Propensity Score 分析

聖路加国際病院一般内科・
聖ルカ・ライフサイエンス研究所臨床疫学センター

高橋　理

内容

* 理論編
  * なぜPSか？
  * PSとは？
  * PSを利用した分析法は？

* 実践編
  * ロジスティック回帰分析によるPS推定
  * PSモデルの評価
  * 主なPS分析法

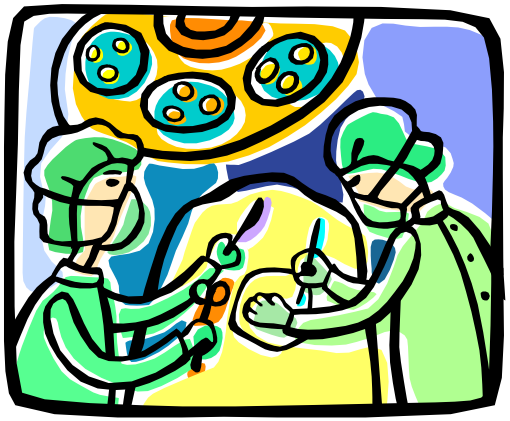
臨床医学分野での PS研究のゴール

医学的介入の因果効果を検討する
質の高い研究

医学的介入

* 薬剤治療

* 手術



* 手技など

ランダム化比較研究:RCT

医学における比較研究の
ゴールドスタンダード

1. ランダム化割り付け
2. (プラセボ)対照
3. 盲検
4. 試験

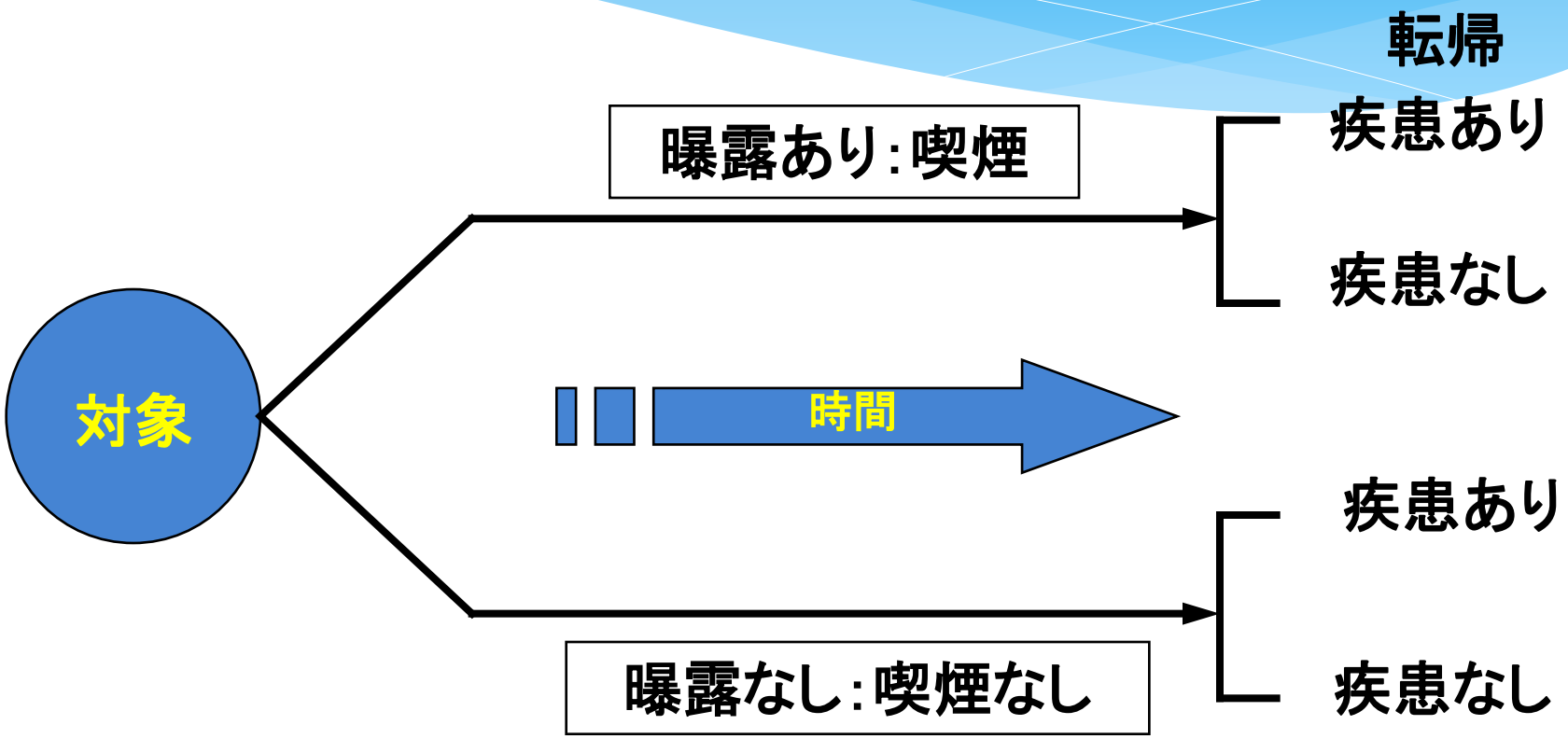
いつもRCT？

無作為化試験がいつも実行可能とは限らない：

- 倫理的（例：喫煙，）
- 計画的に見て実行不可能（時間、費用）
- 治療・手技がすでに日常診療に取り入れられている
  （例：インフルエンザワクチン,手術*）

* Br J Surg. 2006 93(4):389

観察研究：コホート研究の構造



治療研究でのコホート研究の問題点

- 治療が非ランダムに割り付けられている
  - 医師は最も治療を必要としている患者を優先
    （適応がある患者・服薬コンプライアンス ）
  - 治療割り付けのメカニズムは測定が困難で、
    ときにはまったく不可解(local rule)
- 様々な交絡因子（Confounding）が存在する
- Unknown Confounderの調整ができない

交絡のコントロール

- 研究デザイン
  - 従来のデザイン（コホート、ケースコントロール）
  - 特殊デザイン（ケースクロスオーバーなど）

- データ分析方法
  - マッチング
  - 制限
  - 層別化
  - 多変量調整
  - ***傾向スコア(Propensity score: PS) 手法***

Klungel, et al. J Clin Epidemiol 2004

PSとは

* 1983年にローゼンバウムとルービンが導入
* 重回帰のようなモデルによる調整に代わる*手法*
* PS手法は交絡変数の全情報を集約する方法

Rosenbaum & Rubin. *Biometrika* 1983

Propensity Score(PS)

複数の共変量を用いて2群に割り当てられる確率を予測するスコア

$P(y_i)$ = Probability (exposure | $x_i = \boldsymbol{X}$)

A propensity score P(y) is the conditional probability of receiving the exposure given a set of observed covariates ***X***.

$$0<P(y_i)<1$$

PS手法の手順

PS手法は2段階の手順

PSの推定 → 分析

PS analysis using

1. PS modeling using Logistic regression

2. Analyzing using PS

   – Matching

   – Regression

   – Weighting

   – Stratification

PS modeling

Logistic regression: treatment+/-(dependent variable）

& adjust for risk factors

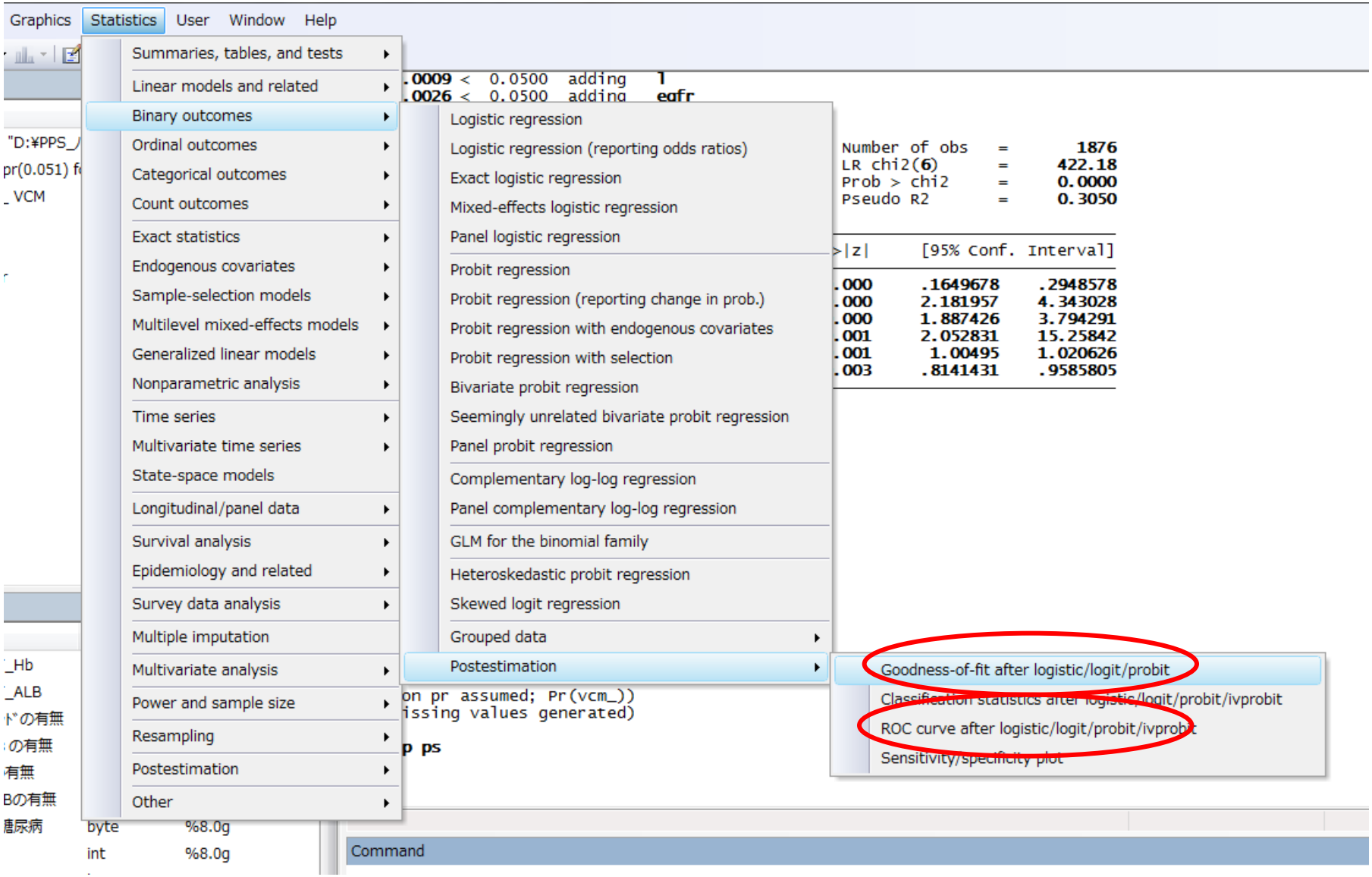
SW, pe(0.05) pr(0.051):logistic VCM Age Gender‥‥

（treatment)

predict PS

Evaluation of model

- Hosmer-Lemeshow: fitness
- ROCとAUC: discrimination


Graphics
Statistics
User
Window
Help
Summaries, tables, and tests
Linear models and related
Binary outcomes
Ordinal outcomes
Categorical outcomes
Count outcomes
Exact statistics
Endogenous covariates
Sample-selection models
Multilevel mixed-effects models
Generalized linear models
Nonparametric analysis
Time series
Multivariate time series
State-space models
Longitudinal/panel data
Survival analysis
Epidemiology and related
Survey data analysis
Multiple imputation
Multivariate analysis
Power and sample size
Resampling
Postestimation
Other
Logistic regression
Logistic regression (reporting odds ratios)
Exact logistic regression
Mixed-effects logistic regression
Panel logistic regression
Probit regression
Probit regression (reporting change in prob.)
Probit regression with endogenous covariates
Probit regression with selection
Bivariate probit regression
Seemingly unrelated bivariate probit regression
Panel probit regression
Complementary log-log regression
Panel complementary log-log regression
GLM for the binomial family
Heteroskedastic probit regression
Skewed logit regression
Grouped data
Postestimation
Goodness-of-fit after logistic/logit/probit
Classification statistics after logistic/logit/probit/ivprobit
ROC curve after logistic/logit/probit/ivprobit
Sensitivity/specificity plot
Number of obs = 1876
LR chi2(6) = 422.18
Prob > chi2 = 0.0000
Pseudo R2 = 0.3050
[95% Conf. Interval]
.1649678 .2948578
2.181957 4.343028
1.887426 3.794291
2.052831 15.25842
1.00495 1.020626
.8141431 .9585805
Command

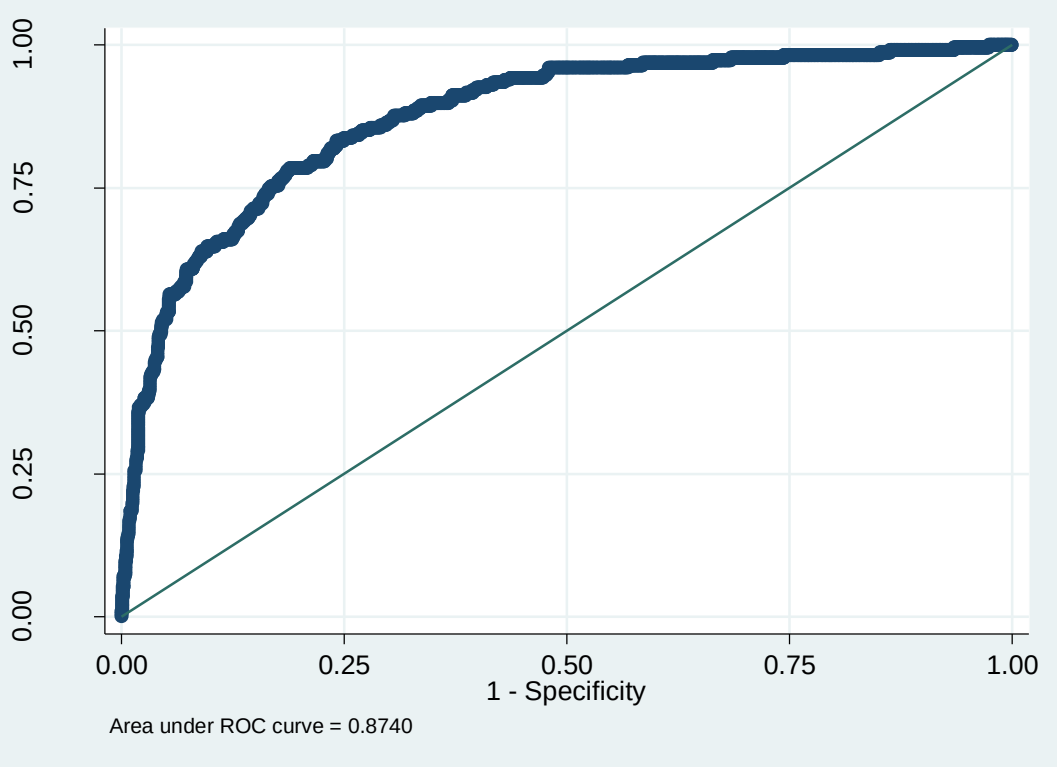
Hosmer-Lemeshow

```
. estat gof, group(10) table

Logistic model for VCM, goodness-of-fit test

(Table collapsed on quantiles of estimated probabilities)
```

Group	Prob	Obs_1	Exp_1	Obs_0	Exp_0	Total
1	0.0099	2	1.3	186	186.7	188
2	0.0139	2	2.2	186	185.8	188
3	0.0202	3	3.1	184	183.9	187
4	0.0284	2	4.5	186	183.5	188
5	0.0415	4	6.4	183	180.6	187
6	0.0649	13	9.8	175	178.2	188
7	0.1075	19	15.8	169	172.2	188
8	0.1916	26	26.8	161	160.2	187
9	0.3591	42	48.7	146	139.3	188
10	0.9968	114	108.3	73	78.7	187

```
number of observations =      1876
      number of groups =        10
Hosmer-Lemeshow chi2(8) =      6.64
           Prob > chi2 =      0.5760
```

ROCとAUC

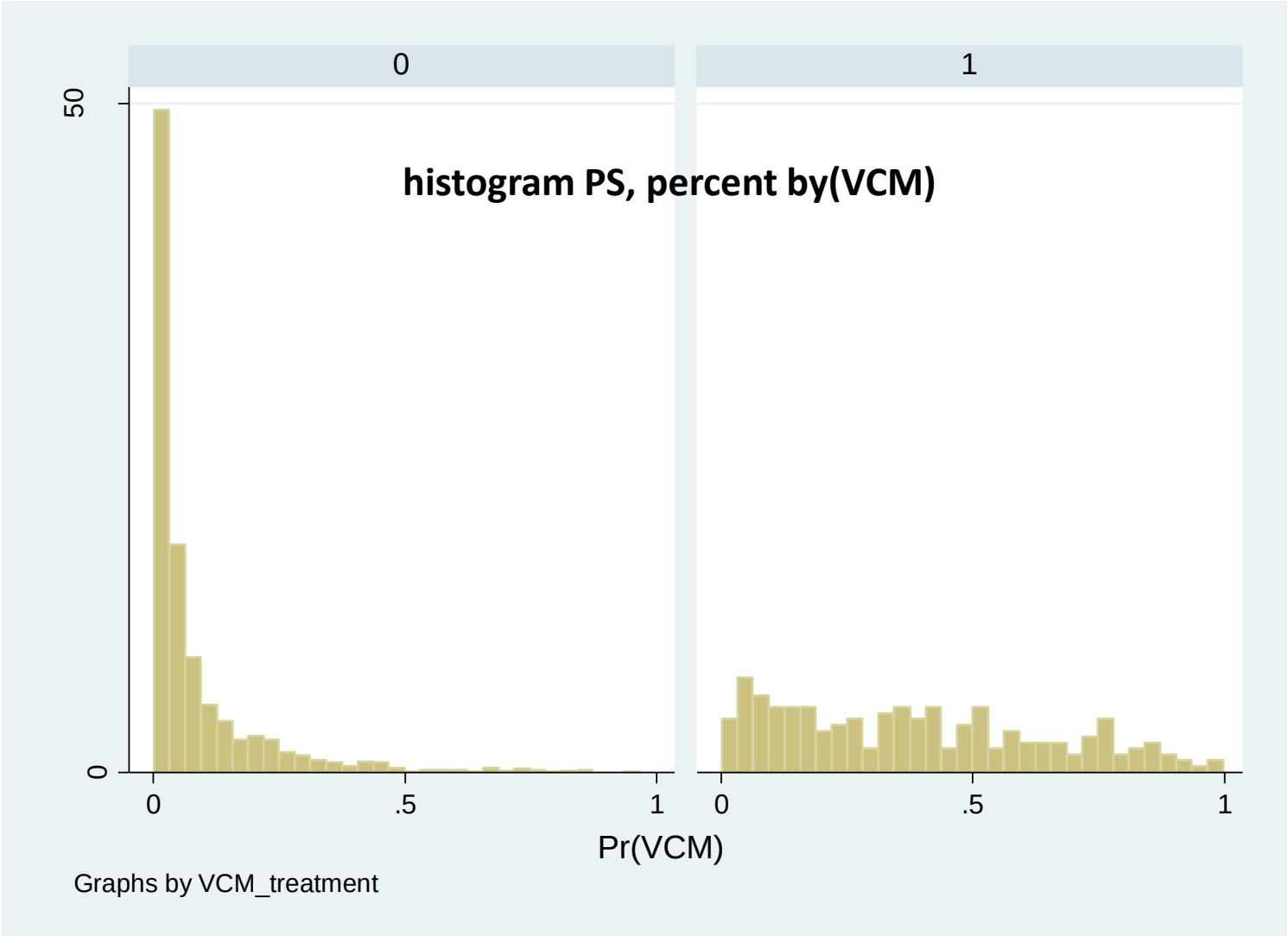


. lroc

Logistic model for VCM

number of observations = 1876
area under ROC curve = 0.8740

Histogram of Propensity Score

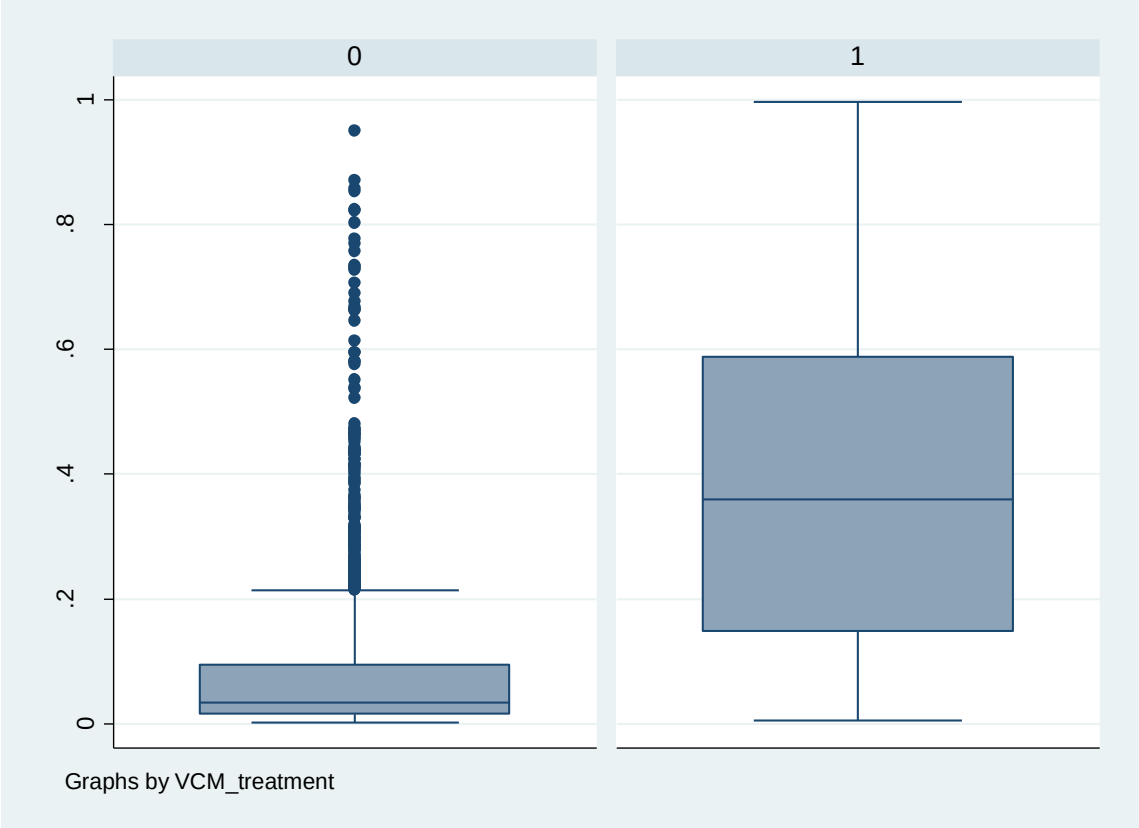


VCM -

VCM +

Box plot of PS by VCM

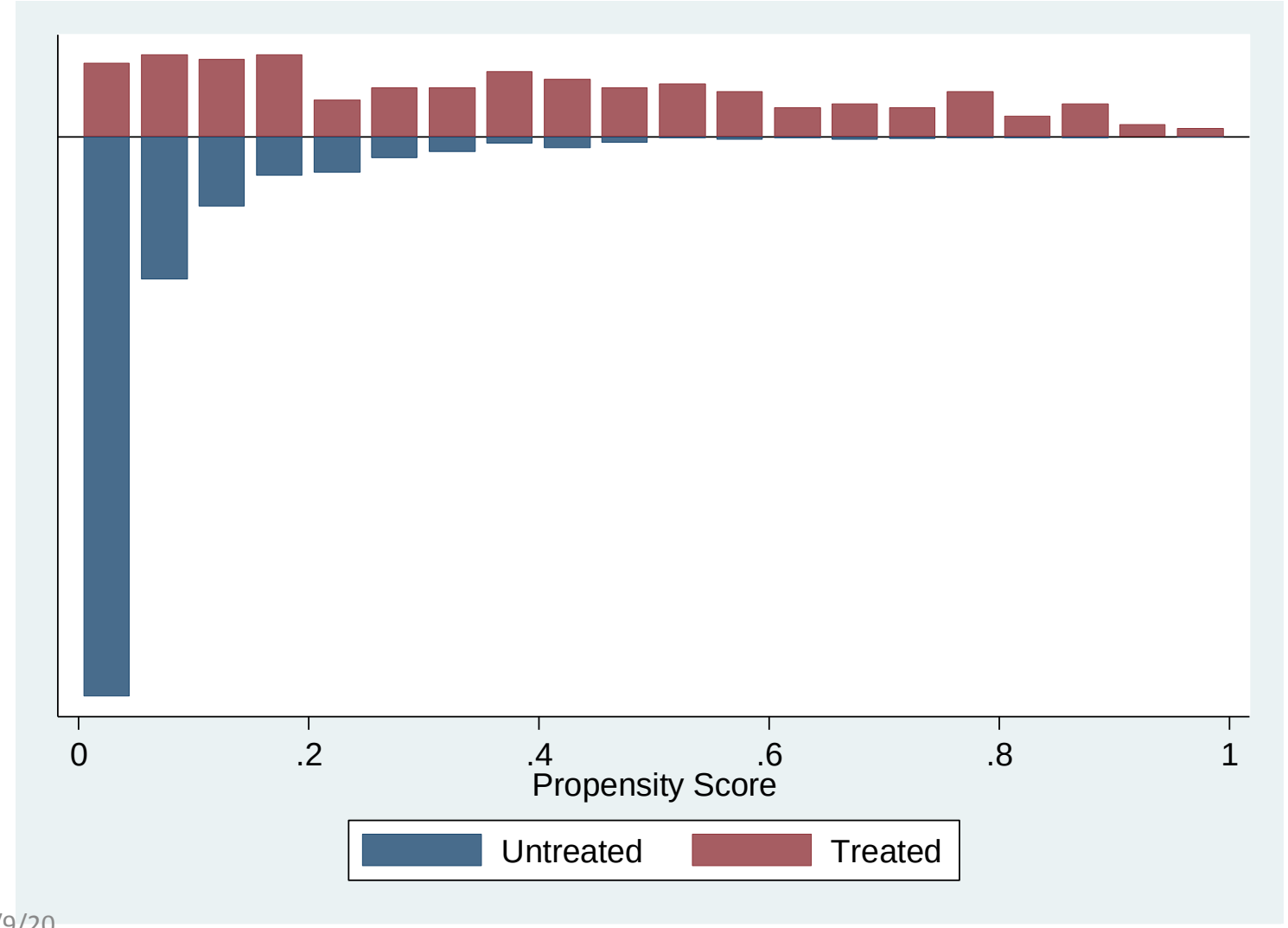
graph box PS, by(VCM)



VCM -

VCM +

Stata command: psgraph

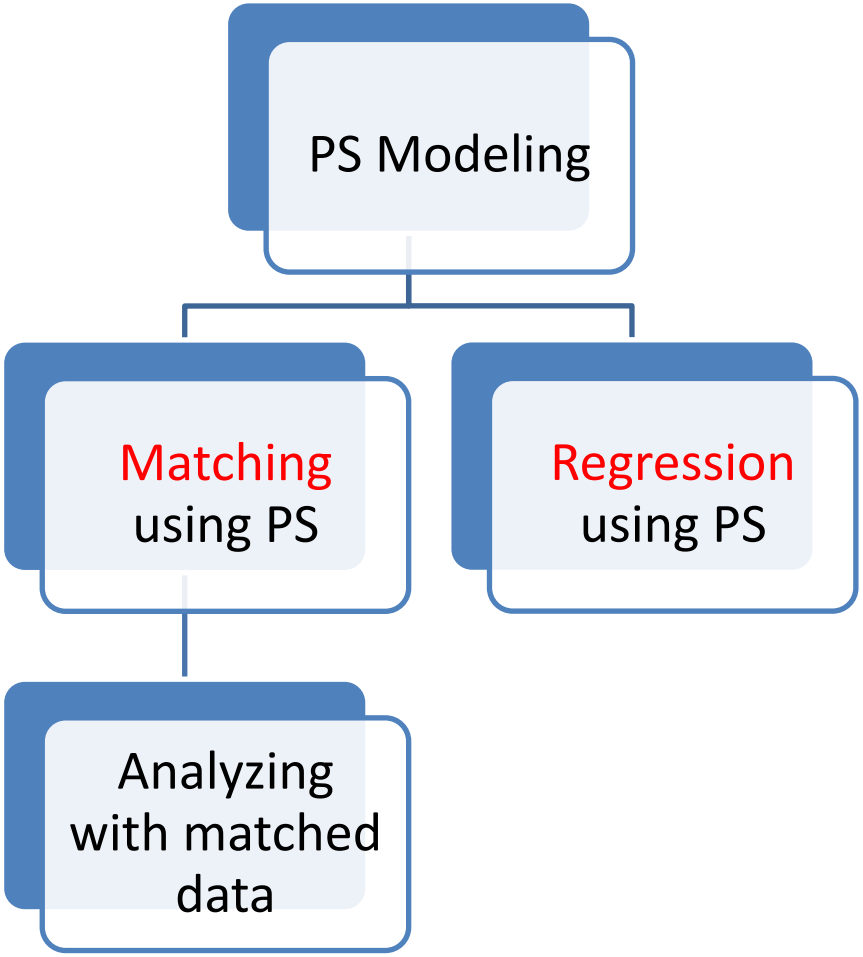


Propensity Score (PS) Analysis

1st step

2nd step

3rd step



PS matching using STATA

PS model

- Summary data: d, sum
- Modeling: Sw,pe()pr():logistic, Model checking: lroc, estat
- Overall PS checking: sum, graph (hist, graph box), psgraph
- Checking each variable: quintile (regress, logistic, ttest, tabulate, pstest)

Matching

- Psmatch2, pstest, psgraph
- teffects psmatch (Ver. 13)

Analyzing

- Regreession (clogit after reshape)
- clogit with other confounders

Matching in Stata

- psmatch2 package created by Edwin Leuven and Barbara Sianesi.

- Matching: psmatch2 implements various types of propensity score matching estimators.

  - one-to-one, *k*-nearest neighbors, radius, kernel,local linear regression, spline, Mahalanobis

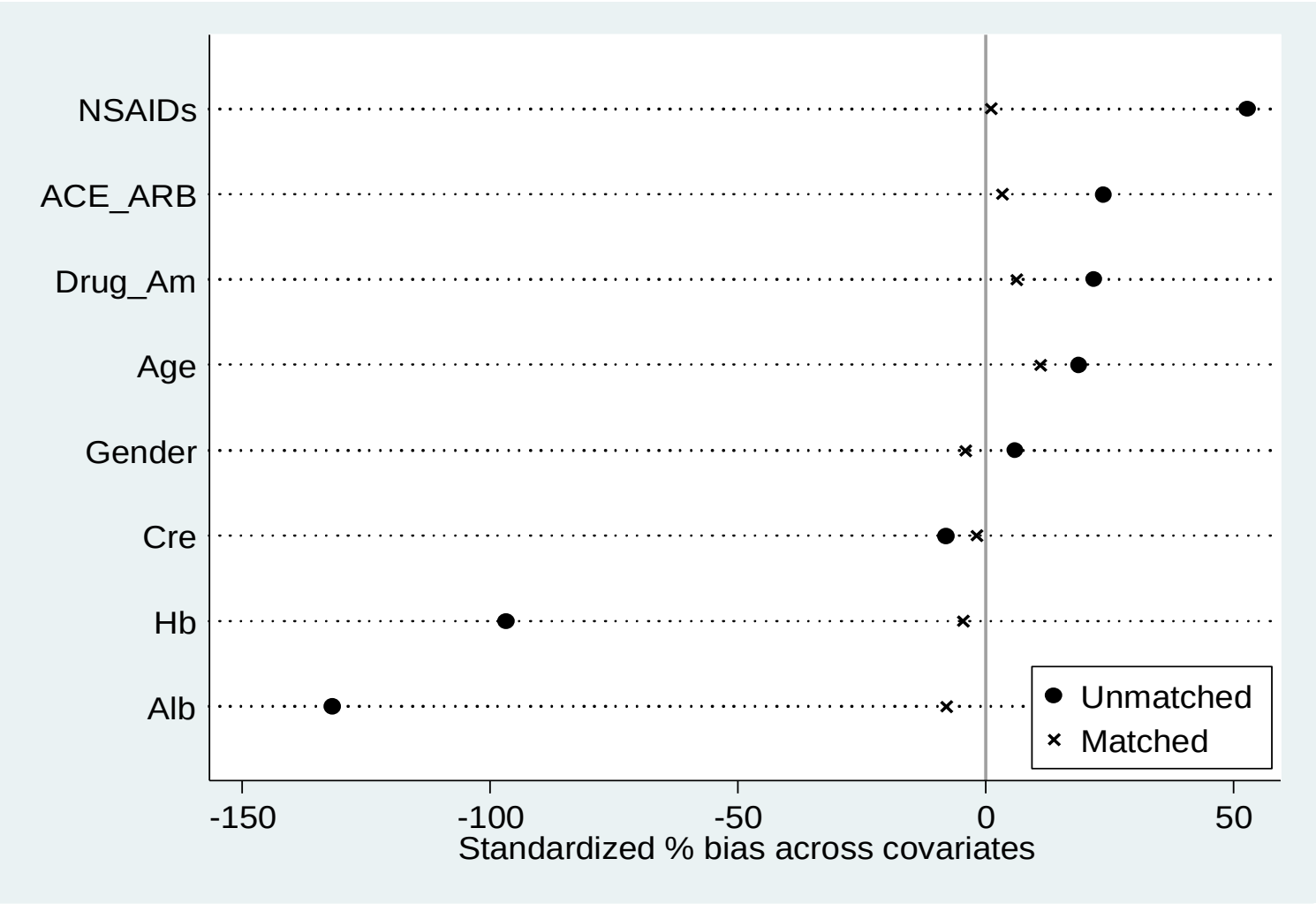
Nearest Neighbor Matching within a caliper

- Typically, one treatment case is matched to several control cases, but one-to-one matching is also common and may be preferred (Glazerman, Levy, and Myers 2003).
  - Caliper: (SD of log(PS/1-PS))x0.25

Stata command

- Matching
  - psmatch2 VCM, pscore(logitPS) caliper (0.4) noreplacement descending
- Checking the balance after matching
  - pstest Gender Age Cre Hb Alb Drug_Am NSAIDs ACE_ARB, both t(VCM) graph

Checking the balances after matching



Analysis after matching

- Conditional logistic analysis
  - clogit aki pair, group(group)
  - clogit,or
- McNemar test

PS regression using STATA

PS model

- **Summary data: d, sum**
- **Modeling: Sw,pe()pr():logistic, Model checking: lroc, estat**
- **Overall PS checking: sum, graph (hist, graph box), psgraph**
- **Checking each variable: quintile (regress, logistic, ttest, tabulate, pstest)**

Analyzing

- **Regression (tabulate, K-M curve, ttest)**
- **Multivariate regression (Logistic, Cox, regress with other confounders**

1. Crude odds ratio

```
. logistic aki_outcome_doubling VCM, or

Logistic regression                               Number of obs   =       1992
                                                  LR chi2(1)      =      96.88
                                                  Prob > chi2     =     0.0000
Log likelihood = -578.67306                       Pseudo R2       =     0.0772

------------------------------------------------------------------------------
aki_outcome_doubling | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
---------------------+--------------------------------------------------------
                 VCM |   6.111996   1.051483    10.52   0.000     4.362604    8.562889
               _cons |   .0729483   .0068979   -27.69   0.000     .0606076    .0878018
------------------------------------------------------------------------------
```

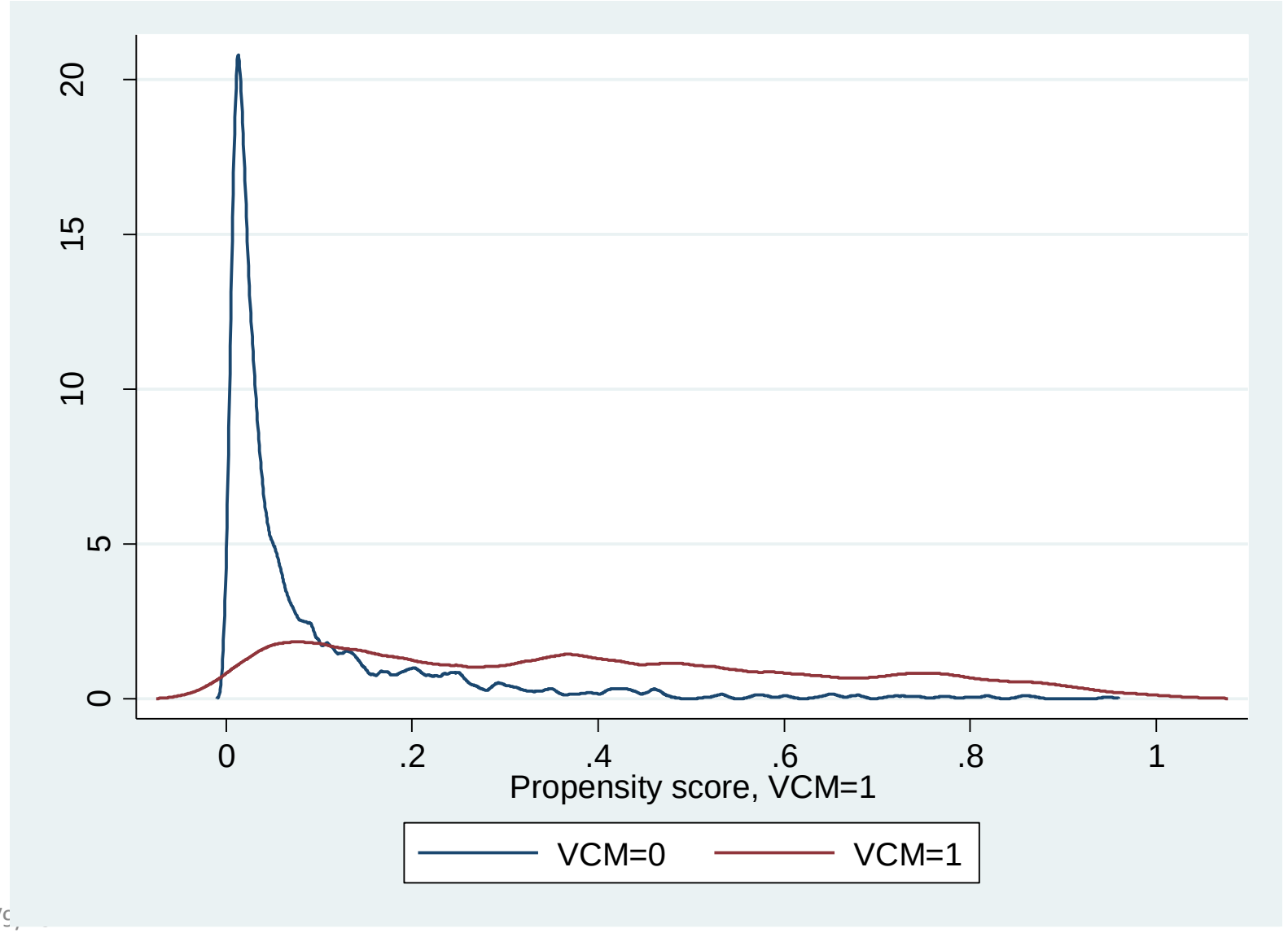
2. Adjusted odds ratio by Propensity score

```
. logistic aki_outcome_doubling VCM PS , or

Logistic regression                               Number of obs   =       1896
                                                  LR chi2(2)      =     138.76
                                                  Prob > chi2     =     0.0000
Log likelihood = -545.65981                       Pseudo R2       =     0.1128

------------------------------------------------------------------------------
aki_outcome_doubling | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
---------------------+--------------------------------------------------------
                 VCM |   2.495089   .5405971     4.22   0.000     1.631778    3.815146
                  PS |   15.61008   6.096279     7.04   0.000     7.260751    33.56054
               _cons |   .0570277   .0062395   -26.18   0.000     .0460208     .070667
------------------------------------------------------------------------------
```

Propensity scores – command summary

- *psmatch2*
- *pstest*
- *psgraph*
- *pscore*
- *teffects psmatch (stata13)*

Stata command: teffects psoverlap



まとめ

- 観察研究で治療効果を推定する際に有用
(特にイベント数が変数の7倍以下の場合）

- PS手法は2段階
  - 治療を従属変数、共変量を独立変数とするロジスティック回帰によって*PSを推定する*
  - PSを基に様々な方法で分析（マッチング、層別化、多変量解析、ウェイト付け）

- 共変量*のバランス*をチェックすることが重要

References

- Cepeda MS, Boston R, Farrar JT, Strom BL. Comparison of logistic regression versus propensity score when the number of events is low and there are multiple confounders. *Am J Epidemiol* 2003; 158: 280-287.
- Connors Jr AF, Speroff T, Dawson NV, et al. The effectiveness of right heart catheterization in the initial care of critically ill patients. *JAMA* 1996; 276: 889-897.
- D'Agostino Jr, RB. Tutorial in biostatistics: propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Stat Med* 1998; 17: 2265-2281.
- Gum PA, Thamilarasan M, Watanabe J, Blackstone EH, Lauer MS. Aspirin use and all-cause mortality among patients being evaluated for known or suspected coronary artery disease. *JAMA* 2001; 286: 1187-1194.
- Harrell FE, Lee KL, Matchar DB, Reichart TA. Regression models for prognostic prediction: advantages, problems, and suggested solutions. *Cancer Treatment Reports* 1985: 69: 1071-1077.
- Kurth T, Walker AM, Glynn RJ, Chan KA, Gaziano JM, Berger K, Robins JM. Results of multivariable logistic regrssion, propensity matching, propensity adjustment, and propensity-based weighting under conditions of nonuniform effect. *Am J Epidemiol* 2006; 163: 262-270.
- Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. *J Clin Epidemiol* 1996; 49: 1373-1379.
- Robins JM, Hernan MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology* 2000; 11: 550-560.
- Rosenbaum PR. *Observational Studies*. New York, NY: Springer-Verlag, 2002.
- Rosenbaum PR, Rubin DB. The central rol of the propensity score in observational studies for causal effects. *Biometrika* 1983; 70: 41-55.
- Rubin DB. Estimating causal effects from large data sets using propensity scores. *Annals of Internal Medicine* 1997; 127: 757-763.
- Sato T, Matsuyama Y. Marginal structural models as a tool for standardization. *Epidemiology* 2003; 14: 680-686.
- Propensity Score Analysis Shenyang guo et al. SAGE 2010
- 傾向スコアを用いた共変量調節による因果効果の推定と臨床医学・疫学・公衆衛生分野での応用について　星野崇宏　他　J Natl. Inst.Public Health 2006; 55(3): 230-242
- 調査観察データの統計科学　　星野　崇宏　岩波書店