Paper PR05

A Method/Macro Based on Propensity Score and Mahalanobis Distance to Reduce Bias in Treatment Comparison in Observational Study

Wuwei Wayne Feng MS, Eli Lilly & Company, Indianapolis, IN
Yu Jun MS, MedFocus Ltd. Des Plaines, IL
Rong Xu MS, Eli Lilly & Company, Indianapolis, IN

ABSTRACT

In observational studies, investigators usually do not have the same control over the treatment assignment as they do with randomized controlled studies. As a result, the treatment and control groups may have a large difference on their observed covariates. These differences could lead to bias in estimating treatment effects. There are several propensity score based methods that could reduce the bias caused by these differences and make the two groups comparable. One method is the nearest available Mahalanobis metric matching within the calipers defined by the propensity score. This paper will present and demonstrate the matching algorithm based on this method. A macro is developed to implement the matching algorithm; a growth hormone observational study is used as an example to demonstrate the bias before the match and percentage of bias reduction after the match.

Key words: observational study, propensity score, Mahalanobis distance

INTRODUCTION

In observational studies, unlike controlled randomized clinical studies, frequently there are large differences in participants' characteristics between treatment and control (or another treatment) groups. These differences may lead to bias in the direct comparison of treatment effect, especially when there is a strong relationship between these characteristics and the outcome variable. Traditionally covariates adjustment and various matching algorithms were used to reduce the bias [1,8]; however, in many situations these methods are not adequate to address the issue. For example, using a model with a large number of covariates may lead to inefficient estimates of the treatment effect. Propensity score, defined as the conditional probability of receiving a particular treatment $(Z_i = 1)$ versus control or anther treatment group $(Z_i = 0)$ given the study participants' covariates, X_i, $\Pr(Z_i = 1 \mid X_i = x_i)$ was first introduced by Rosenbaum and Rubin[6] in 1983, and now is commonly used as a building block in many methods of bias reduction in the analysis of observational data. Matched pairing, stratification (sub-classification) and covariance adjustment are the three commonly used propensity score based techniques. Propensity score can be estimated from a logit model using PROC LOGISTIC[11] procedure in SAS/STAT.

Mahalanobis distance is the distance between two *N* dimensional points scaled by the statistical variation in each component of the point. For example, if X and Y are two points from the same distribution with covariance matrix C, then the Mahalanobis distance can be expressed as $D(X,Y) = (X - Y)^t C^{-1} (X - Y)$. When the covariance matrix is the identity matrix, Mahalanobis distance specializes to the Euclidean distance. Mahalanobis Metric Matching was used as one method of matching observations based on Mahalonobis distance for bias reduction in observational studies [9].

Matching is a method of sampling from a large reservoir of potential candidates in one group to produce a group of modest size in which the distribution of covariates is similar to that in another group. Various methods were developed by combining ideas of matching and propensity score[5,7,8,10] Specifically, propensity scores can be used to construct matching cohorts using three methods: (i) Nearest available matching on the estimated propensity score, (ii) Mahalanobis metric matching including the propensity score and (iii) the nearest available Mahalanobis metric matching within calipers defined by the propensity score. All three methods are useful techniques with different properties. The first method is simple and incurs less computation. Lori [3,4] developed and presented two macros based on the first method. The second method has the effect of "equal percent bias reducing" (EPBR) [9]. The third method produces the best balance for the covariates between two treatment groups, and is considered to be superior to the other two methods[5,8] .

In this paper we present a SAS macro that was developed to implement the third method listed above - the nearest available Mahalanobis metric matching within calipers defined by the propensity score - and use the macro to illustrate the methodology .

Data from a growth hormone study - a large observational study, is used throughout the paper. The objective of the study is to evaluate long-term safety outcomes in growth hormone deficiency (GHD) adult patients received with treatment A compared with treatment B. A total of 2430 (1988 in treatment group A, 442 in treatment group B) GHD adult patients were enrolled in the study. It has been determined based on expert's opinion that 37 variables at baseline may be related to making decision on treatment assignment; therefore these variables were included in logit model to estimate propensity score. Five of these variables, including age, status of diabetic insipidus, onset of disease (adult=1, 0ther=0) and cause of GH deficiency (1=tumor/adenoma, 0=no) and logit of propensity score were considered as key variables for treatment choice and were used in computing the Mahalanobis distance. One quarter of standard deviation of logit of propensity score (about 0.23) was used as a caliper.

MATCHING ALGORITHM

The matching algorithm and macro proposed in the paper are illustrated in **Chart 1**.

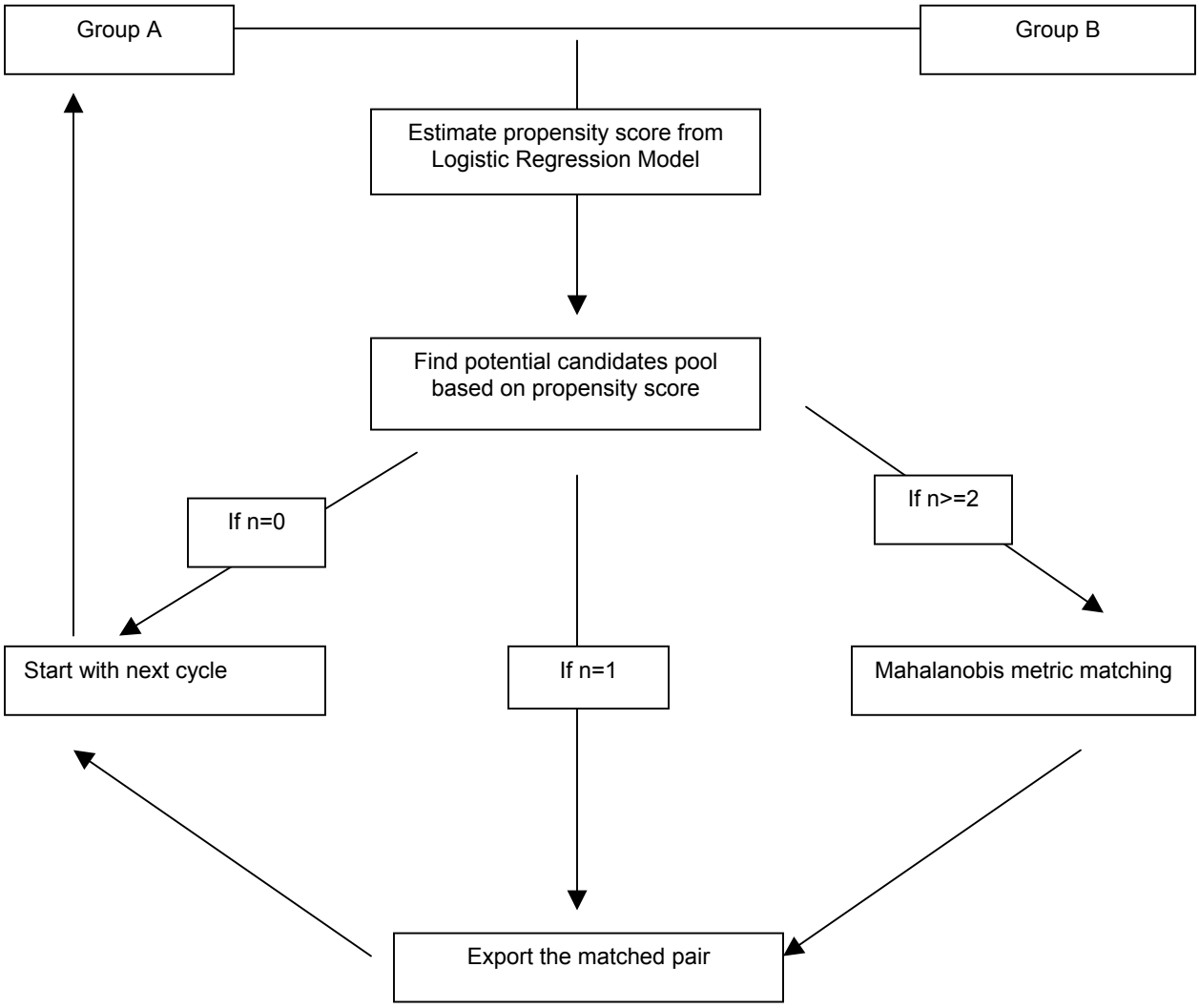
Step 1. Propensity scores are computed for every subject using a logistic regression model with all possible independent variables that may have affected choosing treatment for study participants included.

Step 2. The subjects in treatment group **A** (the group with less study participants) are randomly ordered and the first subject from group **A** is selected. Subjects in group **B** whose propensity score is within the caliper (one quarter of standard deviation of the logit of the propensity score as suggested by Rosenbaum and Rubin [2,]) are identified as initial matched candidates. Three situations may occur in this step: (I) No candidate is located within the caliper, then this round of matching will stop here and the next round will start (i.e. determine the match for the next subject in group **A**). (II) Only one candidate is identified within the caliper and this candidate is considered as final. (III) More than two candidates were identified, and then the search process will move to Step 3.

Step 3. Mahalanobis distances based on a smaller numbers of key variables and propensity score are calculated between the subject in group **A** and those initially selected subjects in group **B**. The subject with the smallest distance to the subject in group **A** is selected as a final matched candidate. The matched pair is then removed from the pool, and the process will repeat for the next subject in group A. All remaining subjects in group **B** are available for the remaining matching rounds. The matching process will repeat until the group **A** is exhausted.

Notice that direct computation of Mahalanobis distances involves inverting variance-covariance matrix **C**, which is both numerically unstable and computationally expensive. This can be avoided by transforming the raw data **X** into standardized **X*** having an identity covariance (via a spectral decomposition of **C**). Then the Mahalanobis distance for data-points in **X** will be identical to the Euclidean distance for data-points in the standardized space, X*. The computations can be implemented in SAS by a variety of ways. For example, one can use PROC IML subroutine SVD. Since some SAS users may have no access to SAS/IML software, in the macro we chose to only use SAS procedures available in the SAS/STAT software. Specifically, we use PROC PRINCOMP and PROC SCORE to obtain the principal components scores with an identity covariance matrix for all data-points and PROC FASTCLUS to compute the Euclidean distances from each observation to a specific reference point by setting it as a seed.

Chart 1 Matching Scheme



RESULTS

As referenced in the Introduction, a growth hormone observational study is used here as an application example for match algorithm and the proposed macro. **Table I** describes the original population and contains all of covariates that were used in the multiple logistic regression model to compute propensity score. Difference between two treatment groups is evaluated using T-test for continuous variables and Chi-Square test for categorical variables. P values for majority of variables (23 out of 38) are significant at 0.05 level which means there is a significant difference between the two groups with respect to these characteristics. A direct comparison between two groups based on this study population would lead to biased and invalid inference if the outcome measure is related to any of these baseline characteristics.

Table I: Group Comparisons Prior to Matching

Variables	Group A (Mean + SD)	Group B (Mean + SD)	P-value
Total patients	1988	442	
Age	46.08 (14.83)	54.82 (15.98)	<.001
Baseline body mass index	31.27 (7.02)	30.08 (6.22)	<.001
Baseline diastolic blood pressure	77.98 (10.31)	77.72 (10.50)	0.635
Baseline systolic blood pressure	123.07 (16.99)	126.90 (19.28)	<.001
Number of pituitary hormonal disorders	2.15 (1.28)	2.17 (1.14)	0.732
Number of smoking years	7.31 (11.96)	9.21 (13.95)	0.008
Logit of propensity score	-1.86 (0.88)	-1.16 (0.82)	<.001
	N** (%)	N** (%)	
Pre-existing coronary artery disease (0 = O/W,1 = Y)	127 (6.39)	50 (11.31)	<.001
Pre-existing diabetes insipidus (0 = O/W,1 = Y)	435 (21.88)	58 (13.12)	<.001
Pre-existing diabetes mellitus (0 = O/W,1 = Y)	166 (8.35)	64 (14.48)	<.001
Pre-existing hypertension (0 = O/W,1 = Y)	493 (24.80)	143 (32.35)	0.002
Pre-existing hyperlipidemia (0 = O/W,1 = Y)	808 (40.64)	214 (48.42)	0.003
Pre-existing pituitary microadenoma (0 = O/W,1 = Y)	237 (11.92)	36 (8.14)	0.024
Pre-existing pituitary macroadenoma (0 = O/W,1 = Y)	713 (35.87)	231 (52.26)	<.001
Onset type for GH deficienc (adult/childhood) (0 = O/W,1 = Adult)	1674 (84.21)	391 (88.46)	0.023
Pre-existing pathological bone fracture (0 = O/W,1 = Y)	57 (2.87)	1 (0.23)	<.001
Pre-existing visual impairment (0 = O/W,1 = Y)	527 (26.51)	147 (33.26)	0.005
Cause of GH deficiency (idiopathic) (0 = O/W,1 = Idiopathic)	365 (18.36)	48 (10.86)	<.001
Cause of GH deficiency (empty sella) (0 = O/W,1 = Empty Sella)	96 (4.83)	12 (2.71)	0.055
Cause of GH deficiency (trauma-shee) (0 = O/W,1 = Trauma-Shee)	89 (4.48)	14 (3.17)	0.242
Cause of GH deficiency (tumor/adenoma) (0 = O/W,1 = Tumor/Adeno)	1243 (62.53)	335 (75.79)	<.001
Cause of GH deficiency (Other) (0 = O/W,1 = Other)	195 (9.81)	33 (7.47)	0.149
Family history of cerebrovascular disease (0 = O/W,1 = Y)	588 (29.58)	141 (31.90)	0.359
Gender (0 = F,1 = M)	1105 (55.58)	273 (61.76)	0.020
Pre-existing malignant tumor (0 = O/W,1 = Y)	91 (4.58)	36 (8.14)	0.004
Pre-existing non-functional pituitary adenoma (0 = O/W,1 = Y)	546 (27.46)	175 (39.59)	<.001
Pre-existing functional pituitary adenoma	411 (20.67)	93 (21.04)	0.846

(0 = O/W, 1 = Y)			
Pre-existing history of radiotherapy (0 = O/W, 1 = Y)	580 (29.18)	141 (31.90)	0.274
Ethnicity (Caucasian Descent) (0 = O/W,1 = Y)	1728 (86.92)	356 (80.54)	<.001
Ethnicity (East/Southeast Asian Descent) (0 = O/W,1 = Y)	23 (1.16)	19 (4.30)	<.001
Ethnicity (African Descent) (0 = O/W,1 = Y)	127 (6.39)	36 (8.14)	0.206
Ethnicity (Hispanic Descent) (0 = O/W,1 = Y)	85 (4.28)	25 (5.66)	0.207
Family history of diabetes mellitus (0 = O/W, 1 = Y)	858 (43.16)	178 (40.27)	0.288
Family history of cardiovascular disease (0 = O/W, 1 = Y)	1121 (56.39)	236 (53.39)	0.266
Family history of osteoporosis (0 = O/W, 1 = Y)	343 (17.25)	60 (13.57)	0.066
Family history of cancer/other neoplasms (0 = O/W, 1 = Y)	1138 (57.24)	226 (51.13)	0.020
Pre-existing Craniopharyngioma (0 = O/W,1 = Y)	210 (10.56)	36 (8.14)	0.139
Pre-existing Acromegaly (0 = O/W, 1 = Y)	55 (2.77)	13 (2.94)	0.873

** N is number of patients with value equals 1

Table II describes the matched population based on the matching algorithm implemented in our macro. In summary, 441 patients out of 442 (99.8%) from group B were successfully matched by patients from group A. Only 882 patients (36.3%) out of 2430 total patients will be used for the final analysis to estimate treatment effect. Difference between the matched pairs is evaluated using T-test for continuous variables and Chi-Square test for categorical ones. For none variables, there is a significant difference between two matched groups, which shows that the two matched groups are well balanced.

Table II: Group Comparisons after Matching

Variables	Group A (Mean + SD)	Group B (Mean + SD)	P-value
Total patients	441	441	
Age	54.77 (15.55)	54.77 (15.95)	0.996
Baseline body mass index	30.13 (6.14)	30.09 (6.23)	0.911
Baseline diastolic blood pressure	77.64 (10.00)	77.77 (10.46)	0.853
Baseline systolic blood pressure	126.50 (18.64)	126.89 (19.30)	0.761
Number of pituitary hormonal disorders	2.16 (1.15)	2.17 (1.14)	0.907
Number of smoking years	9.23 (14.11)	9.23 (13.95)	1.000
Logit of propensity score	-0.016 (0.18)	0.016 (0.18)	0.009
	N** (%)	N** (%)	
Pre-existing coronary artery disease (0 = O/W,1 = Y)	51 (11.56)	50 (11.34)	1.000
Pre-existing diabetes insipidus (0 = O/W,1 = Y)	57 (12.93)	58 (13.15)	1.000
Pre-existing diabetes mellitus (0 = O/W,1 = Y)	65 (14.74)	64 (14.51)	1.000
Pre-existing hypertension (0 = O/W,1 = Y)	147 (33.33)	142 (32.20)	0.774
Pre-existing hyperlipidemia (0 = O/W,1 = Y)	212 (48.07)	213 (48.30)	1.000
Pre-existing pituitary microadenoma (0 = O/W,1 = Y)	35 (7.94)	36 (8.16)	1.000
Pre-existing pituitary macroadenoma (0 = O/W,1 = Y)	228 (51.70)	230 (52.15)	0.946
Onset type for GH deficiency (adult/childhood) (0 = O/W,1 = Adult)	388 (87.98)	390 (88.44)	0.917
Pre-existing pathological bone fracture	1 (0.23)	1 (0.23)	1.000

(0 = O/W,1 = Y)			
Pre-existing visual impairment (0 = O/W,1 = Y)	150 (34.01)	147 (33.33)	0.887
Cause of GH deficiency (idiopathic) (0 = O/W,1 = Idiopathic)	50 (11.34)	48 (10.88)	0.915
Cause of GH deficiency (empty sella) (0 = O/W,1 = Empty Sella)	10 (2.27)	12 (2.72)	0.830
Cause of GH deficiency (trauma-shee) (0 = O/W,1 = Trauma-Shee)	15 (3.40)	14 (3.17)	1.000
Cause of GH deficiency (tumor/adenoma) (0 = O/W,1 = Tumor/Adeno)	335 (75.96)	334 (75.74)	1.000
Cause of GH deficiency (Other) (0 = O/W,1 = Other)	31 (7.03)	33 (7.48)	0.897
Family history of cerebrovascular disease (0 = O/W,1 = Y)	156 (35.37)	141 (31.97)	0.319
Gender (0 = F,1 = M)	281 (61.72)	273 (61.90)	0.626
Pre-existing malignant tumor (0 = O/W,1 = Y)	40 (9.07)	36 (8.16)	0.719
Pre-existing non-functional pituitary adenoma (0 = O/W,1 = Y)	164 (37.19)	174 (39.46)	0.533
Pre-existing functional pituitary adenoma (0 = O/W, 1 = Y)	100 (22.68)	93 (21.09)	0.625
Pre-existing history of radiotherapy (0 = O/W, 1 = Y)	137 (31.07)	140 (31.75)	0.885
Ethnicity (Caucasian Descent) (0 = O/W,1 = Y)	360 (81.63)	355 (80.50)	0.731
Ethnicity (East/Southeast Asian Descent) (0 = O/W,1 = Y)	12 (2.72)	19 (4.31)	0.272
Ethnicity (African Descent) (0 = O/W,1 = Y)	36 (8.16)	36 (8.16)	1.000
Ethnicity (Hispanic Descent) (0 = O/W,1 = Y)	28 (6.35)	25 (5.67)	0.777
Family history of diabetes mellitus (0 = O/W, 1 = Y)	171 (38.78)	178 (40.36)	0.680
Family history of cardiovascular disease (0 = O/W, 1 = Y)	238 (53.97)	236 (53.51)	0.946
Family history of osteoporosis (0 = O/W, 1 = Y)	61 (13.83)	60 (13.61)	1.000
Family history of cancer/other neoplasm (0 = O/W, 1 = Y)	238 (53.97)	226 (51.25)	0.458
Pre-existing Craniopharyngioma s(0 = O/W,1 = Y)	34 (7.71)	36 (8.16)	0.901
Pre-existing Acromegaly (0 = O/W, 1 = Y)	14 (3.17)	13 (2.95)	1.000

** N is number of patients with value equals 1

Table III shows the percentage bias reduction for five key variables after matching. As we can see from the table, the percentage of bias reduction for covariates varies from 89.18% to 99.89%. Although the T-test for the Logit of propensity score is statistically significant (p-value = 0.009), the mean difference between two groups is only 0.03 and the percentage of bias reduction is 95.71%, probably the significance of the test is due to a large sample size which could result in detecting a very small difference even in a population that is well balanced.

Table III Percent Reductions in Bias for Key Variables

Variables	Bias before matching	Bias after Matching	Percent Reduction
Age	8.74	0.01	99.89
Pre-existing diabetes insipidus	8.76	0.22	97.49
Onset type for GH deficiency (adult/childhood)	4.25	0.46	89.18
Cause of GH deficiency (tumor/adenoma)	13.26	0.45	96.61
Logit of propensity score	0.70	0.03	95.71

**Percentage bias reduction is calculated by $(1- D_i)/D_j$ *100% where D_i and D_j are group difference in covariates means after matching and before matching, respectively

CONCLUSION

By applying the macro presented in this paper, subjects in one treatment group can be successfully matched to another treatment group using Mahalanobis metric matching within the calipers defined by propensity score method. Such matching process would in general result in two treatment groups having very similar baseline characteristics, thus treatment difference can be appropriately estimated. The source SAS code is attached in the end of this paper and SAS users could apply this macro in their own research work.

REFERENCE

1. Carpter, R. G. Matching when covariables are normally distributed. Biometrika, 64, 299 – 307, 1977
2. Cochran, W. G. and Rubin, D. B. “Controlling bias in observational studies: A review”. Sankhya, Ser. A. 35 (1973), 417-446
3. Lori S Parsons, “Using SAS software to perform a case-control match on propensity score in an observational study”, SUGI 1999, 225-229,
4. Lori S. Parsons, “Reducing Bias in a propensity score matched-pair sample using greedy matching techniques”, Proceeding of the twenty-fifth annual SAS users group internal conference, Cary, NC: SAS Intistitue Inc. 2000, 1166-1171.
5. Ralph B Dagostino. Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. Statistics in Medicine, 1998, 17, 2265—2281.
6. Robenbaum, P. R. and Rubin, D. B. “The central role of the propensity score in the observational studies for causal effects’, Biometrika, 70, 41—55 (1983).
7. Robenbaum, P. R. and Rubin, D. B. “Constructing a control group using multivariate matched sampling methods that incorporate the propensity score”, The American Statistician, 39, 33-38 (1985) .
8. Rubin, D. B. Matching to remove bias in observational studies: a review. Sankhya, series A 35, 417-446. 1973a
9. Rubin, D.B. “Bias reduction using Mahalanobis metric matching” Biometrics, 36, 293-298 (1980).
10. Rubin. D.B. “Estimating causal effects from large data sets using propensity score”. Annals of Internal medicine, October, 1997, 127: 757-763.
11. SAS Instiute Inc. (2002), SAS/STAT User’s guide, Version 8, sixth edition, Volume 1, Cary, NC: SAS Inistitue Inc.

ILLUSTRATION OF CODE

```
options macrogen mlogic mprint;

libname XXXX 'X:\XXXX'; **Please specify the libname name and location***;

proc datasets lib=work kill;run;

*******************************************************************************
*** This macro is to calculate Mahalanobis distance from each point to a reference ***
*** point. For example, reference point can be a patient from case group, other    ***
*** points can be the patients from control group who meet the criteria within     ***
*** 1/4 standard deviation of Logit of propensity score                            ***
*******************************************************************************;


%macro Mahalanobis(data, var, refdata);

proc princomp data=&data std out=out outstat=outstat noprint;
     var &var;
 run;

proc score data=&refdata score=outstat out=reference_point;
```

```
     var &var;
run;

proc append data=out base=reference_point;
run;

proc fastclus data=reference_point maxc=1 replace=none maxiter=0 noprint
     out=mahalanobis_to_point(drop=cluster);
     var prin:;
run;

proc sql;
  create table mahalanobis_to_point
  as select a.*
  from mahalanobis_to_point a, &refdata b
  where a.&id^=b.&id;
quit;

proc sort data=mahalanobis_to_point;
  by distance;

data mahalanobis_to_point;
  set mahalanobis_to_point;
  by distance;
  if _n_=1;
  run;

data case_ctrl;
  set &refdata(keep=&id) mahalanobis_to_point(keep=&id);
  newid=&i;
  run;

data case_ctrl_together;
  set case_ctrl_together case_ctrl;
  if &id=. then delete;
  run;

*** Exclude patients in case group which are selected in the Mahalanobis macro ***;

proc sql;
  create table case_temp
  as select a.*
  from case_temp a,  mahalanobis_to_point b
  where a.&id^=b.&id;
  quit;

%mend Mahalanobis;

***Set up a dataset which will contain all control patients and matched case
patients***;


data case_ctrl_together; run;


*********************************************************************************
***    data  --- input dataset name                                            ***
***    class --- categorical variables included in the logistic model          ***
***    yvar  --- dependent variable (1=case group, 0=control group)            ***
***    xvar  --- independent variables                                         ***
***    id    --- patient identification number                                 ***
*********************************************************************************;
```

```
%macro match(data, class, yvar, xvar, id);

**Perform logistic regression, select independent variables, and create the propensity
score***;

proc logistic descending data=&data noprint;
   &class;
   model &yvar=&xvar;
   output out=propen(drop=_level_) prob=prob xbeta=logit;
   run;

proc univariate data=propen noprint;
  var logit;
  output out=propen_sd std=sd;
  run;

***create datasets of case and control, which contain patients from case and control
groups ***;

data ctrl case;
       set propen;

       if &yvar=0 and prob ne . then do;
       rannum=ranuni(1234567);
              label rannum='Uniform Randomization Score';
              output ctrl;
       end;

              else if &yvar=1 and prob ne . then do;
       output case;
       end;

data ctrl;
       if _n_=1 then set propen_sd(keep=sd);
              set ctrl;
run;

***sort ctrl by random number generated ***;

proc sort data=ctrl;
  by rannum;
  run;

data _null_;
  set ctrl;
  by sd;
  if last.sd then call symput('tot',put(_n_,8.));
  run;

data case_temp;
  set case;


*** select one patient once time from dataset ctrl, and search for matched-pair in
case group ***;

%let j=%eval(&tot);

%do i=1 %to &j;
  data ctrl_temp;
    set ctrl;
    if _n_=&i;
```

```
      low=logit - 0.25*sd;
      up =logit + 0.25*sd;
run;

  data match_temp;
    if _n_=1 then set ctrl_temp(keep=low up);
                  set case_temp(drop=rannum);
      if low<=logit<=up;
      run;


*****************************************************************************;
*** calculate number of patients who were included in the match dataset      ***
*** 1) if there is no match found, then select next patient from case group  ***
*** 2) if there is one matched, select this patient and go to next round     ***
*** 3) if there are more than 1 patients, calculate Mahalanobis distance and ***
***    select the patient with smallest distance, then go to next round      ***
*****************************************************************************;
  proc sql;
    create table casen
      as select count(&id) as n
      from match_temp;
      quit;

  data _null_;
    set casen;
      if _n_=1 then call symput('casen',put(n,8.));
      run;

  %if %eval(&casen)=0 %then %do;
    data case_ctrl_together;
         set case_ctrl_together match_temp(in=a keep=&id);
         if &id=. then delete;
         if a then newid=&i;
         run;
  %end;

    %else %if %eval(&casen)=1 %then %do;

      data match_temp1;
             set ctrl_temp match_temp;
             newid=&i;
      run;

    data case_ctrl_together;
         set case_ctrl_together match_temp1(keep=&id newid);
         if &id=. then delete;
         run;

      proc sql;
             create table case_temp
             as select a.*
             from case_temp a,  match_temp b
             where a.&id^=b.&id;
      quit;


  %end;

  %else %do;

***please specify the key variables (including propensity score as one key
variable) which are used to calculate the mahalanobis distance ****;
```

```
    %Mahalanobis(match_temp, XXXXX, ctrl_temp);
  %end;

%end;

***Create test dataset that contains patients from case_control_together, and all
variables and propensity scores as well***;


proc sql;
  create table test
  as select a.*, b.newid
  from propen a, case_ctrl_together b
  where a.&id=b.&id
  order by newid,&yvar;
  quit;

%mend match;

                          *** This is the end of macro ***
```


ACKNOWLEDGMENT

Thanks are given to Ilya Lipkovich, Rayamajhi Jyoti, Baojin Zhu and Douglas Faries for reviewing the paper and providing valuable comments. We also thank for SAS technical group for their generous support.

CONTACT INFORMATION

Wuwei Wayne Feng M.S.
Lilly Corporate Center, DC: 4136
Eli Lilly & Company
Indianapolis, IN, 46285
cnwwfeng@rocketmail.com

Jun Yu M.S.
2800 River Road
MedFocus Ltd.
Des Plaines, IL, 60018
Jyu1977@yahoo.com

Rong Michelle Xu MS
Lilly Corporate Center, DC: 5114
Eli Lilly & Company
Indianapolis, IN, 46285
rxu@lilly.com