

臨床試験データへの GLMSELECT procedureの適用

横溝孝明¹、三角俊裕²、高橋行雄³

¹大正製薬株式会社 臨床統計部

²アステラス製薬株式会社 データサイエンス部

³中外製薬株式会社 臨床企画推進部

An application of GLMSELECT procedure for clinical trial data

Takaaki Yokomizo¹, Toshihiro Misumi², Yukio Takahashi³

¹Taisho Pharmaceutical Co., Ltd.

²Astellas Pharma Inc.

³Chugai Pharmaceutical Co., Ltd.

1/23

要旨:

得られた臨床試験データから、
反応変数に影響を与えうる変数を効率的に
選択する方法について
GLMSELECT procedureを用いて検討した

キーワード: 臨床試験、探索的解析、変数選択、
交互作用、GLMSELECT procedure

2/23

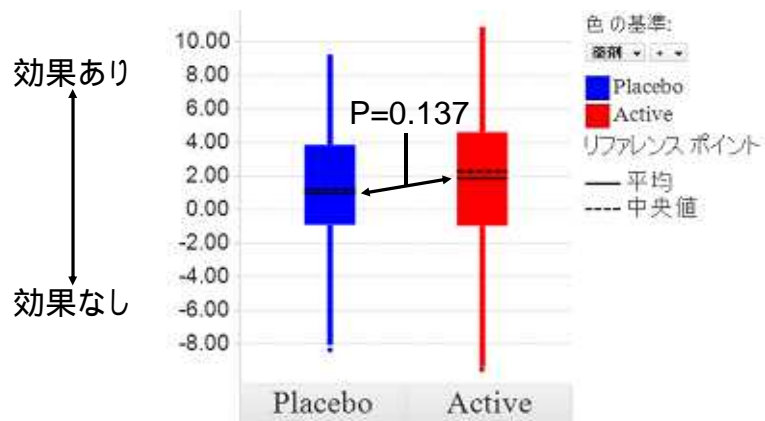
謝辞

この研究は、財団法人 日本科学技術連盟の
「臨床試験セミナー 統計手法専門コース（略称：BioS）
（運営委員長：大橋靖雄）」
のフォロー研修の一環として行われたものである

3/23

本研究の背景 (1/4)

例．臨床試験 (ActiveとPlaceboの2群比較)



全集団で、ActiveとPlaceboの間に差がなかった

4/23

本研究の背景 (2/4)

例 . 臨床試験 (ActiveとPlaceboの2群比較)

探索的解析として背景項目ごとの結果をみていき、
原因究明や次試験のために、効果がうかがえる集団を探す
(たとえば、性別)



5/23

本研究の背景 (3/4)

例 . 臨床試験 (ActiveとPlaceboの2群比較)

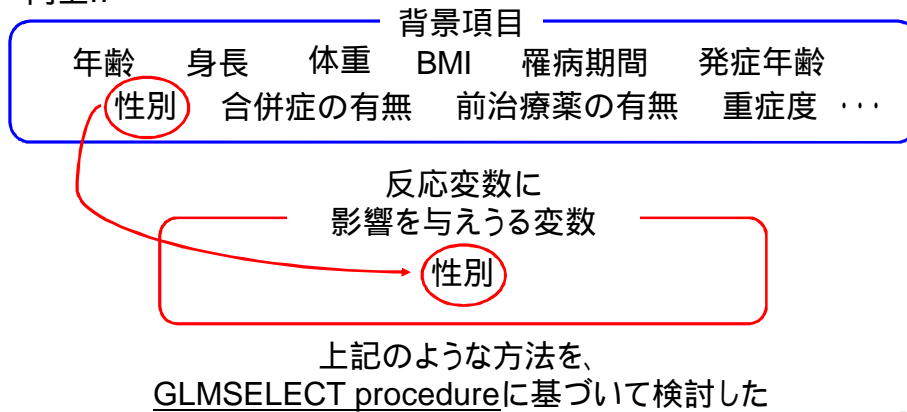
しかしながら、試験によっては背景項目が100個にもおよぶ
臨床試験データでは、一つ一つの背景項目ごとに
効果を見ていくことは大変!!



6/23

本研究の背景 (4/4)

得られた臨床試験データから、**反応変数に影響を与えうる変数**を効率的に選択できる方法があれば、解析業務の質および生産性向上!!



7/23

本研究の目的

得られた臨床試験データから、**反応変数に影響を与えうる変数**を効率的に選択する方法について検討する

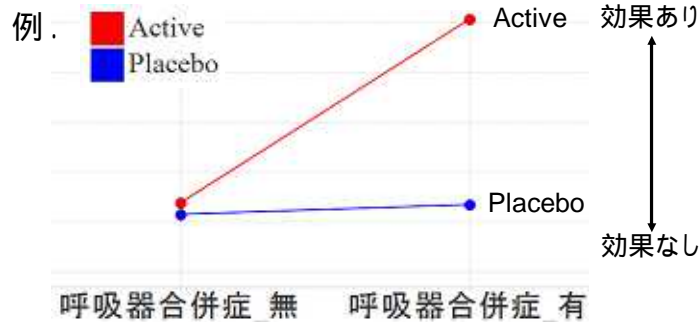
選択された変数について、**医学的に意味のある変数**ということが確認できれば、その後の研究仮説を引き出すために大いに役立てることができる

ICH E9「臨床試験のための統計的原則」においても、探索的解析の必要性について記載されている

では、**反応変数に影響を与えうる変数**とはどのようなものか？

8/23

反応変数に影響を与える変数の考え方



「薬剤」と「呼吸器合併症の有無」の間に**交互作用**あり

↓
「呼吸器合併症の有無」は反応変数に影響を与える変数

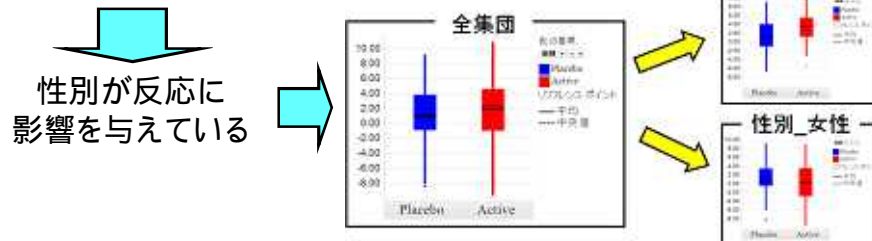
薬剤との交互作用を見つけてやればよい!!!

薬剤との**交互作用**を効率的にみつけるために、
以下のような線形モデルに、変数選択法を適用することを考える

線形モデル

$$\text{反応変数} = \text{薬剤} * \text{性別} + \text{薬剤} * \text{年齢} + \text{薬剤} * \text{BMI} + \dots$$

変数選択法により**薬剤*性別**が変数選択された!!



よって、交互作用項の変数選択が行えれば
本研究の目的が果たせると考える **GLMSELECTの登場!!**

GLMSELECT procedureについて

(参考: Cohen, 2006)

- SAS 9.1.3では評価版、9.2で正式版として利用可能
- GLM procedureに変数選択の機能を付加したイメージ
- 説明変数の数が多くなっても変数選択可能
- Hot** □ 説明変数に**交互作用項**があっても変数選択可能
- 近年提案された変数選択法の
LAR (Efron et al., 2004)と
LASSO (Tibshirani et al., 1996)が利用可能
説明変数の数が症例数より多くても ($N < p$ でも)
変数選択可能

11/23

各procedureの比較

- REG procedure:
様々な変数選択法が用意されているが、
交互作用項を含む変数選択が行えない
- GLM procedure:
交互作用項を含む回帰分析は行えるが、
変数選択が全く行えない
- GLMSELECT procedure:
様々な変数選択法が用意されており、
交互作用項を含む変数選択が可能

12/23

REG と GLMSELECT の比較

y: 反応変数 (連続変数)
 x1, x2, x3: 説明変数 (連続変数)
 c1, c2, c3: 説明変数 (カテゴリカル変数)

- REG procedure

```
proc reg data=data;
  model y=x1 x2 x3/selection=stepwise SLE=0.15 SLS=0.15;
run;
```

- GLMSELECT procedure

```
proc glmselect data=data;
  class c1 c2 c3;
  model y = x1 x2 x3
        c1 c2 c3
        x1*x2 x1*c1
        /selection=stepwise(select=SL SLE=0.15 SLS=0.15);
run;
```

CLASSステートメントが
利用可能

交互作用項を含む
変数選択が可能

3/23

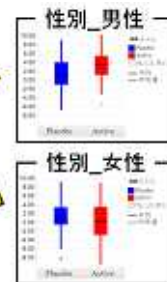
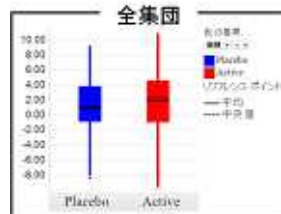
本研究の目的 (再確認)

線形モデル

反応変数 = 薬剤*性別 + 薬剤*年齢 + 薬剤*BMI + ...

薬剤*性別が変数選択された!!

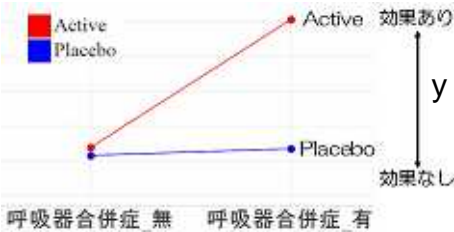
性別が反応に
影響を与えている



臨床試験データ (人工データ) に対して、
 GLMSELECTを適用し、交互作用項が選択されるかを確認した 14/23

臨床試験データ(人工データ)への適用

症例番号	反応変数 y	薬剤	呼吸器合併症の有無	年齢	身長	体重	BMI	罹病期間	重症度	性別	病型
1	5.23	Placebo	呼吸器合併症_無	49	152.2	61.3	24.5	-13	中等度	男	病型1
2	3.80	Placebo	呼吸器合併症_無	45	157.8	90.2	19.7	-12	中等度	男	病型2
3	11.18	Active	呼吸器合併症_有	62	156.1	82.7	24.3	-10	中等度	女	病型1
...				⋮							
199	6.48	Active	呼吸器合併症_有	48	172.3	62.0	24.1	-8	軽度	女	病型3
200	1.59	Active	呼吸器合併症_無	55	164.9	65.7	23.6	-8	中等度	女	病型1



- ・症例数200例、反応変数 y
- ・「薬剤」と「呼吸器合併症の有無」の間に交互作用あり
- ・背景項目25個 (連続変数20個、カテゴリカル変数5個)

上記データに対してGLMSELECTを用い、Stepwise法により、「薬剤」と「呼吸器合併症の有無」の交互作用項が選択されるかを確認した 15/23

SASコード

症例番号	反応変数 y	薬剤	呼吸器合併症の有無	年齢	身長	体重	BMI	罹病期間	重症度	性別	病型
1	5.23	Placebo	呼吸器合併症_無	49	152.2	61.3	24.5	-13	中等度	男	病型1
2	3.80	Placebo	呼吸器合併症_無	45	157.8	90.2	19.7	-12	中等度	男	病型2
3	11.18	Active	呼吸器合併症_有	62	156.1	82.7	24.3	-10	中等度	女	病型1
...				⋮							
199	6.48	Active	呼吸器合併症_有	48	172.3	62.0	24.1	-8	軽度	女	病型3
200	1.59	Active	呼吸器合併症_無	55	164.9	65.7	23.6	-8	中等度	女	病型1

```
proc glmselect data=data;
  class 薬剤 呼吸器合併症の有無 重症度 性別 病型 ...;
  model y = 薬剤 呼吸器合併症の有無 年齢 身長 体重 BMI
            罹病期間 重症度 性別 病型 ...;
  /selection=stepwise(select=SL SLE=0.15 SLS=0.15);
run;
```

主効果

薬剤との交互作用

結果

選択された変数

Parameter Estimates

```

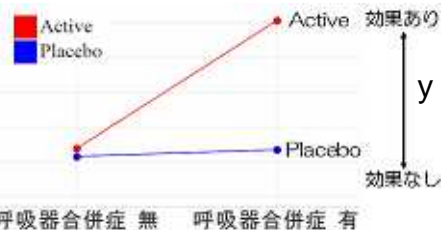
Intercept
key      Active
key      Placebo
umu      呼吸器合併症_無
umu      呼吸器合併症_有
e0?
e1?
e12
e13
key*umu  Active 呼吸器合併症_無
key*umu  Active 呼吸器合併症_有
key*umu  Placebo 呼吸器合併症_無
key*umu  Placebo 呼吸器合併症_有
    
```

薬剤との
交互作用項

薬剤との交互作用として、
「呼吸器合併症の有無」が選択されることを確認した

シミュレーションの実施

症例番号	反応変数 y	薬剤	呼吸器合併症の有無	年齢	身長	体重	BMI	罹病期間	重症度	性別	病型
1	5.23	Placebo	呼吸器合併症_無	49	152.2	61.3	24.5	-13	中等度	男	病型1
2	3.80	Placebo	呼吸器合併症_無	45	157.8	90.2	19.7	-12	中等度	男	病型2
3	11.18	Active	呼吸器合併症_有	62	156.1	82.7	24.3	-10	中等度	女	病型1
				⋮							
199	6.48	Active	呼吸器合併症_有	48	172.3	62.0	24.1	-8	軽度	女	病型3
200	1.59	Active	呼吸器合併症_無	55	164.9	65.7	23.6	-8	中等度	女	病型1



- ・症例数200例、反応変数 y
- ・「薬剤」と「呼吸器合併症の有無」の間に交互作用あり
- ・背景項目25個
(連続変数20個、カテゴリカル変数5個)

上記のようなデータを人工的に1000組作成し、「薬剤」と「呼吸器合併症の有無」の交互作用項がどのくらい選択されるか検証した

シミュレーションの設定条件

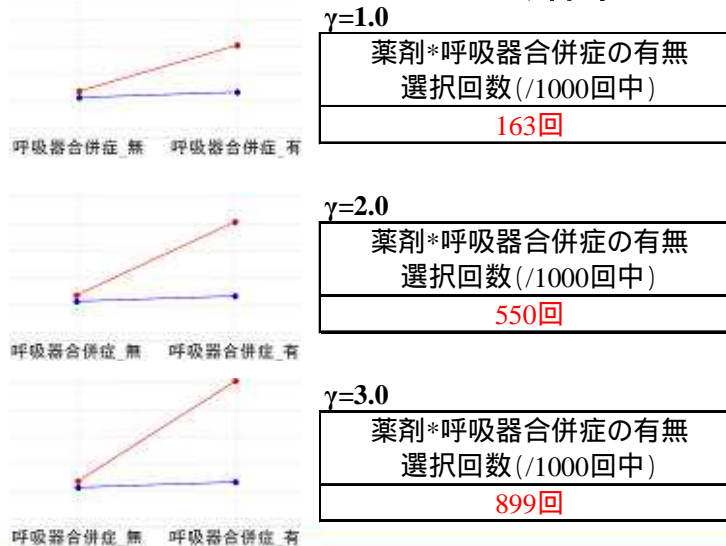
・真のモデル: $y = 1.0 + \gamma \times \text{薬剤} * \text{呼吸器合併症の有無} + \varepsilon$, $\varepsilon \sim N(0, 4^2)$

薬剤(=Active)*呼吸器合併症(=有) 1、その他 0

- ・身長、年齢、性別などの背景項目:
正規乱数、一様乱数をもとに薬剤間で偏りなく作成
(連続変数20変数、カテゴリカル変数5変数の計25変数)
- ・サンプルサイズ: 200例 (Active: Placebo=1:1)
- ・変数選択法として、Stepwise法を使用
(変数の組み入れ・除外水準SLを通常使われる0.15と設定)
- ・ $\gamma = 1.0, 2.0, 3.0$ の3パターンで1000回データを発生させ、
「薬剤*呼吸器合併症の有無」の選択回数を比較

19/23

シミュレーションの結果



20/23

シミュレーションの結果

- 交互作用の傾き (γ) がある程度大きければ、目的の交互作用項が選択される
- 変数の組み入れ・除外水準 (SL) をいくつに設定するかは課題はあるが、本研究の目的は真のモデルの推定ではなく、**薬効に影響を与える変数の探索**であるので、見逃しを防ぐためにも、SLをある程度大きく設定すべきであると考えられる (よく用いられるSL=0.15が妥当と考える)

21/23

まとめ

- 得られた臨床試験データから、反応変数に影響を与える変数を効率的に選択する方法についてGLMSELECT procedureを用いて検討した
なお、反応変数がカテゴリカル変数の場合は、LOGISTIC procedureで同様のことを行うことが可能 (高橋, 2006)
- 場合によっては背景項目が100個以上にもなる臨床試験データにおいて、検討手法を用い、**薬効に影響を与える変数に当たりをつけることができれば、効率的に探索的解析を行える**と考えられる

22/23

参考文献

1. 医薬審第1047号(平成10年11月30日付)「臨床試験のための統計的原則」
2. Cohen, R. (2006), Introducing the GLMSELECT Procedure for Model Selection, Proceedings of the Thirty-First Annual SAS Users Group International Conference.
3. Efron, B., Hastie, T., Johnstone, I., and Tibshirani, R. (2004), Least Angle Regression (with Discussion), *Annals of Statistics*, 32, 407–499.
4. Tibshirani, R. (1996), Regression Shrinkage and Selection via the Lasso, *Journal of the Royal Statistical Society Series B*, 58, 267–288.
5. 高橋 行雄(2006), 順序カテゴリー反応の探索的データ解析
<http://www.yukms.com/biostat/takahasi/rec/023.htm>