

# 医学統計勉強会

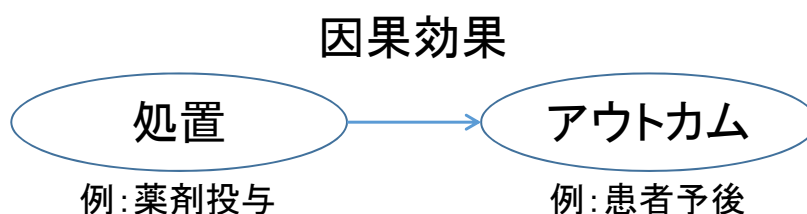
東北大学病院循環器内科・東北大学臨床研究推進センター 共催

東北大学大学院医学系研究科EBM開発学寄附講座

宮田 敏

## 傾向スコア

新しい処置 = 医学的介入の因果効果の検証



「処置」と「アウトカム」は、様々な交絡因子（予後因子）の影響を受ける。

無作為化比較試験 (Randomized Controlled Trial, RCT) では、介入のランダムな割り付けで予後因子を均等化する。

無作為化が困難な場合 ⇒ 傾向スコア

## 無作為化比較試験 (Randomized Controlled Trial, RCT)

- 新しい処置を受ける群 + 対照群 (control)
- 処置群と対照群の割り付けは、無作為割り付け。(randomization)
- ヒトを対象とする場合、情報開示と文書による同意。(informed consent)

無作為割り付けにより、**既知**および**未知**の交絡因子に関して、処置群と対照群が確率的な誤差を除いて**同等**。

観察研究では、無作為割り付けが可能とは限らない。

- 臨床では状態の悪い患者さんに、優先して投薬。
- 割り付けの基準はときに不明確、不可解。
- 喫煙など有害な習慣。

## 交絡因子による影響の調整

- **層別, マッチング**：共変量が「似た」サブグループに分割し、サブグループごとに解析  
⇒ 共変量が多いと層の数が多くなり、各層ごとの標本数が少なくなる。
- **多変量解析**：回帰分析, ロジスティック回帰, Cox比例ハザードモデル, 等。  
⇒ 共変量とアウトカムの関係性を、適切にモデル化する必要がある。**変数選択, 線形モデル**の仮定等。
- **傾向スコア (Propensity Score, PS)**：「処置」の有無に関わる全交絡因子の情報を集約する。

## 例：lindnerデーターPCI施行後のAbciximab投与の効果

lifepres : PCI施行6ヶ月経過後の平均余命。死亡=0, 生存=11.4  
 cardbill : PCI施行後6ヶ月間の心臓病関連医療費（ドル）  
 abcix : Abciximab投与有り=1, 投与なし=0  
 stent : ステンント  
 height : 身長  
 female : 性別（女性）  
 diabetic : 糖尿病  
 acutemi : 7日以内のAMI  
 ejecfrac : 左室駆出率

	lifepres	cardbill	abcix	stent	height	female	diabetic	acutemi	ejecfrac	
1		0	14301	1	0	163	1	1	0	56
2	11.6	3563		1	0	168	0	0	0	56
3	11.6	4694		1	0	188	0	0	0	50

## 例：lindnerデーターPCI施行後のAbciximab投与の効果

Name	abcix=0 (n=298)	abcix=1 (n=698)	p-value
cardbill	14614.22 ± 840.77	16126.68 ± 355.18	0.098
height	171.45 ± 0.61	171.44 ± 0.4	0.996
ejecfrac	52.29 ± 0.6	50.4 ± 0.39	0.009
stent	174 (58.4%)	492 (70.5%)	0
female	115 (38.6%)	231 (33.1%)	0.11
diabetic	80 (26.8%)	143 (20.5%)	0.031
acutemi	18 (6%)	125 (17.9%)	0

身長を除き、全ての共変量でAbciximab投与群と非投与群に有意差がある。

# 傾向スコア

【定義】 **傾向スコア (Propensity Score, PS)**

処置の有無に關与する複数の共変量を用いた、処置群、対照群への**割り付け確率を予測するスコア**。

$$PS = P(\text{treatment group} \mid x_i = X)$$

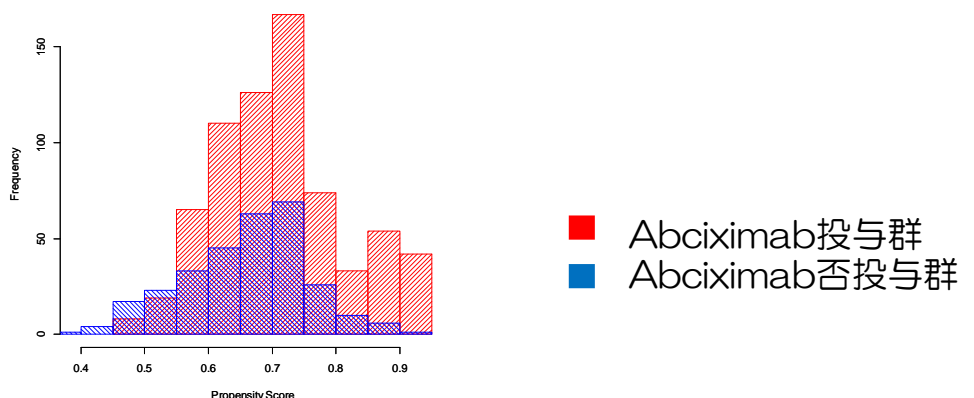
通常は、PSは**ロジスティック回帰モデル**によって推定される。処置の有無に關与する共変量は全てPS予測に用いられ、PS予測に用いられた共変量の情報は、全てPSに**吸収**されたと解釈する（それ以降の解析では、PS予測に用いた共変量は使用しない。）

例：lindnerデータのPropensity Score

ロジスティック回帰モデルによるPS

$$\log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1(\text{height})_i + \beta_2(\text{stent})_i + \beta_3(\text{female})_i \\ + \beta_4(\text{diabetic})_i + \beta_5(\text{acutemi})_i + \beta_6(\text{ejectfrac})_i$$

$p_i$ : Abciximab投与群への割り付け確率 = **傾向スコア**



## 傾向スコアの特徴：

- 群の割り付けに関する多くの交絡因子を、一次元の情報に集約し調整する。（交絡因子の数がイベントに比べて多いとき、次元の縮小に有用）
- 同じ傾向スコアを持つ標本は、処置群と対照群の多くの共変量が同時にバランスされている。

傾向スコアマッチング後の共変量の比較

Name	abcix=0 (n=275)	abcix=1 (n=275)	p-value
cardbill	14841.53 ± 902.45	16009.2 ± 536.21	0.267
height	171.5 ± 0.64	171.37 ± 0.63	0.884
ejecfrac	51.83 ± 0.62	51.41 ± 0.61	0.631
stent	174 (63.3%)	185 (67.3%)	0.37
female	99 (36%)	109 (39.6%)	0.429
diabetic	62 (22.5%)	63 (22.9%)	1
acutemi	18 (6.5%)	15 (5.5%)	0.72

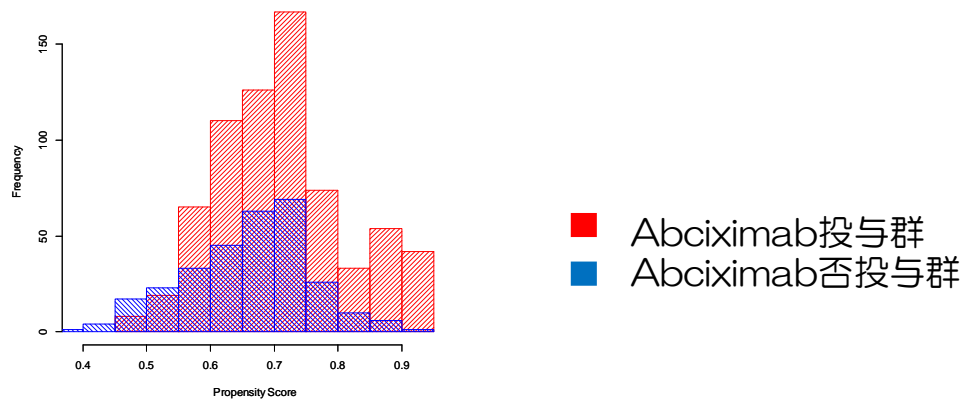
## 傾向スコアのマッチング (Propensity Score Matching)

傾向スコアが同じサンプルは、交絡因子の分布が同じになる傾向がある。 傾向スコアは実数値であるので、適当な有効桁数に丸める (caliper)。

処置群と対照群で、同じ傾向スコアを持つサンプルをマッチさせていく。もし、同じ傾向スコアを持つサンプルが複数あれば、順番はランダムにマッチングさせる。処置群と対照群で同じ傾向スコアを持つ相手がいない場合は、データから除かれる。

# 例：lindnerデータのPropensity Score

## 傾向スコアの分布



処置群、対照群のPS分布の重なり部分のサンプルがマッチングされる。重なっていない部分のサンプルは、マッチングデータからは削除される。

## 傾向スコアモデルのチェック

- 傾向スコアを推定するのに用いたロジスティック回帰モデルの、パフォーマンスを評価する
  - ⇒ Hosmer-Lemeshowの適合度検定
  - ⇒ ROC曲線
- 同じ傾向スコアを持つサンプルは、交絡因子の分布が同じになる傾向がある。

### 傾向スコアマッチング後の共変量の比較

Name	abcix=0 (n=275)	abcix=1 (n=275)	p-value
cardbill	14841.53 ± 902.45	16009.2 ± 536.21	0.267
height	171.5 ± 0.64	171.37 ± 0.63	0.884
ejecfrac	51.83 ± 0.62	51.41 ± 0.61	0.631
stent	174 (63.3%)	185 (67.3%)	0.37
female	99 (36%)	109 (39.6%)	0.429
diabetic	62 (22.5%)	63 (22.9%)	1
acutemi	18 (6.5%)	15 (5.5%)	0.72

## 傾向スコアモデルのチェック（続き）

- マッチングされた後のサンプル数を確認する。処置群と対照群で傾向スコアの分布の乖離があまりに大きいと、マッチされるサンプルが少なくなり、解析に適さない。

⇒ 傾向スコアの乖離が大きい場合は、マッチングを行わず他の方法で傾向スコアを用いることを検討すべき。（傾向スコアによる層別化、重み付け、共変量としての利用、他）

## 傾向スコアを用いた解析法

- Matching: 傾向スコアマッチングを行い、その後は、処置群、対照群の別による **log-rank 検定** を行う。
- 層別：処置群、対照群をプールした上で、傾向スコアの値で層別する（通常5層に分ける）。その上で、層別に解析を行う。（例：stratified Cox proportional hazard model）
- 共変量としての傾向スコア：マッチングは行わず、全てのサンプルを用い、PSを共変量とした多変量モデルを当てはめる。

$$\log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \gamma G_i + \beta_1(\text{PS})_i, G_i = \begin{cases} 1: \text{treatment group} \\ 0: \text{control group} \end{cases}$$

$$\log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \gamma G_i + \beta_1(\text{PS})_i + \beta_2(\text{other})_i + \dots$$

## 傾向スコアを用いた解析法

- 重み付け：Inverse-probability score-based weighted methods (IPTW)

傾向スコアの逆数を重み付けとして、各種多変量モデルを推定する。線形回帰モデル、ロジスティック回帰モデル、Cox比例ハザードモデルなどいずれの解析モデルにも、重み付き回帰モデルが存在するので、その重みに傾向スコアの逆数を用いるということ。

傾向スコアを用いた解析法には、未だ最終的な答えはない。上記いずれかの方法を、少なくとも複数用いて、同じような結果が出ることを確かめるべき。

### Take Home Message

1. 因果効果と交絡因子
2. 傾向スコア
3. 傾向スコアマッチング
4. 傾向スコアモデルのチェック
5. 傾向スコアを用いた解析法
  - i. 傾向スコアマッチング：対応づけられたデータに対する単変量解析
  - ii. 層別
  - iii. 共変量としての傾向スコア
  - iv. 重み付け：Inverse-probability score-based weighted methods (IPTW)

### 参考文献

Rosenbaum & Rubin, “The central role of the propensity score in observational studies for causal effects” *Biometrika* (1983) 70 (1): 41-55.

星野 崇宏「調査観察データの統計科学—因果推論・選択バイアス・データ融合」  
岩波書店 (2009/7/29) ISBN-10: 4000069721