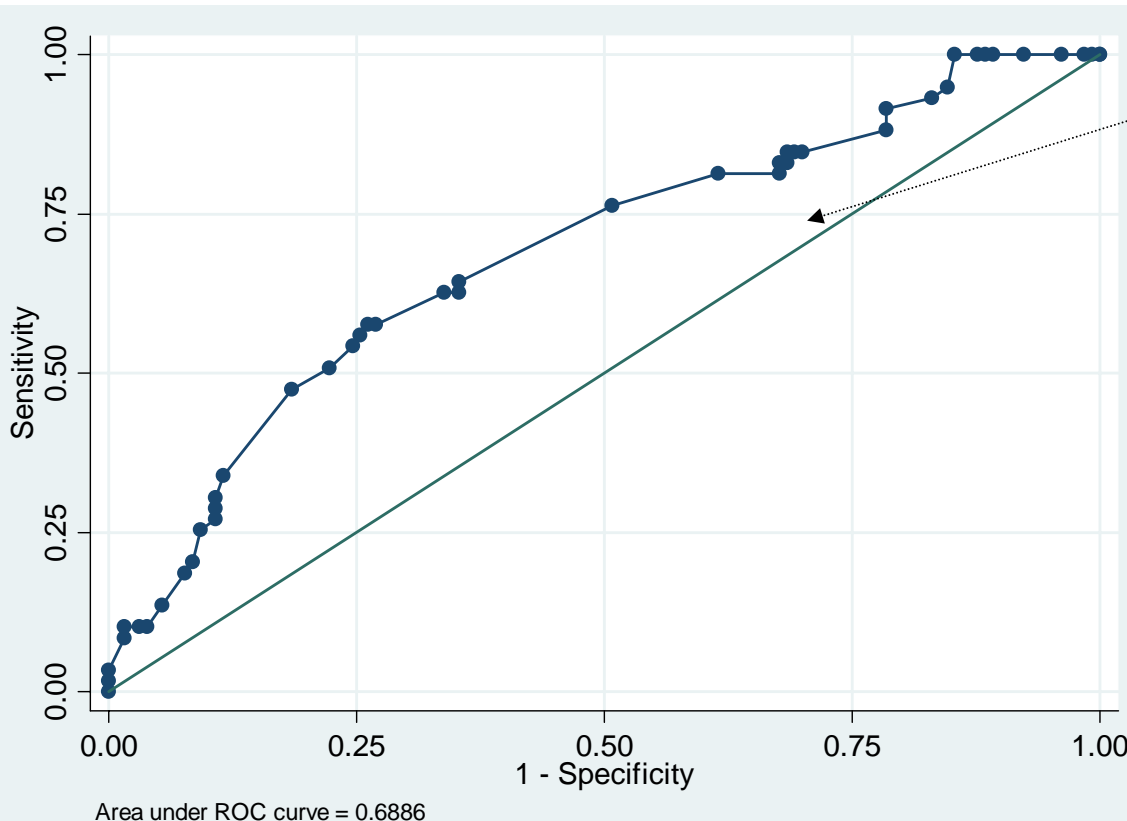
```
. lroc
```

Logistic model for low

number of observations = 189

area under ROC curve = 0.6886

ここが大きい方が良い



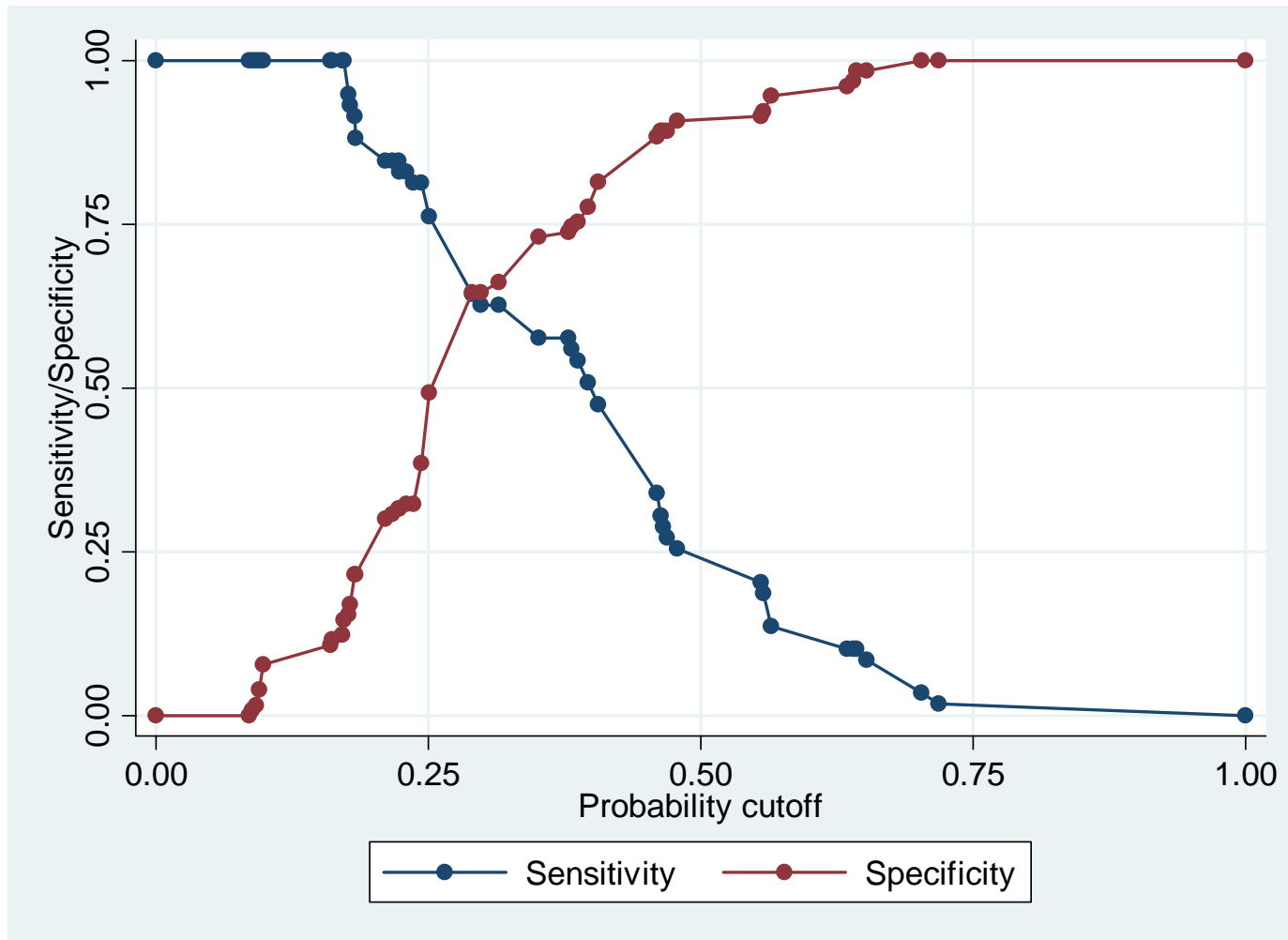
この面積を併せた値を**AUC**
=Area Under Curve

という。

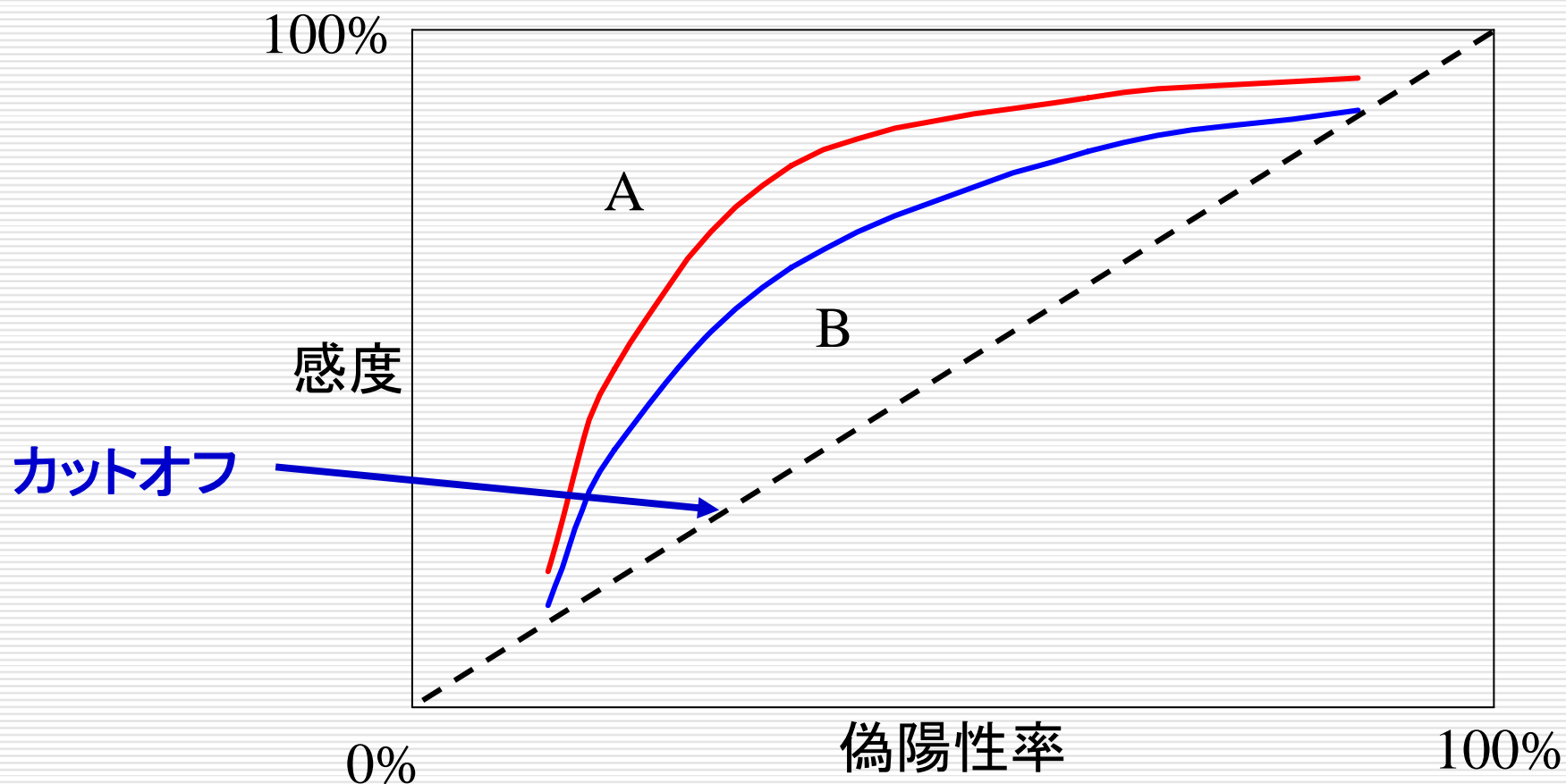
別名：**c statistics (c 統計量)**

カットオフ値を動かしたときの 感度、特異度の動きを表すグラフ

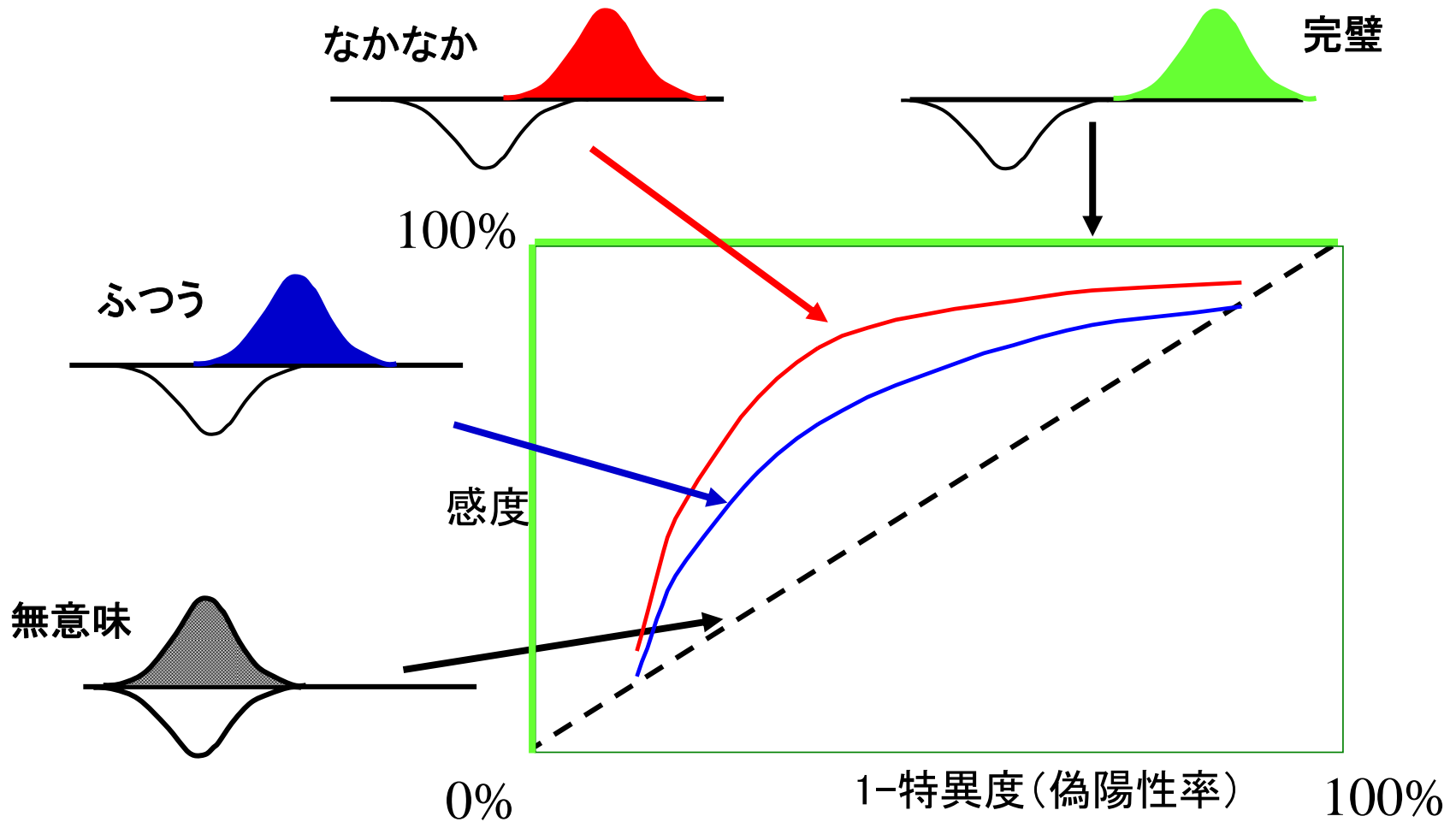
. lsens



ROC曲線によって 異なる検査法の判別能を比較できる

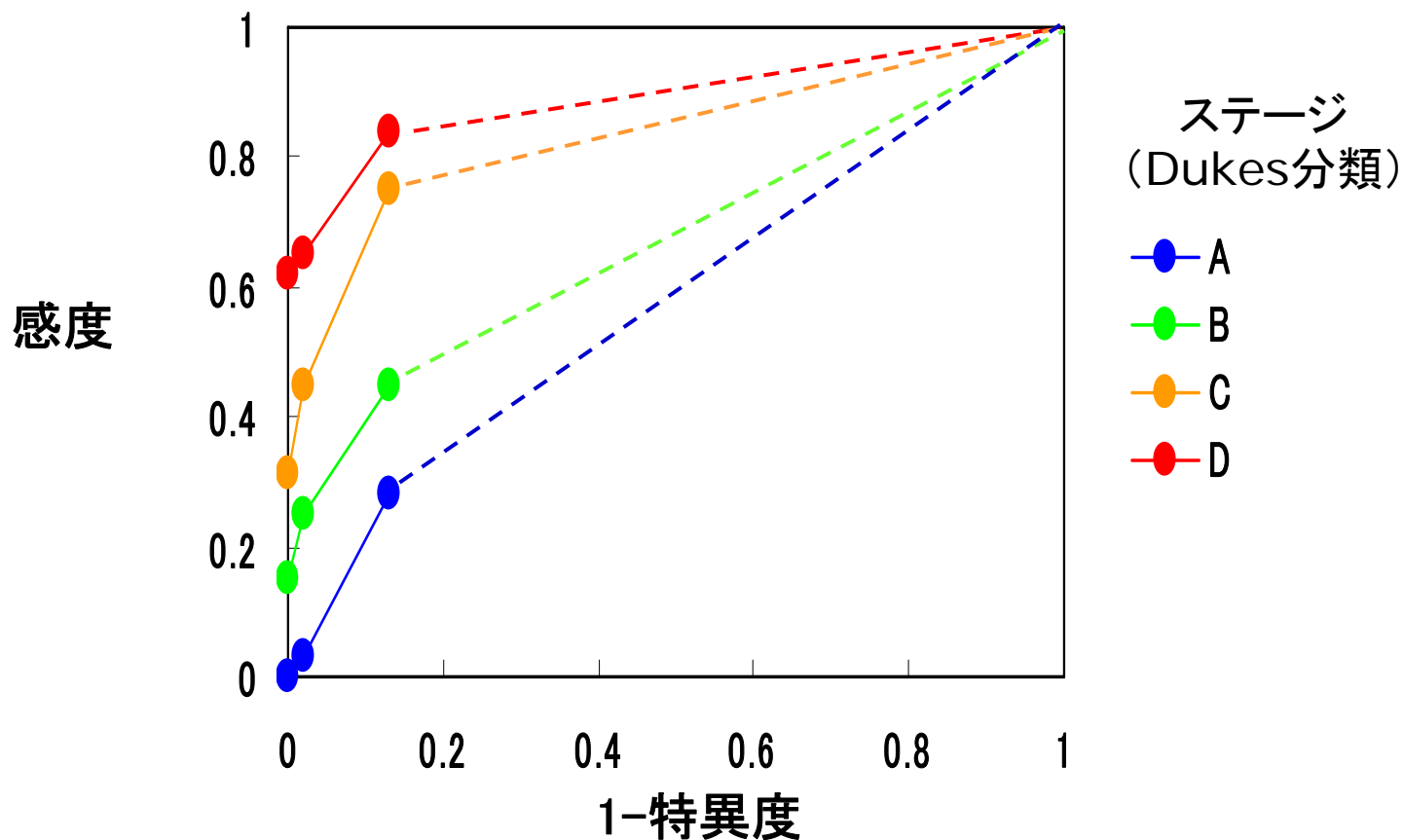


データの分布とROC曲線



検査の感度や特異度は 疾患のステージにより変化する

大腸癌の診断検査としてのCEAのROC曲線



ロジスティック回帰モデルの評価

■ モデル適合度 Goodness of Fit

モデルが予測した確率が観察されたアウトカム確率と適合するか(確率の適合度)

- Pearson's Goodness of Fit
- Hosmer-Lemshow Goodness of Fit

■ 判別力 Discriminatory ability

モデルが予測した確率にカットオフ値を加えた分類は、実際にアウトカムと合っているか

- AUC(c statistics)

臨床場面で本当に知りたいこと

「検査の結果が陽性(陰性)の時に
病気である(ない)可能性が
どれくらいあるか？」

ということ

話は「尤度比」へと続く・・・

- ② 診断テストの結果を解釈する時に
検査前確率と検査後確率 を考える
- ② 尤度比Likelihood ratioで
診断テストを評価・選択する「検査のインパクト指標」
簡便に、検査後確率を推定する
Bayes理論では、オッズ比を用いると、
検査前オッズ × 尤度比 = 検査後オッズ

サンプルサイズの設定の意義

- ② 統計的推測の精度はサンプルサイズに依存
 - 点推定の標準誤差、信頼区間の幅
 - 検定の検出力(統計学的パワー)
- ② 必要最低限の精度を確保
 - 精度の低い研究の排除
 - 過度に精度の高い研究の排除
- ② デザインの早期の段階で、エンドポイントと解析方法に基づいたサンプルサイズの設定を行う必要がある

サンプルサイズの設定～2つのアプローチ

① 推定精度に基づく設定（記述研究）

- 推定精度がある水準以上になるようにサンプルサイズを設定

② 検出力に基づく設定（実験的・分析的研究）

- 検出力がある水準以上になるようにサンプルサイズを設定

推定精度に基づく設定（信頼区間を用いる）

㊦ 連続変数の場合

- 変数の**標準偏差(S)**を推定
- 信頼区間の幅(W)を決める
- 信頼水準(1 - α)を決める(例: 95%)

$$N = 4z_{\alpha}^2 S^2 \div W^2$$

※ $z_{\alpha} = \alpha$ の標準正規偏差
(対立仮説が両側の場合、 $\alpha = 0.05$ のとき $Z_{\alpha} = 1.96$)

推定精度に基づく設定（信頼区間を用いる）

② 2区分変数の場合（割合を推定する場合）

- その要因を持つ人の**期待割合(P)**を見積もる
- 信頼区間の幅(W)を決める
- 信頼水準(1- α)を決める(例: 95%)

$$N = 4z_{\alpha}^2 P(1-P) \div W^2$$

P: 半数以上がその要因を保有している場合には、
有していない割合に基づいて計算

Z_{α} : α の標準正規偏差 ($\alpha = 0.05$ のとき $Z_{\alpha} = 1.96$)

割合の推定に必要なサンプルサイズの設定例

新しい検査の感度を決定する記述的研究

- パイロット研究では、80% (0.8) が陽性
- 感度の95%信頼区間を 0.8 ± 0.05 としたい
- 必要な患者数は？

1. 期待割合 (P) = 0.8

2. 信頼区間の幅 (W) = 0.10 ($\because \pm 0.05$)

3. 信頼水準 = 95%

$$N = 4z^2 P(1-P) \div W^2$$

$$= 4 \times 1.96^2 \times 0.8(1-0.8) \div 0.1^2 = 246 \text{人}$$

特異度を研究するには、疾患を有しない人のサンプルサイズを推定する 66

検出力に基づく設定

- ② 2群のアウトカムの割合を比較するとき
 - 検定の α レベル(有意水準)を決める(例:両側5%)
 - 検定の検出力($1 - \beta$)を決める(例:80%)
 - コントロール群のアウトカムの割合(P_0)
 - 比較群のアウトカムの割合(P_1)
- どのくらい上回れば臨床的に差があると考えられるか？
- (両群の割合の標準誤差)

検出力に基づく設定

(2群間の反応割合を比較)

$$N = \frac{[Z_{\alpha/2} \sqrt{2\bar{P}(1-\bar{P})} + Z_{\beta} \sqrt{P_0(1-P_0) + P_1(1-P_1)}]^2}{(P_0 - P_1)^2}$$

P_0 : コントロール群での反応割合

P_1 : 治療群での反応割合

$\bar{P} = (P_0 + P_1) / 2$

$Z_{\alpha/2}$: $\alpha/2$ の標準正規偏差、 Z_{β} : β の標準正規偏差

必要サンプルサイズの感度分析

- ◎ 必要サンプルサイズはパラメータの値に大きく依存する
 - 大ざっぱな見積もりと捉えるべき
- ◎ パラメータ値を変えたとき結果がどう変化するか評価すべき
 - 検出力を固定したサンプルサイズの変化
 - サンプルサイズを固定した検出力の変化

見積もりを変えた時のサンプルサイズ(1群)

(例) 禁煙プログラムの効果を調べる研究

α レベル: 両側5%、検出力: 90%

コントロール群の禁煙割合: 10%

禁煙プログラム群の禁煙割合20%とした場合

試験グループの禁煙割合

		50%	40%	30%	20%
検出力	90%	26	42	82	266
	80%	20	32	62	199
	70%	16	26	49	157
	60%	13	21	39	125
	50%	11	17	31	98

研究におけるセオリーの重要性

『我流は攻において威を発するが、
守に転じて威を失う』



ラ王(北斗の拳)

∴ 失敗をした時、短時間・短距離で修正可能

参考図書

- 井村裕夫. 臨床研究イノベーション. 東京: 中山書店; 2006
- 中山健夫、津谷喜一郎. 臨床研究と疫学研究のための国際ルール集. ライフサイエンス社; 2008
- 開原成允、浅井泰博 (監訳). JAMA 医学文献の読み方. 中山書店; 2001
- Hulley SB. *Designing Clinical Research: An Epidemiologic Approach 3rd ed.* Chapter 12
Lippincott Williams&Wilkins Philadelphia, 2007
(木原雅子、木原正博 訳: 医学研究デザイン 研究の質を高める疫学的アプローチ 第3版 メディカル・サイエンス・インターナショナル)

謝辞

▶ 京都大学 大学院医学研究科 社会健康医学専攻

- 健康情報学: 中山 健夫
- 予防医療学: 川村 孝、安藤昌彦
- 医療疫学: 福原 俊一、森田 智視、東 尚弘
- 医療統計学: 佐藤 俊哉、大森 崇
- 薬剤疫学: 松井 茂之



▶ 同 臨床研究者養成者コース (MCR) 1期生

- 石見拓
- 片多史明
- 川口武彦
- 北村和也
- 西内辰也
- 能城 毅
- 杉岡隆
- 白井貴子
- 松澤重行
- 松田二三子



**Master of
Clinical
Research**