

# SASによる傾向スコアマッチング

○魚住 龍史<sup>1</sup> 矢田 真城<sup>2</sup> 山本 倫生<sup>3</sup> 川口 淳<sup>4</sup>

<sup>1</sup>京都大学大学院医学研究科 医学統計生物情報学

<sup>2</sup>東京理科大学大学院 工学研究科 経営工学専攻

<sup>3</sup>岡山大学大学院 環境生命科学研究科

<sup>4</sup>佐賀大学医学部附属 地域医療科学教育研究センター

## Propensity score matching using SAS

Ryuji Uozumi<sup>1</sup>, Shinjo Yada<sup>2</sup>, Michio Yamamoto<sup>3</sup>, and Atsushi Kawaguchi<sup>4</sup>

<sup>1</sup>*Department of Biomedical Statistics and Bioinformatics, Kyoto University Graduate School of Medicine*

<sup>2</sup>*Department of Management Science, Graduate School of Engineering, Tokyo University of Science*

<sup>3</sup>*Graduate School of Environmental and Life Science, Okayama University*

<sup>4</sup>*Faculty of Medicine, Saga University*

## 要旨

臨床研究では、ランダム化比較試験によって得られたデータの場合、群間比較を行う際の比較可能性が担保される。しかし、ランダム化比較試験でなく、観察研究によって得られたデータの場合、比較可能性が担保されず、群間比較としてはバイアスが混在した結果となってしまう場合がある。比較可能性を担保する方法のひとつとして、傾向スコアによるマッチングがよく行われている。これまで、SAS/STAT で提供されている LOGISTIC プロシジャで傾向スコアを算出して、マッチングに関しては DATA ステップでプログラムを構築しなければならなかった。しかし、SAS/STAT 14.2 (SAS 9.4 TS1M4) では、新たに PSMATCH プロシジャが追加され、傾向スコアによる分析を行うためのさまざまなツールが提供された。本稿では、PSMATCH プロシジャを用いて、傾向スコアによるマッチングを行う方法に関して解説を行い、JMP や R を用いて実施する場合との比較に関しても報告する。

キーワード：観察研究 傾向スコア 平均治療効果 ATE ATT マッチング 最近傍マッチング PSMATCH キャリパー PSMODEL MATCH ASSESS ODS GRAPHICS 箱ひげ図 度数分布 クラウドプロット

## 1 はじめに

臨床研究では、ランダム化比較試験によって得られたデータの場合、群間比較を行う際の比較可能性が担保される。しかし、ランダム化比較試験でなく、観察研究 (observational study) によって得られたデータの場

合，比較可能性が担保されず，群間比較としてはバイアスが混在した結果となってしまう．比較可能性を担保する方法として，マッチング，層別解析，および回帰モデルによる解析などが挙げられる (Faries et al., 2010).

近年，観察研究において傾向スコアによる解析がよく行われている．著者らも，臨床研究者から傾向スコアマッチングや重み付きの調整解析に関する統計相談に対応した経験を有している．本稿では，傾向スコアマッチング (propensity score matching) による調整解析に焦点を当てることとする．観察研究において，試験治療と対照治療の比較を行うと仮定すると，背景因子で条件付けたもと，試験治療を受ける確率が傾向スコアである．SAS/STAT では，LOGISTIC プロシジャにおける OUTPUT ステートメントのオプションを指定すれば，ロジスティックモデルによる予測確率が出力される．すなわち，応答変数として治療を表す変数を指定すれば，傾向スコアを出力させることが可能である．しかし，傾向スコアによるマッチングを行う場合，SAS/STAT のプロシジャで処理することはできず，DATA ステップでマッチングを行うためのプログラムを構築しなければならなかった．このため，傾向スコアによるマッチングを実施している臨床研究では，SAS 以外の統計解析ソフトとして，例えば R のパッケージ “Match” を用いた結果が報告されている (Hamatani et al., 2016; Hida et al., 2017)．また，JMP のアドインを用いた傾向スコアによるマッチング解析の結果も報告されている (Yamashita et al., 2017)．しかし，JMP のアドインを用いたマッチングでは，JMP スクリプト言語を用いない限りプログラムによる効率化ができない．加えて，JMP のアドインによる方法では，マッチングの条件指定の定義が特殊であることに留意しなければならない．したがって，SAS ユーザーであれば，SAS を用いて傾向スコアマッチングを行うことができれば，より効率的に実施できるといえる．

SAS/STAT 14.2 (SAS 9.4 TS1M4) では，新たに PSMATCH プロシジャが追加され，傾向スコアによる分析を行うためのさまざまなツールが提供された (SAS Institute Inc., 2016)．本稿では，PSMATCH プロシジャを用いて，傾向スコアによるマッチングを行う方法に関して解説を行い，JMP や R を用いて実施する場合との比較についても報告を行う．

## 2 傾向スコアとマッチング

本節では，ランダム化比較試験および観察研究における平均治療効果の推定について考えたもとで，傾向スコアおよびマッチングに関して解説する．

### 2.1 観察研究における平均治療効果の推定

2 値の指示変数  $Z$  として， $Z=1$  を治療群， $Z=0$  を対照群とする．さらに，潜在的結果変数 (potential outcome) として，治療群であった場合の結果変数  $Y(1)$ ，対照群であった場合の結果変数  $Y(0)$  のいずれかをとり，観察される結果変数は

$$Y = ZY(1) + (1-Z)Y(0) = \begin{cases} Y(0) & (Z=0) \\ Y(1) & (Z=1) \end{cases}$$

となる．このとき，

$$ATE = E[Y(1) - Y(0)] \tag{1}$$

は平均治療効果 (average treatment effect; ATE) は呼ばれ，

$$ATT = E[Y(1) - Y(0) | Z=1] \tag{2}$$

は治療群に対する平均治療効果 (average treatment effect on the treated; ATT) と呼ばれる．

ランダム化臨床試験のもとでは、 $\{Y(1), Y(0)\} \perp Z$ であり、ATEはATTに一致する。このとき、観測データから、ATEおよびATTは $E[Y(1)|Z=1] - E[Y(0)|Z=0]$ と表せ、平均治療効果の不偏推定を可能とする (Lunceford and Davidian, 2004).

一方、観察研究のもとでは、 $\{Y(1), Y(0)\} \perp Z$ が成り立たず、ATEはATTと異なる。さらに、結果変数は直接的に治療効果を推定できない。ここで、ATTは、

$$\begin{aligned} \text{ATT} &= E[Y(1) - Y(0) | Z = 1] \\ &= E[Y(1) | Z = 1] - E[Y(0) | Z = 0] + E[Y(0) | Z = 0] - E[Y(0) | Z = 1] \end{aligned}$$

と分解することができ、 $E[Y(1) | Z = 1] - E[Y(0) | Z = 0]$ は推定可能である一方、選択バイアスである  $E[Y(0) | Z = 0] - E[Y(0) | Z = 1]$ は推定不能である。観察研究において、治療効果を正しく推定する方法として、傾向スコアに基づく調整解析が挙げられる。

## 2.2 傾向スコアとモデルによる推定

$X_i$ を被験者 $i \in \{1, 2, \dots, N\}$ において観測される共変量ベクトル、 $Z_i$ を被験者 $i$ の群を表す指示変数とする。と、 $X_i$ で条件付けたもとで治療群である確率 $e_i = \Pr(Z_i = 1 | X_i)$ が傾向スコアと定義される (Rosenbaum and Rubin, 1983)。2群間で割付比が1:1のランダム化比較試験の場合、傾向スコアは $e_i = 1/2$ となる。また、傾向スコアはバランススコア (balancing score) とも呼ばれており、傾向スコアで条件付けたもとで、各群で観察される共変量の分布が似通っていることが望まれる ( $Z_i \perp X_i | e_i$ )。

傾向スコアは、治療群を表す変数を応答変数としたロジスティックモデルから推定され、SAS/STATではLOGISTICプロシジャによって算出可能である。ここで、群を表す変数をZ (Z=1:治療群, Z=0:対照群)、調整する共変量として、性別、年齢、BMIを表す変数をそれぞれGender, Age, Bmiとする。このとき、プログラム1を実行させることで傾向スコア $e_i$ を推定することが可能であり、傾向スコアを表す変数PSを含んだデータセットが出力される。傾向スコアを算出するためのロジスティックモデルは「傾向スコアモデル (propensity score model)」と呼ばれる。

プログラム1: LOGISTICプロシジャによる傾向スコアの算出プログラム例

---

```
proc logistic data=<入力データセット名>;
  class Gender;
  model z(event='1')= Gender Age Bmi;
  output out=<出力データセット名> pred=ps;
run;
```

---

なお、傾向スコア推定の際のロジスティックモデルの説明変数として、どのような変数を指定するかについてはコンセンサスが取れていない。Rosenbaum (2010)は、「傾向スコアによるマッチングでバランスを取りたい変数で考えるべきである。」と述べている。Austin et al. (2007)は、「ランダム化比較試験と同じように被験者背景の表を群別に作成し、アンバランスなベースラインの共変量を選択する。選択する共変量は重要な予後因子であるべきであり、この手順は傾向スコアによるマッチングでも用いると効果的である。」と述べている。

傾向スコア $e_i$ を用いて、マッチングや重み付きの調整解析を行うことが可能である。重み付きの調整解析に用いる「重み」として、傾向スコアの逆数で表されるIPTW (inverse probability of treatment weighting) やATT Weightingがよく用いられ、それぞれ

$$\text{IPTW} \quad : \quad w_{\text{IPTW},i} = \frac{Z_i}{e_i} + \frac{1-Z_i}{1-e_i} \quad (3)$$

$$\text{ATT Weighting} \quad : \quad w_{\text{ATT},i} = Z_i + \frac{e_i(1-Z_i)}{1-e_i} \quad (4)$$

として表される．なお，IPTW は ATE を推定するために，ATT Weighting は ATT を推定するために用いられる．SAS による傾向スコアを用いた重み付き解析として，IPTW 法による解析方法について，SAS ユーザー総会においても報告されている (古川, 2011)．

### 2.3 傾向スコアによるマッチング

傾向スコアによるマッチングでは，治療群と対照群における被験者のうち，傾向スコアが似通った被験者をペアとして，新たな解析対象集団を構成する (“matched sample” と呼ばれる)．マッチングにより解析対象集団が構成された後，群間の治療効果を推定すれば，ATT の推定を可能とする．なお，本稿では，治療群よりも対照群の被験者数が多い状況を想定する．

傾向スコアによるマッチングを行うためには，被験者の傾向スコア間の差異を測るための尺度が必要となる．その理由は，傾向スコアモデルに連続変数を含めた場合，ある被験者と全く同一の傾向スコアを持つ被験者を見つけることは困難であるためである．被験者  $i$  と被験者  $j$  の傾向スコア間の差異を測る尺度として，傾向スコア自体の差  $|e_i - e_j|$  や，傾向スコアのロジット変換した値の差  $|\text{logit}(e_i) - \text{logit}(e_j)|$  が用いられる．

傾向スコアによるペアを構成するために，よく用いられる方法として，最近傍マッチング (greedy nearest neighbor matching) が挙げられる (Rosenbaum and Rubin, 1985)．最近傍マッチングでは，治療群の被験者に対して最も差異の小さい被験者を対照群から選択してペアを構成する．例えば，治療群 1 名に対して対照群から 1 名を選択する方法は 1:1 マッチングと呼ばれる．対照群から  $k$  名を選択する  $k:1$  マッチングとして実施することも可能である．

最近傍マッチングによってペアを構成する場合，特定のキャリパー (caliper) を指定し，キャリパーの値に基づいて算出される幅以下であればペアの構成が行われる．傾向スコアマッチングの場合，キャリパーの値として，傾向スコアの推定値をロジット変換した値の標準偏差に 0.25 を掛けた値を推奨する報告がある (Rosenbaum and Rubin, 1985)．近年の報告としては，傾向スコアの推定値をロジット変換した値の標準偏差に 0.20 を掛けた値を用いることで，多くの状況下でバイアスが小さくなることが示されている (Austin, 2011)．キャリパーが傾向スコアの推定値をロジット変換した値の標準偏差に 0.20 を掛けた値で設定された事例は多く報告されており，例えば Bangalore et al. (2015) とその解説記事として田栗 (2017) を参照されたい．ただし，実際にマッチング実施後に，傾向スコアで調整した変数の分布に群間でバイアスがなくなったか確認することが重要であり，バイアスが生じたままである場合には，キャリパーの値を再考する必要がある (山本・森田, 2015)．

SAS を用いて傾向スコアマッチングを行う場合，プロシジャでサポートされていない分析手法であるため，DATA ステップで最近傍マッチングのアルゴリズムを組まなければならなかった．DATA ステップで傾向スコアマッチングを行うための SAS マクロが報告されているが (Coca-Perraillon, 2007; Lanehart et al., 2012)，キャリパーの設定値が傾向スコアの推定値をロジット変換した値の標準偏差を掛けた値となっていない．そこで，3 節で解説するプロシジャの動作を理解する手助けとなるよう，過去に報告された SAS マクロを改良して，SAS の DATA ステップ上で，傾向スコアによる最近傍マッチングを行うための SAS マクロを開発した．詳細は，付録 A を参照されたい．

## 2.4 JMP を用いた傾向スコアによる 1:1 最近傍マッチング

SAS の代わりに JMP で分析を行う場合、次の URL からアドインをダウンロードすれば、傾向スコアによる 1:1 最近傍マッチングを行うことができる。ただし、JMP 11.0 以上のバージョンでなければならない。

アドイン入手 URL : <http://www.jmp.com/content/dam/jmp/documents/jp/support/propensityaddin.zip>

傾向スコアによる 1:1 最近傍マッチングを行うにあたっては、事前に傾向スコアを表す変数もデータセットに用意する必要がある。SAS を用いる場合は LOGISTIC プロシジャの OUTPUT ステートメント、JMP を用いる場合は「名義ロジスティックのあてはめ」のオプション「確率の計算式の保存」で対応できる。

以下の手順で 1:1 最近傍マッチングを行うことができる。

1. [アドイン] → [最近傍マッチング] → [最近傍マッチングの実行] を選択して、図 1 の入力画面を表示させる。
2. 「Y, 応答変数」 に適当な変数を指定する。
3. 「T, 処置変数 (2水準)」に群を表す変数を指定する。
4. 「マッチングに使うスコア (数値)」に傾向スコアを表す変数を指定する。
5. 「乱数シード値」を設定する。
6. 「Caliper 係数 : a」を設定する。
7. 「元データにマッチング結果の列を追加する」にチェックする。
8. 「OK」をクリックして出力結果を示す。

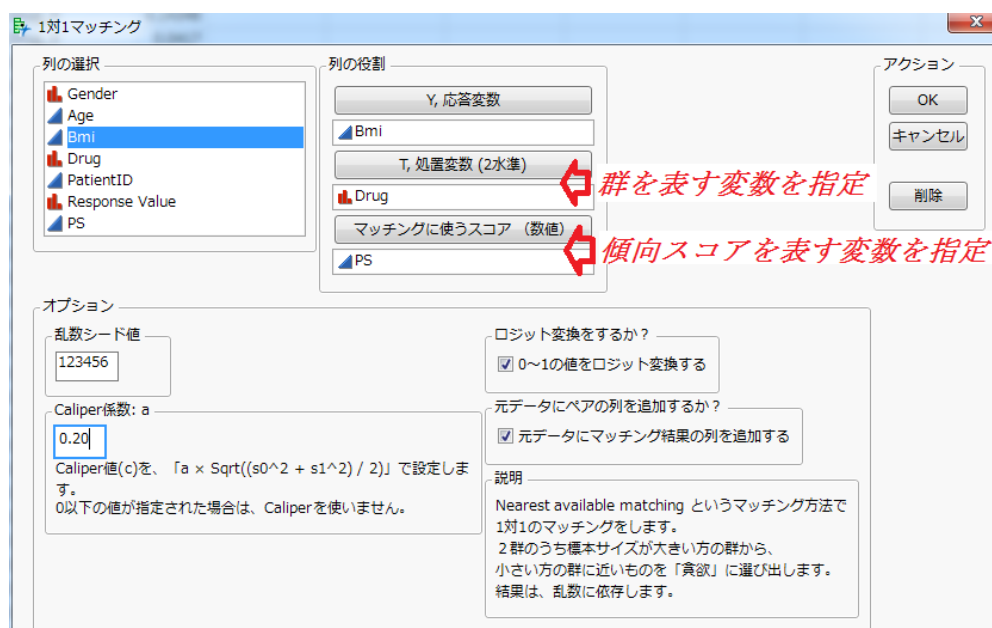


図 1 : JMP アドインによる最近傍マッチング実行画面

なお、手順 6 におけるキャリパーの値は 0.20 とすることが推奨されている (JMP ジャパン事業部, 2014)。ただし、Austin (2011) は傾向スコアの推定値のロジット変換後の標準偏差を 0.20 倍した値をキャリパーとして利用することが推奨している一方、JMP アドインの実行画面で指定するキャリパーの値は群別に求めた不偏分散の平均値の平方根である。そのため、手順 6 のキャリパーの値を 0.20 と指定することは、Austin (2011) によって推奨されている 0.20 とは異なる。

以上の手順で、マッチングの結果として、マッチング前後におけるロジット変換後の傾向スコアの分布を確認でき、元データのすべてのオブザベーションに対して、「マッチした」あるいは「マッチなし」を表す変数が追加される。

SAS プログラムに関する素養がなく、JMP で統計解析を行っているユーザーであれば、上記の方法によって、傾向スコアマッチングを行う選択肢もある。ただし、JMP のアドインでは 1:1 マッチングのみで、k:1 マッチングはサポートされていない。また、SAS に限らず、例えば R プログラムに関する素養があれば、R のパッケージ “Match” を用いて、傾向スコアマッチングを行うことが可能である。付録 B には、R のパッケージ “Match” を用いて、傾向スコアによる最近傍マッチングを行うための詳細を示している。しかし、SAS ユーザーであれば、SAS 上の DATA ステップあるいは SAS/STAT のプロシジャを用いて傾向スコアマッチングを行いたいと考えるであろう。

### 3 PSMATCH プロシジャを用いた傾向スコアによるマッチング

傾向スコアマッチングは、SAS/STAT のプロシジャでサポートされていない分析手法であるため、SAS で対応するためには、DATA ステップで最近傍マッチングのアルゴリズムを組まなければならなかった。しかし、SAS/STAT 14.2 (SAS 9.4 TS1M4) から PSMATCH プロシジャが追加され、傾向スコアマッチングができるようになった (SAS Institute Inc., 2016)。PSMATCH プロシジャを用いれば、傾向スコア推定後の最近傍マッチングを行う場合に、前節で述べたような JMP あるいは R にデータセットを移動させて、マッチングを行う必要がなくなる。また、SAS の DATA ステップでマッチングのアルゴリズムを組むことなく、傾向スコアの推定からマッチングまで、SAS のプロシジャですべて完結できる。

本節では、PSMATCH プロシジャを用いて、傾向スコアによるマッチングをどのように実施するか解説し、出力結果の解釈を述べる。傾向スコアによるマッチングを行うための対象データとして、観察研究のデータ (データセット名: Drugs) を用いる ( $n=486$ )。データセット Drugs は、新規治療を行った「治療群」(変数: Drug = 'Drug\_X') と標準治療を行った「対照群」(変数: Drug = 'Drug\_A') を比較するために収集されたデータであるが、被験者あるいは主治医がどちらの治療を行うか選択できたため、観察研究のデータとして扱う。ベースラインの被験者背景として、年齢 (変数名: Age)、性別 (変数名: Gender; Gender = 'Male'「男性」, 'Female'「女性」), BMI (変数名 Bmi) が挙げられる。

3.1 節から 3.3 節では、PSMATCH プロシジャの主なステートメントによる機能を詳述する。図 2 は、傾向スコアマッチングの手順における PSMATCH プロシジャの主なステートメントの役割を概念図として示している。

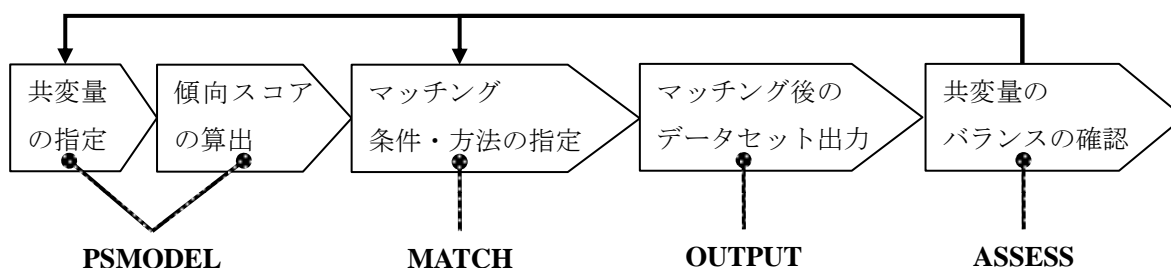


図 2 : 傾向スコアマッチングの手順における PSMATCH プロシジャの主なステートメントの役割

### 3.1 PSMODEL ステートメントによる傾向スコアモデル

LOGISTIC プロシジャでサポートされている MODEL ステートメントと同じように、PSMODEL ステートメントで傾向スコアモデルを指定すれば、傾向スコアの推定結果が出力される (プログラム 2)。

プログラム 2 : PSMATCH プロシジャにおける PSMODEL ステートメントの利用例

```
proc psmatch data=drugs region=allobs;  
  class z Gender;  
  psmodel z(treated='1')= Gender Age Bmi;  
  output out(obs=all)=psout00 ps=ps attwgt=attw atewgt=atew;  
run;
```

- PROC PSMATCH ステートメント

PSMATCH プロシジャでは、REGION = で層別やマッチングを行う場合のオブザベーションのサポート範囲を指定できる。プログラム 2 では、簡便に REGION = ALLOBS と指定し、すべてのオブザベーションをマッチングのサポート範囲として取り扱っている。デフォルトは REGION = CS であり、治療群および対照群における傾向スコアについて、共通のサポート (common support) 範囲が選択される。

- PSMODEL ステートメント

傾向スコアモデルの指定を行っている。LOGISTIC プロシジャなどの MODEL ステートメントにおける指定と異なり、PSMODEL ステートメントの傾向スコアモデルで指定する応答変数も CLASS ステートメントに指定する必要がある。また、応答変数の水準指定には、TREATED = “治療群の水準” を指定する。なお、LOGISTIC プロシジャなどで予め傾向スコアを含んだデータセットを用いれば、PSMODEL ステートメントの代わりに PSDATA ステートメントを指定する方法もある。

- OUTPUT ステートメント

プログラム 2 では、OUTPUT ステートメントにおいて OUT = PSOUT00 と指定し、出力させるデータセット名を指定している。ここで、PSMATCH プロシジャでは、マッチング後の対象データを出力データセットとして保存することもできるため、OUT = のオプションとして OBS = が指定できる。OBS = ALL と指定すると、PROC PSMATCH ステートメントの DATA = で指定したデータセットすべてが出力される。

表 1 : プログラム 2 を実行させて得られたデータセット PSOUT00 (最初の 3 オブザベーション)

Gender	Age	Bmi	Drug	PatientID	PS	ATEW	ATTW
Male	29	22.02	Drug_X	284	0.36444	2.744	1
Male	45	26.68	Drug_A	201	0.22296	1.2869	0.28694
Male	42	21.84	Drug_A	147	0.11323	1.1277	0.12768

さらに、OUTPUT ステートメントにおいて、PS = を指定すれば、指定した変数名で傾向スコアの推定値が出力データセットに含まれる。また、ATEWGT = や ATTWGT = を指定すれば、式 (3) および (4) で示した IPTW および ATT Weighting が指定した変数名で出力データセットに含まれる (表 1)。

### 3.2 MATCH ステートメントによる傾向スコアマッチング

PSMATCH プロシジャにおいて、MATCH ステートメントを指定すれば、傾向スコアによるマッチングができる。図3は、PSMATCH プロシジャでサポートされているマッチング方法である。デフォルトは METHOD = OPTIMAL である。入力データセットとキャリパーの値の指定によっては、マッチングができない場合もある。例えば、キャリパーの値が小さすぎると、ペアの構成数が少なくなってしまうため、治療群と同じ数の対照群をペアとして構成することが困難となる。このような場合に METHOD = OPTIMAL を指定して Fixed Ratio Matching を実行すると、”A feasible optimal fixed ratio matching that has the specified parameters does not exist.” というエラーが出て、マッチングできない。

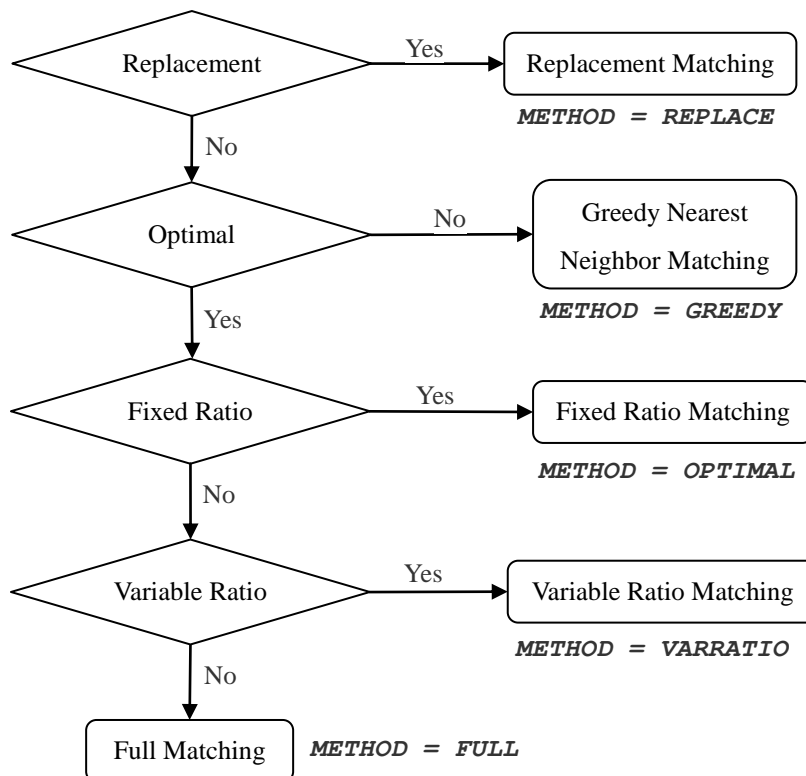


図3: MATCH ステートメントで指定可能なマッチング方法

データセット Drugs に対する傾向スコアマッチングとして、プログラム3は、前節で取りあげた最近傍マッチング “Greedy Nearest Neighbor Matching” を行うためのプログラムである。MATCH ステートメントにおいて METHOD = GREEDY を指定したプログラム3を実行すると、表2の出力データセット PSSORT00 が得られる。オプションで k=1 を指定することで 1:1 マッチングを行っている。また、ORDER = DESCENDING を指定することで、治療群における傾向スコアが大きい順に対照群のペアを探している。ORDER = RANDOM(SEED = 乱数シード) と指定すれば、乱数シードに基づき治療群における傾向スコアのランダムな順で対照群のペアを探すことが可能である。なお、デフォルトは k=1, ORDER = DESCENDING である。

- キャリパーの指定

プログラム3では、キャリパーとして CALIPER = 0.20 と指定している。STAT = LPS および CALIPER のオプションとして MULT = STDDEV と指定することで、キャリパーの値としては、推定された傾向スコアのロ



ジット変換した値の標準偏差に 0.20 を掛けた値を用いることを意味しており，Austin (2011) において推奨されている値に相当する．なお，STAT=LPS および MULT=STDDEV は指定しなくてもデフォルトとして実行される．

プログラム 3 : PSMATCH プロシジャによる最近傍マッチング

```

proc psmatch data=Drugs region=allobs;
  class Drug Gender;
  psmodel Drug(treated='Drug_X')= Gender Age Bmi;
  match method=greedy(k=1 order=descending)
        stat=lps caliper(mult=stddev)=0.20;
  output out(obs=match)=psmatch00 ps=ps lps=lps
        matchwgt=match matchid=matchsort;
run;
/* ペアの確認のために並び替え */
proc sort data=psmatch00 out=pssort00;
  by descending ps;
run;

```

● OUTPUT ステートメント

プログラム 3 のように，OBS=MATCH と指定することで，マッチングされたオブザベーションのみを含んだデータセットが出力される．なお，OBS=ALL と指定した場合，MATCHWGT= で指定した変数で，マッチングされたオブザベーションを識別することが可能である．プログラム 3 の場合，マッチングされたオブザベーションのみが変数 MATCH=1 と出力される．

さらに，MATCHID= で指定した変数 MATCHSORT で，マッチングした順およびペアの構成を確認することができる．そのために，出力データセットを傾向スコア順に並び替えたデータセット PSSORT00 で確認すると，最初の 5 オブザベーションは表 2 のようになる．傾向スコアの大きい順にマッチングが行われ，被験者番号 (変数名 : PatientID) として 87 と 123，200 と 349 のペアが構成されたことがわかる．

表 2 : プログラム 3 を実行させて得られたデータセット PSSORT00 (最初の 5 オブザベーション)

Gender	Age	Bmi	Drug	PatientID	PS	match	matchsort
Male	27	25.87	Drug_X	87	0.64115	1	1
Male	27	25.76	Drug_A	123	0.63513	1	1
Female	30	27.75	Drug_A	200	0.61831	1	2
Male	33	28.06	Drug_X	349	0.60481	1	2
Male	27	25.14	Drug_X	114	0.60047	1	3

### 3.3 ASSESS ステートメントによる変数の群間差の評価

PSMATCH プロシジャの ASSESS ステートメントで、すべての解析対象データセット、あるいはマッチされた対象データセットに対して、指定した変数の群間差を評価できる。

プログラム 4 では、MATCH ステートメントに関して 3.2 節と同様の指定を行ったもとの、ASSESS ステートメントで LPS を指定して、傾向スコアをロジット変換した値の群間差を評価する。もし PS も指定すれば、傾向スコアの群間差も評価可能である。さらに、VAR = で指定した変数についても、群間差の評価を行う。また、VARINFO オプションを指定することで、各群に対する変数情報を出力させる (表 4)。もし WEIGHT = オプションで重みを指定すれば、式 (3), (4) で示した IPTW, ATT Weighting を用いた重み付きで変数の評価を行う。プログラム 4 では、最後に、PLOTS = で ODS GRAPHICS による出力結果を得ている。

プログラム 4 : PSMATCH プロシジャにおける ASSESS ステートメントの利用例

```
ods graphics on;
proc psmatch data=drugs region=allobs;
  class Drug Gender;
  psmodel Drug(treated='Drug_X')= Gender Age Bmi;
  match method=greedy(k=1 order=descending)
        stat=lps caliper(mult=stddev)=0.20;
  assess lps var=(Gender Age Bmi)
        / weight=none varinfo plots=(all);
run;
```

表 3 : 傾向スコアの情報

Propensity Score Information						
Observations	Treated (Drug = Drug_X)			Control (Drug = Drug_A)		
	N	Mean	Std Dev	N	Mean	Std Dev
All	113	0.310773	0.132467	373	0.208801	0.131969
Region	113	0.310773	0.132467	373	0.208801	0.131969
Matched	113	0.310773	0.132467	113	0.308755	0.130649

表 3 では、傾向スコアの情報を示しており、紙面スペースの関係上、一部の統計量に関しては出力を省略している。Observation = Region である出力結果も示されているが、PROC PSMATCH ステートメントで REGION = ALLOBS を指定しているため、Observation = All と同様の結果となっている。次に、Observation = All と Observation = Matched の結果を比較すれば、データセットのうち、どの程度のオブザベーションでペアが構成されたか確認できる。データセット Drugs の場合、治療群 : 113 名, 対照群 : 373 名のうち、治療群 : 113 名, 対照群 : 113 名がマッチされたことがわかる。3.2 節のプログラム 3 を実行させて出力したデータセット PSMATCH00 とオブザベーション数が一致していることを確認できるであろう。

表 4 は VARINFO オプションを指定して出力される内容で、ASSESS ステートメントで指定した変数情報の要約統計量が出力される。性別のようなカテゴリカルデータの場合、Mean の列には、平均の代わりに、最初

の順番のカテゴリにおける割合が出力される。表 4 における性別 Gender では、Female「女性」の割合が出力されている。

表 4：変数情報

Variable Information							
Variable	Observations	Treated (Drug = Drug_X)			Control (Drug = Drug_A)		
		N	Mean	Std Dev	N	Mean	Std Dev
LPS	All	113	-0.880615	0.681761	373	-1.520586	0.844486
	Region	113	-0.880615	0.681761	373	-1.520586	0.844486
	Matched	113	-0.880615	0.681761	113	-0.889200	0.674687
Age	All	113	36.309735	5.534114	373	40.404826	6.579103
	Region	113	36.309735	5.534114	373	40.404826	6.579103
	Matched	113	36.309735	5.534114	113	35.884956	5.468664
Bmi	All	113	24.492566	1.863797	373	23.753271	1.980778
	Region	113	24.492566	1.863797	373	23.753271	1.980778
	Matched	113	24.492566	1.863797	113	24.309027	1.882123
Gender	All	113	0.433628	0.495575	373	0.458445	0.498270
	Region	113	0.433628	0.495575	373	0.458445	0.498270
	Matched	113	0.433628	0.495575	113	0.495575	0.499980

表 5 では、ASSESS ステートメントで指定した変数に対して、標準化した群間差が出力される。Mean Difference は単純な群間の平均値の差であり、Standardized Mean Difference は Mean Difference / Divisor の計算結果である。Divisor「除数」は各群の標準偏差をプールした標準偏差で構成される。例えば、年齢を表す変数 Age の All Obs「すべてのオブザベーション」を対象とした出力の場合、表 4 で求めた各群の標準偏差から、

$$\sqrt{\frac{(5.534114^2 + 6.579103^2)}{2}} = 6.079104$$

が Divisor となり、表 5 に出力された Mean Difference より、

$$\frac{\text{Mean Difference}}{\text{Divisor}} = \frac{-4.095091}{6.079104} = -0.673634$$

が Standardized Mean Difference として出力されている。

なお、ASSESS ステートメントで WEIGHT = NONE と指定したため、表 5 は重みを付けない計算結果が出力されている。例えば、WEIGHT = ATEWGT と指定すれば、式 (3) の IPTW で重み付けたした計算結果が出力される。また、ODS GRAPHICS ON の状態で実行しているので、表 5 で示した各変数に対する標準化した群間差に対応する図も出力される (付録 C)。付録 C では、±0.25 がカットオフ値として参照線が出力され、[-0.25, 0.25] であればバランスが取れているという目安として表している (SAS Institute Inc., 2016)。なお、カットオフ値として ±0.10 を推奨した報告も行われている (Normand et al., 2001)。

次に、ASSESS ステートメントで PLOTS = ALL を指定することで、ODS GRAPHICS による出力結果を得ることができる。ODS GRAPHICS による出力として、ASSESS ステートメントで指定した変数に対する群間の分布として、クラウドプロットが出力される (図 4)。

表 5 : 各変数に対する標準化した群間差

Standardized Variable Differences (Treated - Control)							
Variable	Mean Difference			Standardized Mean Difference			
	All Obs	Region Obs	Matched Obs	Divisor	Mean Difference		
					All Obs	Region Obs	Matched Obs
LPS	0.639971	0.639971	0.008584	0.767448	0.833894	0.833894	0.011185
Age	-4.095091	-4.095091	0.424779	6.079104	-0.673634	-0.673634	0.069875
Bmi	0.739296	0.739296	0.18354	1.923178	0.384414	0.384414	0.095436
Gender	-0.024817	-0.024817	-0.061947	0.496925	-0.049941	-0.049941	-0.124661

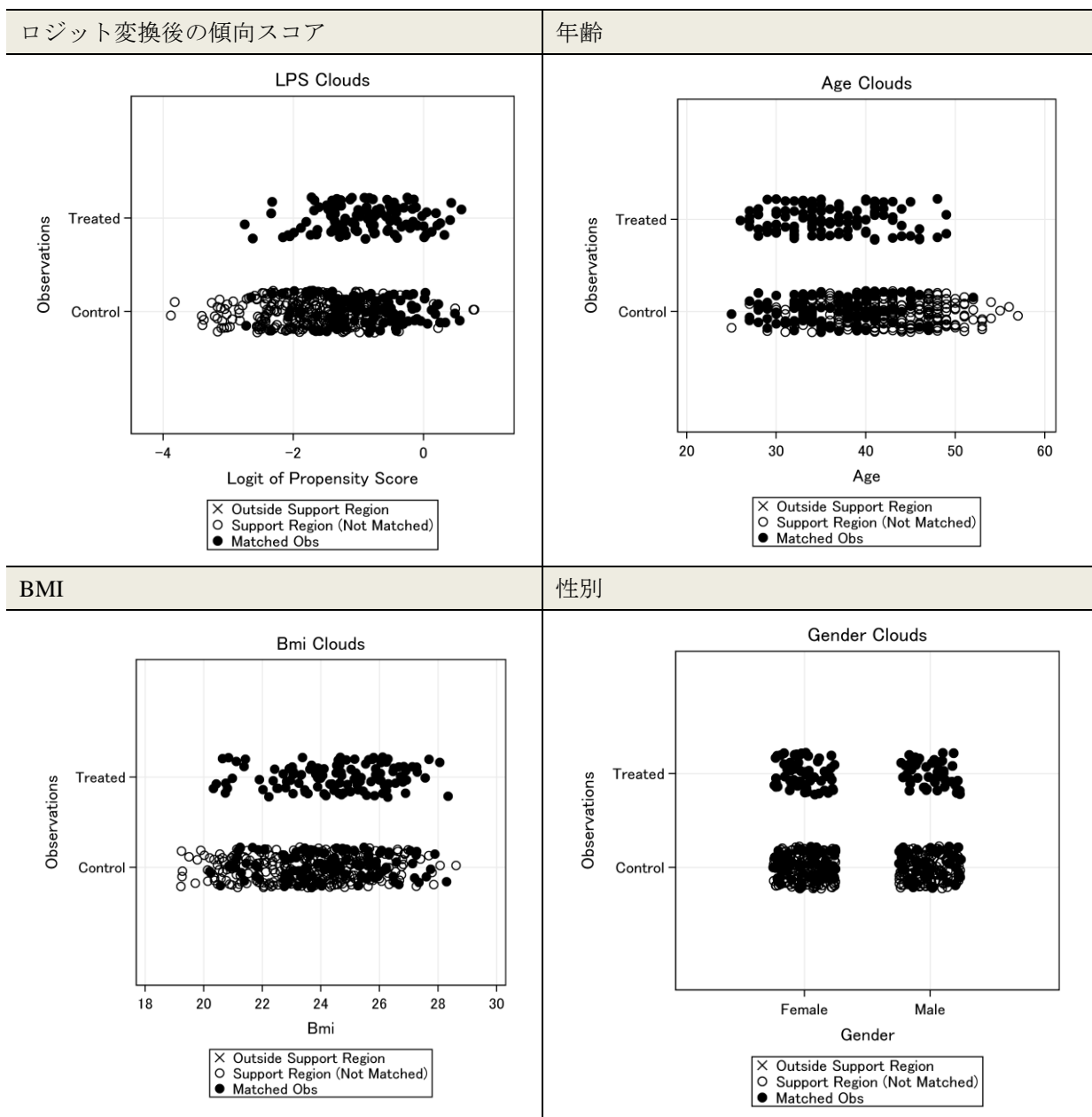


図 4 : ODS GRAPHICS による各変数に対する群別のクラウドプロット

クラウドプロットでは、縦軸に群、横軸に対象となる変数の値をとって散布図を描いており、マッチされたデータとマッチされなかったデータの記号を変えて出力している。クラウドプロットをみれば、傾向スコアがどの程度の値のデータでペアが構成できなかったか、各変数がどの程度の値をとるデータでペアが構成できなかったか、視覚的に確認することができる。

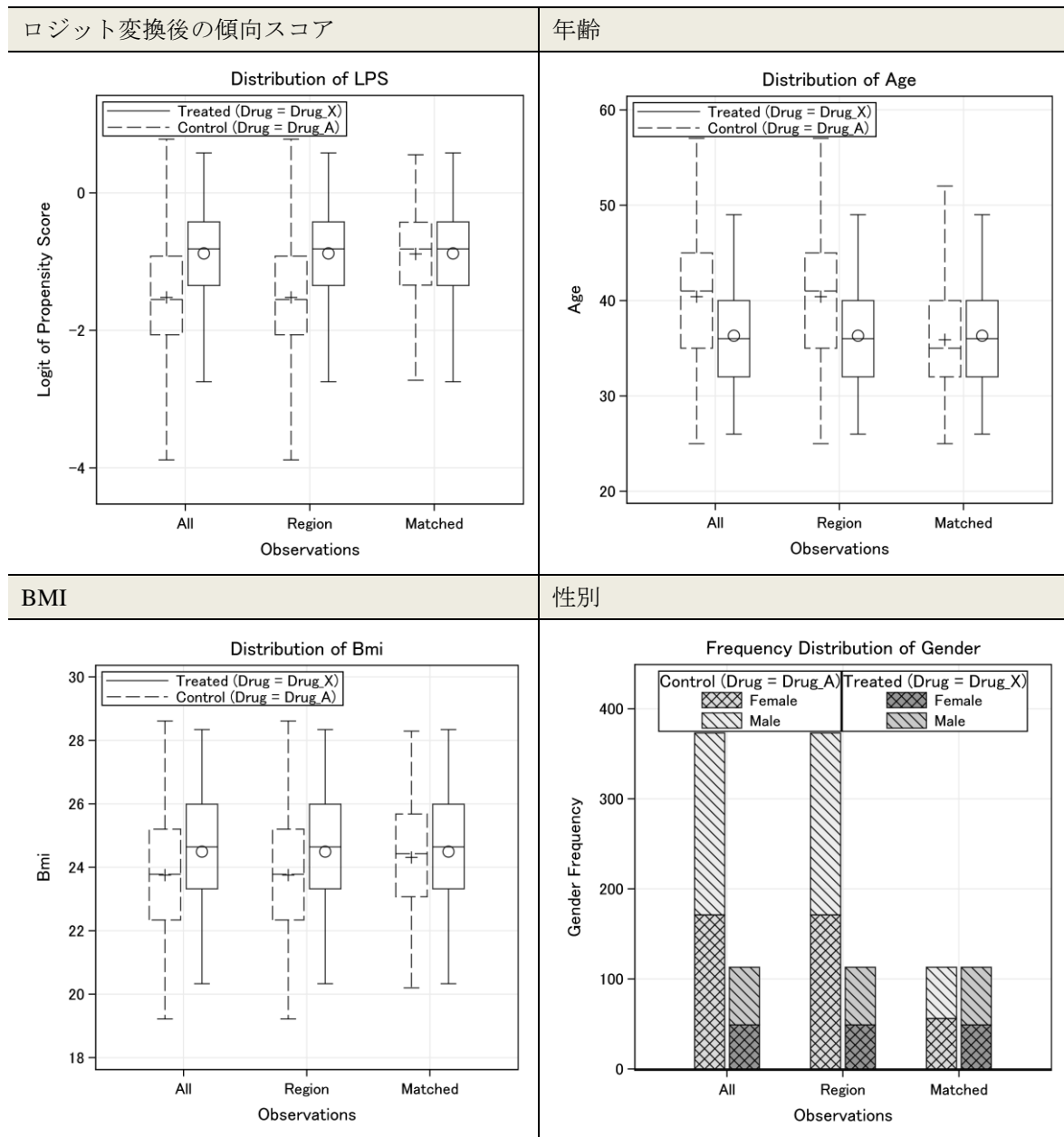


図 5 : ODS GRAPHICS による各変数に対する群別の箱ひげ図・度数分布

また、ODS GRAPHICS による出力として、連続変数に対しては箱ひげ図、カテゴリカル変数に対しては度数分布が出力される (図 5)。図 5 では、プログラム 4 において `PLOTS = ALL` と指定するだけでなく、`PLOTS(ORIENT = VERTICAL) = ALL` と指定し、縦軸に対象となる変数の値、横軸に群をとったグラフを描いている。なお、デフォルトは `ORIENT = HORIZONTAL` である。図 5 の箱ひげ図および度数分布より、各変数におけるマッチング前後で群間のバランスが視覚的に確認できる。なお、`MATCH` ステートメントにおいて、

EXACT=<カテゴリカル変数> と指定すると、指定したカテゴリカル変数は群間で完全にバランスが取れる。プログラム 4 の MATCH ステートメントにおいて、EXACT = Gender と指定すれば、”Some treated units have less than the specified K=1 matched controls because there are not enough available controls for these treated units.” という警告が出て、マッチされた対象データは結果的に 113 名から 112 名に減少してしまう。しかし、群間の性別の度数分布は完全にバランスが取れていることが視覚的に確認できるであろう (付録 D)。

なお、図 4 および図 5 で示した箱ひげ図やクラウドプロットの出力では、モノクロ印刷されても識別しやすいよう、ODS の出力としてジャーナルスタイル "STYLE = JOURNAL2" を指定した。特に、図 5 においては群別のグラフを示す場合、デフォルトでは群の違いが色で分けて出力されてしまうが、ジャーナルスタイルを指定したことで、モノクロ印刷に対応できるよう出力されている (魚住ら, 2016a, 2016b; 大橋ら, 2016)。

## 4 まとめ

本稿では、観察研究データに対する統計解析手法として、傾向スコアによるマッチングに着目した。傾向スコアマッチングを SAS で実行する方法として、SAS/STAT 14.2 から追加された PSMATCH プロシジャによるプログラムについて解説し、実行結果の解釈に関して詳述した。PSMATCH プロシジャにおける PSMODEL ステートメントで、傾向スコアモデルにおける目的変数と説明変数を指定でき、MATCH プロシジャで様々なマッチングがサポートされている。3 節では、特に最近傍マッチングをについて取りあげ、PSMATCH プロシジャによるアウトプット画面および ODS GRAPHICS による出力結果について解説した。なお、2 節において、観察研究における平均治療効果の推定について整理したもとの、傾向スコアによる最近傍マッチングに関しても解説している。さらに、SAS のプロシジャで傾向スコアマッチングが実施できなかった状況下で、JMP や R で実行できることを述べた。そして、JMP で傾向スコアマッチングを行う場合の留意点を取りあげ、加えて R による最近傍マッチングの実行プログラムを付録 B に示した。また、PSMATCH プロシジャで実施している傾向スコアマッチングのアルゴリズムを理解する手助けとなるよう、DATA ステップで最近傍マッチングを実施するための SAS マクロを開発した (付録 A)。この SAS マクロは、キャリパーの設定値として、臨床論文でよく報告されている、傾向スコアの推定値をロジット変換した値の標準偏差を掛けた値となっている点で有用である。

SAS ユーザー総会においても、これまで傾向スコアを用いた IPTW 法などに関して、SAS を用いた研究報告が行われている (古川・杉本, 2010; 古川, 2011)。また、SAS/STAT 14.2 では PSMATCH プロシジャに加えて、平均因果効果を推定するためのプロシジャとして、新たに CAUSALTRT プロシジャも追加されている (SAS Institute Inc., 2016)。PSMATCH プロシジャや CAUSALTRT プロシジャが新たに追加されたことによって、今後 SAS による傾向スコアを用いた統計解析が多く実施され、さまざまな分野におけるデータ解析に応用された結果が、論文・学会発表を通して報告されると予想される。傾向スコアによるマッチングを用いた調整解析を行う実務家にとって、本稿の報告内容がお役に立てれば幸いである。

## 参考文献

- [1] Austin PC, Grootendorst P, Anderson GM. A comparison of the ability of different propensity score models to balance measured variables between treated and untreated subjects: a Monte Carlo study. *Statistics in Medicine*. 26:734–753, 2007.

- [2] Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Statistics in Medicine*. **28**:3083–3107, 2009.
- [3] Austin PC. Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies. *Pharmaceutical Statistics*. **10**:150–161, 2011.
- [4] Bangalore S, Guo Y, Samadashvili Z, Blecker S, Xu J, Hannan EL. Everolimus-eluting stents or bypass surgery for multivessel coronary disease. *New England Journal of Medicine*. **372**:1213–1222, 2015.
- [5] Coca-Perraillon M. Local and global optimal propensity score matching. *Proceedings of the SAS Global Forum*. SAS Institute Inc., Cary, NC, 2007. Available at <http://www2.sas.com/proceedings/forum2007/185-2007.pdf>.
- [6] Faries D, Leon AC, Haro JM, Obenchain RL. *Analysis of Observational Health Care Data Using SAS(R)*. SAS Institute Inc., Cary, NC, 2010.
- [7] Hida K, Okamura R, Sakai Y, Konishi T, Akagi T, Yamaguchi T, Akiyoshi T, Fukuda M, Yamamoto S, Yamamoto M, Nishigori T, Kawada K, Hasegawa S, Morita S, Watanabe M. Open versus laparoscopic surgery for advanced low rectal cancer: a large multicenter propensity score matched cohort study in Japan. *Annals of Surgery*, 2017. DOI:10.1097/SLA.0000000000002329.
- [8] Hamatani Y, Ogawa H, Uozumi R, Iguchi M, Yamashita Y, Esato M, Chun YH, Tsuji H, Wada H, Hasegawa K, Abe M, Morita S, Akao M. Low body weight is associated with the incidence of stroke in atrial fibrillation patients: Insight from the Fushimi AF Registry. *Circulation Journal*. **79**:1009–1017, 2015.
- [9] JMP ジャパン事業部. JMP による傾向スコアを用いたマッチング、層別分析、回帰分析. SAS Institute Japan 株式会社, 2014. Available at <https://www.jmp.com/content/dam/jmp/documents/jp/support/propensityscore201412.pdf>.
- [10] Lanehart RE, de Gil PR, Kim ES, Bellara AP, Kromrey JD, Lee RS. Propensity score analysis and assessment of propensity score approaches using SAS(R) procedures. *Proceedings of the SAS Global Forum*. SAS Institute Inc., Cary, NC, 2012. Available at <http://support.sas.com/resources/papers/proceedings12/314-2012.pdf>.
- [11] Lunceford JK, Davidian M. Stratification and weighting via the propensity score in estimation of causal treatment effects: a comparative study. *Statistics in Medicine*. **23**:2937–2960, 2004.
- [12] Normand ST, Landrum MB, Guadagnoli E, Ayanian JZ, Ryan TJ, Cleary PD, McNeil BJ. Validating recommendations for coronary angiography following acute myocardial infarction in the elderly: a matched analysis using propensity scores. *Journal of Clinical Epidemiology*. **54**:387–398, 2001.
- [13] Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika*. **70**:41–55, 1983.
- [14] Rosenbaum PR, Rubin DB. Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician*. **39**:33–38, 1985.
- [15] Rosenbaum PR. *Design of Observational Studies*. Springer-Verlag: New York, NY, 2010.
- [16] SAS Institute Inc. *SAS/STAT(R) 14.2 User's Guide*. SAS Institute Inc., Cary, NC, 2016.
- [17] Stuart EA. Matching methods for causal inference: A review and a look forward. *Statistical Science*. **25**:1–21, 2010.
- [18] Yamashita Y, Uozumi R, Hamatani Y, Esato M, Chun YH, Tsuji H, Wada H, Hasegawa K, Ogawa H, Abe M,

Morita S, Akao M. Current status and outcomes of direct oral anticoagulant use in patients with atrial fibrillation in the real-world: The Fushimi AF Registry. *Circulation Journal*, 2017 DOI:10.1253/circj.CJ-16-1337.

- [19] 魚住龍史・吉田早織・平井隆幸, 浜田知久馬. Kaplan-Meier プロット・Forest プロット作成の応用: グラフ出力範囲内・範囲外への数値出力. SAS ユーザー総会 論文集. 216-232, 2016a.
- [20] 魚住龍史・矢田真城・浜田知久馬. SAS プロシジャを用いた生存時間データに対する例数設計の変革. SAS ユーザー総会 論文集. 250-267, 2016b.
- [21] 大橋靖雄・浜田知久馬・魚住龍史. 生存時間解析 応用編 -SAS による生物統計. 東京大学出版会, 2016.
- [22] 田栗正隆. 有名論文から統計の基礎を学ぶ: プロペンシティスコア. *Coronary Intervention*, **13**:63-69, 2017.
- [23] 古川敏仁・杉本典子. Propensity Score 法によるバイアスの調整法に関する実務的な問題点. SAS ユーザー総会 論文集. 101-108, 2010.
- [24] 古川敏仁. 観察研究において選択 bias を制御するために用いられる Propensity Score IPTW と層化調整法の、頑健性の観点からの使い分けについて. SAS ユーザー総会 論文集. 409-416, 2011.
- [25] 山本倫生・森田智視. 傾向スコアによる調整解析. *呼吸*. **34**:1187-1193, 2015.

## 連絡先

E-mail : [uozumi@kuhp.kyoto-u.ac.jp](mailto:uozumi@kuhp.kyoto-u.ac.jp)

### 付録 A: DATA ステップによる最近傍マッチングを行うための SAS マクロ

```
*=====*;
* 群を表す数値変数を準備 ;
data Drugs0;set Drugs;
  if Drug="Drug_X" then Group=1;else if Drug="Drug_A" then Group=0;
run;

*=====*;
* 傾向スコアを含んだデータセットを準備 ;
ods listing close;
proc logistic data=Drugs0;
  class Gender; model(event="1")=Gender Age Bmi;
  output out=ps00 pred=ps xbeta=ps_xb stdxbeta=ps_sdx;
run;
ods listing;

*=====*;
* MACRO: PSMATCHING ;
*=====*;
* 傾向スコアによるマッチング後の被験者番号リストを作成する ;
* 傾向スコアによるマッチング後のデータセットを作成する ;
*=====*;
* 引数の説明 ;
*=====*;
* data: 入力対象データセット [傾向スコアなどを変数に含めたデータセットとする] ;
* id: 被験者番号を表すカテゴリ変数 ;
* grp: 群 (治療群と対照群) を表すカテゴリ変数 ;
* tgrp_ind: 治療群を表す値 ;
* cgrp_ind: 対照群を表す値 ;
* ps_pred: 傾向スコアを表す変数 ;
* sort: マッチングを行う傾向スコアの順を指定 ;
* --- sort = descending : 傾向スコアの大きい順 ;
* --- sort = ascending : 傾向スコアの小さい順 ;
```



```

* --- sort = random      : ランダムな順 ;
* caliper: キャリパーの値(傾向スコアのロジット変換した値の標準偏差に掛ける値)を指定 ;
* seed: 乱数のシードを指定(sort = random) ;
* outid: マッチングされた被験者番号を群ごとに出力したデータセット ;
* outdata: 入力対象データセットからマッチングされなかった被験者番号を削除したデータセット ;
*=====*;

*-----;
* マクロ開始 ;
%macro psmatching(data,id,grp,tgrp_ind,cgrp_ind,ps_pred,sort,caliper,seed,outid,outdata);

*-----;
* 傾向スコアの変数名・傾向スコアのロジット変換した変数を準備 ;
data ps_p;set &data; ps_pred=&ps_pred; l_ps_pred=log(ps_pred/(1-ps_pred)); run;

*-----;
* 各群のデータ数を読み込み ;
ods listing close; ods output summary=grp_N;
proc means data=ps_p n; class &grp; var ps_pred; run;
ods listing;
data _null_;set grp_N;
  where &grp=&tgrp_ind; %global tgrp_N; call symput("tgrp_N",ps_pred_N);
data _null_;set grp_N;
  where &grp=&cgrp_ind; %global cgrp_N; call symput("cgrp_N",ps_pred_N);
run;

*-----;
* 傾向スコアのロジット変換した値の標準偏差を準備 ;
ods listing close; ods output summary=lps_p;
proc means data=ps_p stddev; var l_ps_pred; run;
ods listing;

*-----;
* マッチングをランダムに行う場合の乱数を準備 ;
data one;set ps_p; seed=&seed; call streaminit(seed); ranvar=rand("uniform"); run;

*-----;
* マッチングを行う傾向スコアの順を指定 ;
%if "&sort"="descending" %then %do;
proc sort data=one out=one; by &grp descending ps_pred;
proc transpose data=one out = data1; by &grp; run;
%end;

%else %if "&sort"="ascending" %then %do;
proc sort data=one out=one; by &grp ps_pred;
proc transpose data=one out = data1; by &grp; run;
%end;

%else %if "&sort"="random" %then %do;
proc sort data=one out=one; by &grp ranvar;
proc transpose data=one out = data1; by &grp; run;
%end;

*-----;
* マッチングを行うために各群のデータ数だけ変数を準備 ;
data id_t (rename=(COL1-COL%eval(&tgrp_N)=tid1-tid%eval(&tgrp_N)));set data1;
  if &grp=&tgrp_ind and _NAME_="PatientID";
data ps_t (rename=(COL1-COL%eval(&tgrp_N)=tps1-tps%eval(&tgrp_N)));set data1;
  if &grp=&tgrp_ind and _NAME_="ps_pred";
data id_c (rename=(COL1-COL%eval(&cgrp_N)=cid1-cid%eval(&cgrp_N)));set data1;
  if &grp=&cgrp_ind and _NAME_="PatientID";
data ps_c (rename=(COL1-COL%eval(&cgrp_N)=cps1-cps%eval(&cgrp_N)));set data1;
  if &grp=&cgrp_ind and _NAME_="ps_pred";
run;

*-----;
* キャリパーを指定して傾向スコアマッチングを実施 ;
data all;

```

```

merge id_t ps_t id_c ps_c lps_p;
caliper=&caliper;
array t_id {*} tid1-tid%eval(&tgrp_N); array c_id {*} cid1-cid%eval(&cgrp_N);
array t_ps {*} tps1-tps%eval(&tgrp_N); array c_ps {*} cps1-cps%eval(&cgrp_N);
array used_i {*} used1 - used%eval(&cgrp_N);
array matched_t {*} m_tid1-m_tid%eval(&tgrp_N); array matched_c {*} m_cid1-m_cid%eval(&tgrp_N);
array min_diff_i {*} min_diff1-min_diff%eval(&cgrp_N);
match_N=0;
do i=1 to &tgrp_N;
  min_diff=1; best_match=0;
  do j=1 to &cgrp_N;
    if used_i[j]=. then do;
      if abs(log(t_ps[i]/(1-t_ps[i]))-log(c_ps[j]/(1-c_ps[j])))
        < caliper*1_ps_pred_StdDev then do;
        if abs(log(t_ps[i]/(1-t_ps[i]))-log(c_ps[j]/(1-c_ps[j]))) < min_diff then do;
          min_diff=abs(log(t_ps[i]/(1-t_ps[i]))-log(c_ps[j]/(1-c_ps[j])));
          min_diff_i[j]=abs(log(t_ps[i]/(1-t_ps[i]))-log(c_ps[j]/(1-c_ps[j])));
          best_match=j;
        end;
      end;
    end;
  end;
  if best_match > 0 then do;
    match_N=match_N+1; used_i[best_match]=1;
    matched_t[match_N]=t_id[i]; matched_c[match_N]=c_id[best_match];
  end;
end;
run;
data all;set all;
  max_min_diff=max(of min_diff1-min_diff%eval(&cgrp_N));
  sum_min_diff=sum(of min_diff1-min_diff%eval(&cgrp_N));
run;

* -----;
* マッチングされた被験者番号を群ごとに出力したデータセット ;
data &outid;set all;
  array matched_t {*} m_tid1-m_tid%eval(&tgrp_N); array matched_c {*} m_cid1-m_cid%eval(&tgrp_N);
  do match = 1 to match_N; treat_id=matched_t[match]; control_id=matched_c[match]; output; end;
  keep match treat_id control_id;
run;

* -----;
* 群を表す変数を準備 ;
proc transpose data=&outid out=psmatch00; by match; var treat_id control_id;
data psmatch00;set psmatch00;
  if _NAME_="treat_id" then &grp=&tgrp_ind; else if _NAME_="control_id" then &grp=&cgrp_ind;
  &id=COL1; keep match &grp &id;
proc sort data=psmatch00 out=psmatch00; by &grp &id; run;
proc sort data=ps00 out=ps00sort; by &grp &id; run;

* -----;
* 入力対象データセットからマッチングされなかった被験者番号を削除したデータセット ;
data &outdata; merge ps00sort psmatch00(in=in1); by &grp &id; if in1=1 then output;
proc sort data=&outdata out=&outdata; by match &grp &id; run;

* -----;
* マクロ終了 ;
%mend psmatching;

* -----*;
* ランダムな順に傾向スコアマッチング ;
* -----*;
%psmatching(ps00, PatientID, Group, 1, 0, ps, random, 0.20, 123456, matchesr, psmatchdatar);

* -----*;
* 傾向スコアの大きい順にマッチング ;
* -----*;
%psmatching(ps00, PatientID, Group, 1, 0, ps, descending, 0.20, 123456, matchesd, psmatchdatad);

```

## 付録 B : R による最近傍マッチングの実行プログラム

```

### パッケージを使用 ###
library(Matching)

### 傾向スコアの推定結果が含まれたCSVファイルPSOUT00をデータセットとして読み込み ###
data = read.csv("psout00.csv")

### 傾向スコアモデルの分子(治療群のカテゴリ)を指定 ###
Trt = data$Drug == "Drug_X"

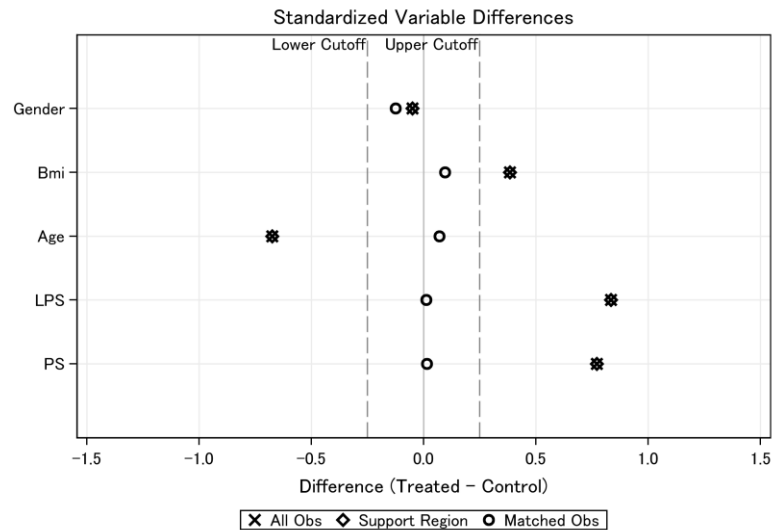
### キャリパーの値を指定 ###
caliper0 = 0.20

### 傾向スコアのロジット変換 ###
data <- transform(data, ps.logit=log(ps / (1 - ps)))

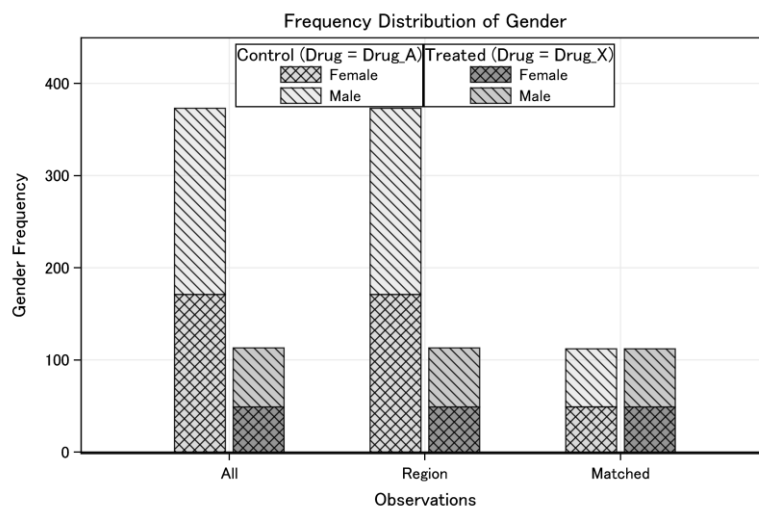
### 乱数のシード ###
set.seed(123456)

### 1:1最近傍マッチングを非復元抽出で実行 ###
Matching = Match(Y=NULL, Tr=Trt, X=data$ps.logit, M=1, replace=FALSE, caliper = caliper0);
summary(Matching)

```



付録 C : 各変数に対する標準化した群間差 (ODS GRAPHICS)



付録 D : MATCH ステートメントで EXACT = Gender とした場合の性別の度数分布