PhUSE EU Connect 2018

Paper RW03

The Propensity Score Matching

Stephanie Fechtner, Bayer AG, Wuppertal, Germany

ABSTRACT

For evaluating the effect of an intervention, a randomized-controlled trial is the gold standard. But what can a statistician do, if the implementation of a randomized-controlled trial is not possible, due to ethical reasons for example? And how can an effect of an observational study be analyzed without getting biased estimates due to unobserved heterogeneity in the data? One possible answer to these questions can be the propensity score. It can be used to match patients of a treatment group with patients of a control group to ensure balanced sample statistics and therefore unbiased estimates. In this paper the propensity score and corresponding terminologies will be defined and different matching methods will be introduced. Furthermore, it will be explained how the propensity score matching can be implemented in R and SAS®. Finally, the advantages and disadvantages of the propensity score method compared to a multiple linear regression will be discussed.

INTRODUCTION

A randomized-controlled trial is the gold standard for evaluating the effect of an intervention or a treatment. It results, on average, in balanced sample characteristics of a treatment and a control group and the effect can be estimated without fearing biased results. Whereas in nonrandomized or observational studies, it is not possible to assume that the results are unbiased when estimating a treatment effect. It is much more likely that the effect is over- or underestimated due to different baseline sample statistics which would therefore lead to a misjudgment of the treatment effect.

The propensity score method was proposed by Rosenbaum and Rubin in 1983 as a tool to adjust for confounding and to reduce selection bias in nonrandomized or observational studies and will be explained in the next section. Propensity scores can be used in different ways to control for confounding in the comparison of outcomes between treatment groups. One way is to create two groups with similar baseline covariates by finding partners with a similar propensity score. This matching method will be also explained in the next section.

It is possible to implement the propensity score matching within the statistical programming languages R and SAS. In the subsequent section, an observational study started in 1948 will be used as an example to show how these implementations can be performed. In the final section, the limitations and the strength of the propensity score method will be discussed.

THE PROPENSITY SCORE

The propensity score E_i was introduced by Rosenbaum and Rubin and is defined as the conditional probability of receiving treatment Z_i, given a vector of observed covariates X_i, i=1,…,n:

$$E_i = P(Z_i \mid X_i=x_i)= \exp(x_i\beta)/[1+\exp(x_i\beta)] = 1/[1+\exp(-x_i\beta)] \Leftrightarrow \log[E_i/(1-E_i)] = x_i\beta. \quad (1)$$

Z_i denotes the binary treatment condition (Z_i=1, if patient i is in the treatment group and Z_i=0, if patient i is in the control group), β the vector of regression parameters and n the number of patients in the study. The covariates X_i are variables which are associated as potential confounders regarding the main analysis. The probability E_i can be estimated by using the maximum likelihood estimator of β resulting from the logistic regression model (1) considering the predefined potential confounders X_i.

Rosenbaum and Rubin defined three key approaches where the estimated propensity scores can be used for the evaluation of a treatment effect. The scores can be used for covariate adjustment, for stratification or for pair matching. To learn more about the covariate adjustment or the stratification, the interested reader is referred to Guo and Fraser (2010) or Rosenbaum and Rubin (1983).

In the next subsection the pair matching of the propensity scores will be explained. This method is used to construct two groups with balanced baseline covariates out of the treatment and the control group.

THE PROPENSITY SCORE MATCHING

If matching is performed by using propensity scores, the usual problem of multidimensionality in finding concurring partners is avoided and reduced to finding another person with a similar propensity score. To ensure that the created samples, which should be used for the estimation of the treatment effect afterwards, have similar distributions of the measured baseline covariates, a sufficient overlapping of the propensity scores is required. If this requirement is fulfilled, two matched samples from the treatment and the control group can be created by finding partners with similar propensity scores.

If the requirement of a sufficient overlapping is not fulfilled, a matching would not make sense or not be possible, if the matching is restricted in some way. In this case, the differences between the baseline covariates of the treatment and the control group are so large, that similar partners are not able to be found; an evaluation of the treatment effect would not make sense with these groups anyway. If, for example, just elderly women are in the treatment group and only younger men belong to the control group, the pure treatment effect can not be assessed; it would always be biased by the age and the gender of the patients.
Figure 1 gives an example of two different samples of propensity scores where the propensity score matching makes sense for one case and where it does not for the other case.

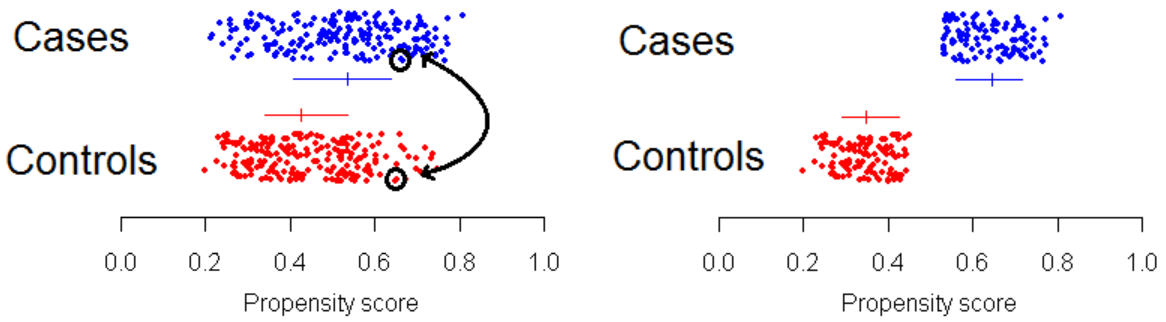

Figure 1: Sufficient overlapping of propensity scores (left side) → matching is useful; No overlapping of propensity scores (right side) → matching is not useful

In case the propensity score matching makes sense and should be used for the creation of two similar groups, the statistician needs to decide how the two groups should be created. He/she can choose between different matching methods like the exact matching, the nearest neighbor matching or the optimal matching. If the exact matching is chosen, for each patient in the treatment group all patients of the control group with exactly the same baseline covariates are matched. If the nearest neighbor matching is chosen, for each patient in the treatment group the best patient in the control group is defined as partner using a specified distance measure. Partners will be found one at a time and at each matching step a control patient is found as partner for a treatment patient who is not yet matched but is closest to that treatment patient (Ho et al., 2011). In contrast, the optimal matching finds the matched samples with the smallest average absolute distance across all matched samples. In addition to these methods, other matching procedures are also available and can be looked up in Ho et al. (2011).
Besides the matching procedure, it needs to be decided about the proportion of the treatment and the control group (1:1, 2:1, …) and if the matching should be performed with or without replacement. Furthermore a maximal distance of the propensity scores of the treatment and the control group, the so-called “caliper”, could be used as an additional restriction for the matching. By determining a caliper the sample size decreases, but at the same time the similarity of cases and controls increases. Austin (2010) investigated the optimal choice of the caliper with the aid of Monte Carlo simulations. He specified the optimal choice of the caliper as 0.2 of the standard deviation of the propensity scores, but any other caliper value greater than 0 is possible.
In the next section, the Framington Heart Study will be investigated as a suitable example for the propensity score matching. The dataset “Heart” within the library Sashelp contains the data of this study and will be used to show how the propensity score matching can be implemented in the statistical programming languages SAS and R.

EXAMPLE – THE FRAMINGTON HEART STUDY

In 1948 the systematic investigation of the population of the city Framingham (Massachusetts) began. This study is still ongoing and the aim was to detect causes and risks of the Coronary heart disease (CHD). The dataset “Heart” in the library Sashelp contains the data of the first 5209 adult patients which were included in the Framington Heart Study and 17 variables related to the CHD (SAS, 2017a). With the aid of `PROC CONTENTS`, the variables of the Heart dataset were explored first and the result of this SAS procedure is displayed in table 1.

Variables in Creation Order				
#	**Variable**	**Type**	**Length**	**Label**
1	Status	Char	5	
2	DeathCause	Char	26	Cause of Death
3	AgeCHDdiag	Num	8	Age CHD Diagnosed
4	Sex	Char	6	
5	AgeAtStart	Num	8	Age at Start
6	Height	Num	8	
7	Weight	Num	8	
8	Diastolic	Num	8	

9	Systolic	Num	8	
10	MRW	Num	8	Metropolitan Relative Weight
11	Smoking	Num	8	
12	AgeAtDeath	Num	8	Age at Death
13	Cholesterol	Num	8	
14	Chol_Status	Char	10	Cholesterol Status
15	BP_Status	Char	7	Blood Pressure Status
16	Weight_Status	Char	11	Weight Status
17	Smoking_Status	Char	17	Smoking Status

Table 1: The output of `PROC CONTENTS DATA=Sashelp.heart;`

One interesting question could be if smoking has a relevant impact on the development of CHD. The easiest way would be to perform a Chi-squared test or a logistic regression model on the given data. But due to the fact, that patients were not allocated to the Smoking status at random (which would be unethical) the investigator faces the problem of analyzing an observational study. This means, differences in observed baseline values could lead to biased results when analyzing the influence of smoking on CHD.

It is known that some other characteristics, like the age of a person, the cholesterol status, the status of the weight or a high blood pressure can have an influence on the development of CHD besides the smoking status (see NHLBI, 2015). These variables can therefore affect the reliability of the results of the main analysis.

Table 2 illustrates seven baseline characteristics of smokers and non-smokers which are recorded in the Sashelp dataset “Heart” and could possibly affect the main analysis. Missing data were not considered in table 2. In this example, a smoker is defined as a person who smokes more than 5 cigarettes during one day.

	Smoker (n = 2093)	**Non-Smoker (n = 3116)**	**P value***
Sex: female, %	35.93	68.07	< 0.01
AgeAtStart, mean (SD)	42.16 (8.04)	45.35 (8.68)	< 0.01
Chol_Status			0.15
% Borderline	38.20	35.85	
% Desirable	27.80	27.77	
% High	34.00	36.38	
Diastolic, mean (SD)	83.92 (12.33)	86.33 (13.30)	< 0.01
Systolic, mean (SD)	133.50 (20.90)	139.20 (25.21)	< 0.01
BP_Status			< 0.01
% High	38.84	46.66	
% Normal	43.72	39.41	
% Optimal	17.44	13.93	
Weight_Status			< 0.01
% Normal	35.79	23.26	
% Overweight	59.76	73.92	
% Underweight	4.45	2.83	

Table 2: Baseline characteristics of smokers (> 5 cigarettes) and non-smokers (≤ 5 cigarettes) recorded in the Sashelp dataset “Heart”; SD = standard deviation

* p-values resulting from a t-test or a chi-squared test

As the baseline values are quite different between the group of smokers and non-smokers (see Table 2), the propensity score will be used to create samples with similar baseline values, in order to minimize the potential influence of these baseline measurements on the main analysis. For that, the seven variables shown in table 2 will be considered as covariates in the propensity score model. Even if no significant differences between smokers and non-smokers in the cholesterol status were found, this variable will also be used as covariate, to ensure that the characteristics are still similar between smokers and non-smokers after the matching. The propensity score

matching can be implemented in both programming languages R and SAS and will be explained in the following subsections with the aid of the previous example.
In each subsection a smoker is defined as a person who smokes more than 5 cigarettes per day. The group of smokers is defined as the treatment group and the group of non-smokers as control group. The propensity scores will be calculated by using a logistic regression model and the matching of the propensity scores will be done using the nearest neighbor method and with replacement, meaning that a non-smoker which was already found as a suitable partner for a smoker, could also be a partner for another smoker. Furthermore a ratio of 2:1 will be used, meaning here that 2 non-smokers will be found as partners for one smoker. To ensure that the created groups will be similar, the recommended caliper of 0.2 will be used (Austin, 2011).

THE PROPENSITY SCORE MATCHING IN SAS

To perform the propensity score matching in SAS, the procedure `PROC PSMATCH` can be used (see SAS, 2017b). Observations without any missing values in the baseline covariates are considered in the propensity score matching. Furthermore a variable `CHD` is created which indicates if a person suffers from a coronary heart disease or not.

```
/*Cholesterol and Weight_Status contain missing data -> exclude these observations*/
data heart;
      set sashelp.heart(where=(Cholesterol ne . and Weight_Status ne ""));
      if Smoking <= 5 then Smoker=0;   /*Create variable 'Smoker'*/
          else if Smoking > 5 then Smoker=1;
      if AgeCHDdiag = . then CHD=0;    /*Create variable 'CHD' (1=Yes, 0=No)*/
          else if AgeCHDdiag ne . then CHD=1;
      keep Sex AgeAtStart Diastolic Systolic Chol_Status BP_Status Weight_Status
           Smoker CHD; /*Keep only variables which are needed for the analysis*/
run;

ods graphics on;
proc psmatch data=heart region=allobs;
   class Smoker Sex Chol_Status BP_Status Weight_Status; /*Define class variables*/
   psmodel Smoker(Treated='1')= Sex Chol_Status BP_Status Weight_Status AgeAtStart
           Diastolic Systolic; /*Specifies logistic regression model for computing the
                                 propensity scores*/
   match method=replace(k=2) distance=ps caliper=0.2; /*Define the matching method*/
   assess ps / plots=(boxplot); /*Assess differences in propensity scores (=PS) before
                                  and after the matching and create boxplots*/
   output out(obs=match)=m_heart matchid=_MatchID; /*Create output dataset 'm_heart'
                                 which contains the data of the matched samples*/
run;
ods graphics off;
```

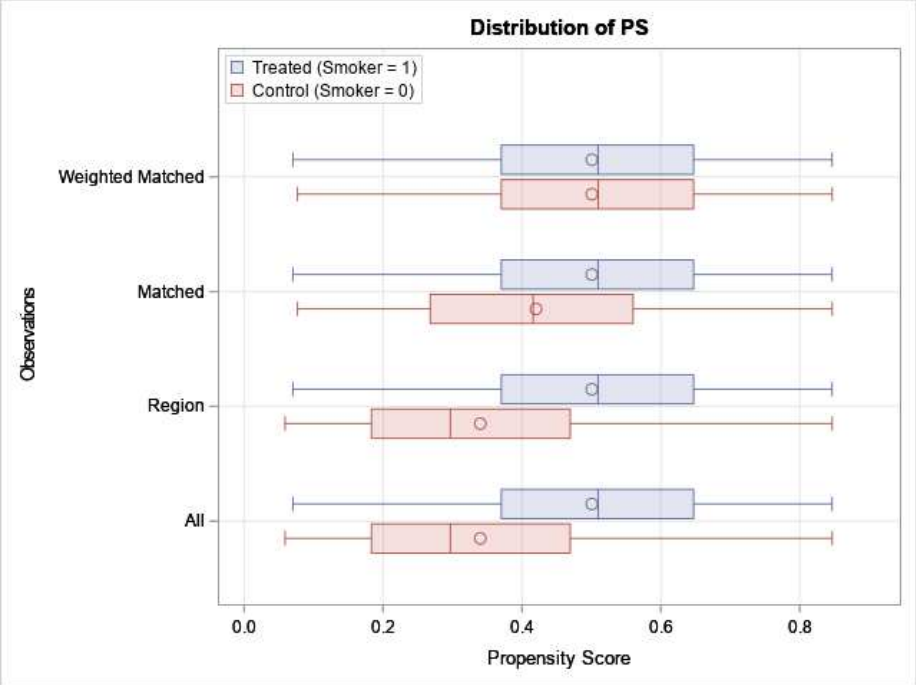

Figure 2: Distributions of the propensity scores before (=All) and after the matching (=Matched & Weighted Matched)

The distributions of the propensity scores are much more similar after the matching than before (see Figure 2). This applies especially for the weighted matched propensity scores which also consider the weight of a matched person. The more often a non-smoker was found as a suitable partner for a smoker, the greater the weight of this non-smoker. With the aid of the procedures `PROC TTEST` and `PROC FREQ` the baseline characteristics included in the created dataset `m_heart` can be investigated by weighted t-tests and weighted chi-square tests. These weighted tests are able to consider the weight of each matched person.
After the matching no significant difference in the seven considered covariates between smokers and non-smokers could be found anymore (p-values > 0.05, see Table 3).

	Smoker (n = 2047)[a]	Non-Smoker (n = 1656)	P value[b]
Sex: female, %	35.56	36.47	0.55
AgeAtStart, mean (SD)	42.11 (8.02)	41.87 (9.37)	0.37
Chol_Status			0.054
% Borderline	38.25	35.83	
% Desirable	27.80	31.17	
% High	33.95	33.00	
Diastolic, mean (SD)	83.98 (12.35)	83.74 (13.27)	0.56
Systolic, mean (SD)	133.59 (20.99)	132.48 (22.41)	0.10
BP_Status			0.67
% High	38.89	37.81	
% Normal	43.77	43.92	
% Optimal	17.34	18.27	
Weight_Status			0.66
% Normal	35.86	36.27	
% Overweight	59.65	58.70	
% Underweight	4.49	5.03	

Table 3: Weighted characteristics of the matched groups of smokers (> 5 cigarettes) and non-smokers (≤ 5 cigarettes) based on the Sashelp dataset “Heart”; mean = weighted mean; SD = weighted standard deviation
[a] For all smokers suitable partners could be found. The number reduced due to the fact that patients with missing data were excluded.
[b] p-values resulting from weighted t-tests or weighted chi-squared tests, see SAS-code below.

```
/*Weighted t-tests to investigate AgeAtStart, Diastolic and Systolic (=VARIABLE)*/
proc ttest data=m_heart sides=2;
      class smoker;
      var VARIABLE;
      weight _MATCHWGT_;
run;

/*Weighted chi-square tests to investigate sex, Chol_Status, BP_Status and
Weight_Status (=VARIABLE)*/
proc freq data=m_heart;
   tables smoker*VARIABLE / expected cellchi2 norow nocol chisq;
   weight _MATCHWGT_;
run;
```

The matched samples within the dataset `m_heart` can be used now to investigate the influence of smoking on the development of CHD. To perform the main analysis a weighted logistic regression model or a weighted chi-square test can be used without fearing biased results due to differences in the baseline values of the seven considered covariates.

THE PROPENSITY SCORE MATCHING IN R

To calculate the propensity scores in R, the package `MatchIt` can be used (Ho et al., 2011). For the propensity score matching in R it is required, that all covariates do not contain any missing values. This means, that a variable which contains a lot of missing values should not be considered as covariate in the propensity score model.

```
set.seed(21122602)
heart <- read.csv("…/heart.csv", sep=",", header=TRUE) # … contains the file path of the dataset
## Add column, if patient has CHD
CHD <- matrix(0, length(heart[,3]))
heart2 <- cbind(heart[,-2], CHD)
heart2[which(is.na(heart[,3])==TRUE), which(names(heart2)=="CHD")] <- 0
heart2[which(is.na(heart[,3])==FALSE), which(names(heart2)=="CHD")] <- 1
## Add column, if smoker (yes or no)
Smoker <- matrix(0, length(heart[,3]))
heart3 <- cbind(heart2, Smoker)
heart3[which(heart3[,which(names(heart3)=="Smoking")]<=5), which(names(heart3)=="Smoker")] <- 0
heart3[which(heart3[,which(names(heart3)=="Smoking")]>5), which(names(heart3)=="Smoker")] <- 1
## Exclude patients with missing data
heart4 <- heart3[which(is.na(heart3$Cholesterol)==FALSE & heart3$Weight_Status!=""),]

### Perform the propensity score matching
library(MatchIt)
formel <- "Smoker ~ as.factor(Sex) + AgeAtStart + Diastolic + Systolic + as.factor(Chol_Status) +
as.factor(BP_Status) + as.factor(Weight_Status)"
heart5 <- heart4[c("Smoker", "Sex", "AgeAtStart", "Diastolic", "Systolic", "Chol_Status",
"BP_Status", "Weight_Status")] # heart5 contains only needed variables for PS matching

## matchit() performs calculation of propensity scores
match_heart <- matchit(as.formula(formel), method="nearest", distance="logit", replace=TRUE,
ratio=2, caliper=0.2, discard="none", data=heart5)

## Create a dataset which just includes matched patients
m.all <- cbind(heart4, match_heart$distance, match_heart$weights) # m.all including propensity
scores and weights
names(m.all)[c(dim(m.all)[2]-1, dim(m.all)[2])] <- c("distances", "weights") # Rename the columns
of the propensity scores and the weights
m.heart <- subset(m.all,weights!=0) # m.heart contains only matched observations (weight > 0)

### Display the distributions of the propensity scores before and after the matching
set.seed(1563)
## Create weighted samples out of the matched smokers and non-smokers
pscore.treated.matched <- sample(m.heart$distances[m.heart,"Smoker"]==1], 10000, replace=TRUE,
prob=m.heart$weights[m.heart,"Smoker"]==1])
pscore.control.matched <- sample(m.heart$distances[m.heart,"Smoker"]==0], 10000, replace=TRUE,
prob=m.heart$weights[m.heart,"Smoker"]==0])
## Create variables containing the propensity scores of all smokers and non-smokers
pscore.control.all <- match_heart$distance[heart4,"Smoker"]==0]
pscore.treated.all <- match_heart$distance[heart4,"Smoker"]==1]

### Create 4 histograms of the propensity scores containing also the density
par(mfcol = c(2, 2))
hist(pscore.treated.all, xlab="Propensity Score", ylab="Density", main ="Unmatched Treated
(Smoker)", freq=FALSE, breaks=10, xlim=c(0,1), ylim=c(0,3))
lines(density(pscore.treated.all))

hist(pscore.control.all, xlab="Propensity Score", ylab="Density", main ="Unmatched Control (Non-
Smoker)", freq=FALSE, breaks=10, xlim=c(0,1), ylim=c(0,3))
lines(density(pscore.control.all))

hist(pscore.treated.matched, xlab="Propensity Score", ylab="Density", main ="Matched Treated
(Smoker)", freq=FALSE, breaks=10, xlim=c(0,1), ylim=c(0,3))
lines(density(pscore.treated.matched))

hist(pscore.control.matched , xlab="Propensity Score", ylab="Density", main ="Matched Control
(Non-Smoker)", freq=FALSE, breaks=10, xlim=c(0,1), ylim=c(0,3))
lines(density(pscore.control.matched))
```

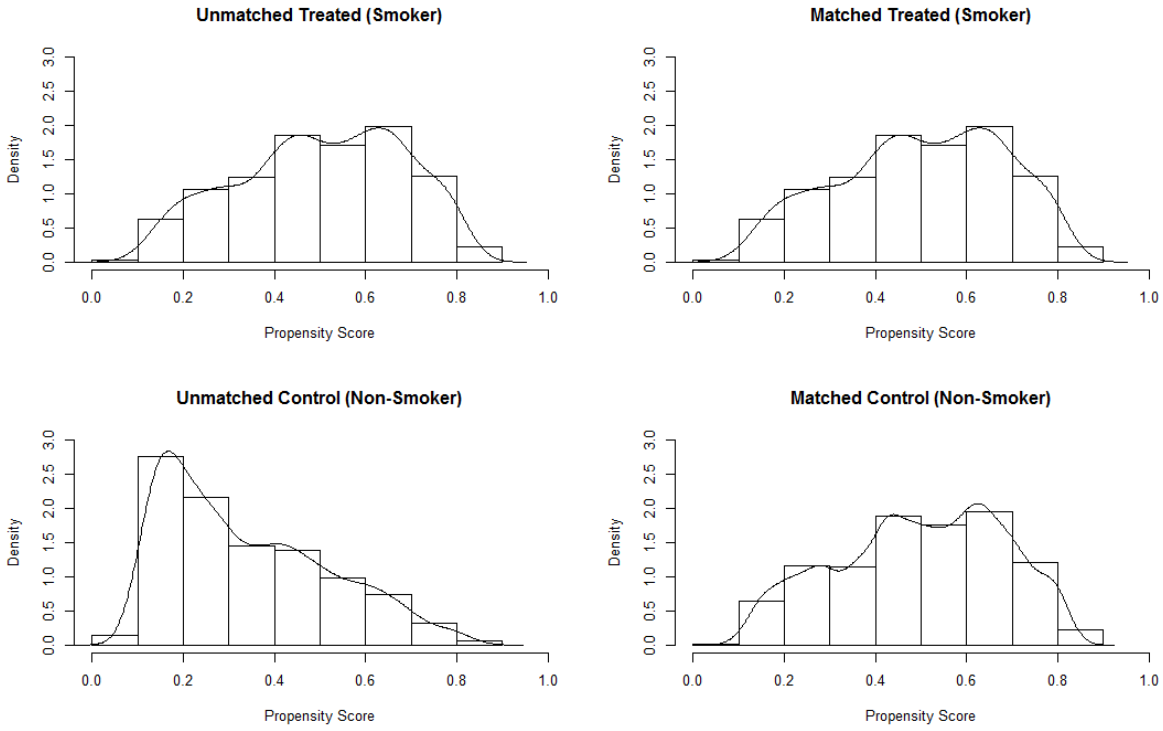

Figure 3: Distributions of the propensity scores before (=Unmatched) and after the matching (=Matched)

Like in SAS, the distributions of the propensity scores of smokers and non-smokers are quite different before the matching and very similar after the matching (see Figure 2 and 3). With the aid of `wtd.t.test` and `wtd.chi.sq` of the R-package `weights`, weighted t-tests and weighted chi-square tests showed that there was no significant difference of the baseline values between the smokers and the non-smokers anymore (all p-values > 0.05, see Table 4).

	Smoker (n = 2047)[a]	**Non-Smoker (n = 1804)**	**P value[b]**
Sex: female, %	35.56	35.74	0.91
AgeAtStart, mean (SD)	42.11 (8.02)	42.18 (8.40)	0.78
Chol_Status			0.52
% Borderline	38.25	39.28	
% Desirable	27.80	28.51	
% High	33.95	32.22	
Diastolic, mean (SD)	83.98 (12.35)	84.09 (11.91)	0.77
Systolic, mean (SD)	133.59 (20.99)	133.54 (20.00)	0.94
BP_Status			0.76
% High	38.89	39.30	
% Normal	43.77	44.26	
% Optimal	17.34	16.44	
Weight_Status			0.54
% Normal	35.86	34.83	
% Overweight	59.65	59.99	
% Underweight	4.49	5.18	

Table 4: Weighted characteristics of the matched groups of smokers (> 5 cigarettes) and non-smokers (≤ 5 cigarettes) based on the Sashelp dataset "Heart"; mean = weighted mean; SD = weighted standard deviation

[a] For all smokers suitable partners could be found. The number reduced due to the fact that patients with missing data were excluded.

[b] p-values resulting from weighted t-tests or weighted chi-squared tests, see R-code below.

```
library("weights")
# Weighted chi-square tests to investigate sex, Chol_Status, BP_Status, Weight_Status (=VARIABLE)
wtd.chi.sq(m.heart$Smoker, m.heart$VARIABLE, weight=m.heart$weights)
# Weighted t-tests to investigate AgeAtStart, Diastolic and Systolic (=VARIABLE)
wtd.t.test(m.heart[which(m.heart$Smoker==0),]$VARIABLE,
           m.heart[which(m.heart$Smoker==1),]$VARIABLE,
           weight=m.heart[which(m.heart$Smoker==0),]$weights,
           weighty=m.heart[which(m.heart$Smoker==1),]$weights, samedata=FALSE)
```

The matched samples within the dataset `m.heart` can be used now to investigate the influence of smoking on the development of CHD without fearing biased results due to differences in the baseline values of the seven considered covariates.

ADVANTAGES AND DISADVANTAGES OF THE PROPENSITY SCORE MATCHING

If propensity scores are used for the matching of two groups, a treatment and a control group are created that have similar covariate values. One advantage of the propensity score matching is, that a large number of covariates can be considered and that subsequent comparisons are not confounded by differences in the distributions of these covariates anymore. Results based on propensity score methods are more robust, more precise and less biased than results based on regression models, especially when a large number of covariates are considered or the number of observations is low (Cepeda et al., 2003). Furthermore the propensity score analysis does not rely on the correct specification of the functional form between the covariates and the outcome. This is also an advantage over the implementation of a linear regression model, where the functional form between the covariates and the outcome needs to be linear. Another advantage of the propensity score matching is the separation of the confounder adjustment and the outcome analysis. This idea could be easier to explain and is more intuitive than in a regression model (Zanutto, 2006).

Unfortunately there are also disadvantages, which need to be considered when implementing the propensity score matching. First, the propensity score matching leads to a selective loss of study participants because just those controls are considered in the analysis which are similar to the treatment group. No statement can be made about unmatched patients. Second, the two created groups are only balanced in all observed covariates. The analysis could still be biased due to unobserved heterogeneity between the propensity score matched samples. Therefore the implementation of a randomized controlled trial should always be preferred, if possible, because it results, on average, in balanced sample characteristics of the treatment and the control group, also in unobserved covariates. Third, the effects of the confounders or any interactions on the outcome are not estimated when the propensity score matching is performed. In this case, a regression model has the advantage over the propensity score matching where any effects are measured (Zanutto, 2006).

CONCLUSION

The propensity score matching is a very valuable method, if a statistician wants to evaluate the treatment effect in an observational or a nonrandomized study, and can easily be implemented within both programming languages R and SAS. This method offers some advantages over the implementation of a regression model and its idea is easy to understand even for a nontechnical audience.

Nevertheless, the analyst needs to be careful with the interpretation of the subsequent main analysis because the evidence level of the analysis results is still as high as the evidence level of an observational study.

REFERENCES

Austin PC (2011). *Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies*. Pharm Stat 10: 150-161.

Cepeda MS, Boston R, Farrar JT, Strom BL (2003): *Comparison of logistic regression versus propensity score when the number of events is low and there are multiple confounders*. Am J Epidemiol. 158(3): 280-287.

Ho DE, Imai K, King G, Stuart EA (2011). *MatchIt: nonparametric preprocessing for parametric causal inference*. J Stat Softw 42:1–28.

NHLBI – National Heart, Lung and Blood Institute (2015). *Coronary Heart Disease*. Bethesda, Maryland. https://www.nhlbi.nih.gov/health-topics/coronary-heart-disease (from 26.09.2018)

R Core Team (2015). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rosenbaum PR, Rubin DB (1983). *The central role of the propensity score in observational studies for causal effects*. Biometrika 70:41–55.

SAS Institute Inc. (2017a). *Sashelp Data Sets*. Cary, NC, USA.

SAS Institute Inc. (2017b). *SAS/STAT 14.3 User's Guide, The PSMATCH Procedure*. Cary, NC, USA.

Zanutto EL (2006). *A comparison of Propensity Score and Linear Regression Analysis of Complex Survey Data*. Journal of Data Science 4: 67-91.

RECOMMENDED READING

Guo S, Fraser MW (2010). *Propensity score analysis: statistical methods and applications*. Los Angeles, CA: Sage Publications.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Author Name	Stephanie Fechtner
Company	Bayer AG
Address	Aprather Weg 18a
City / Postcode	Wuppertal, 42096, Germany
Work Phone:	+49 202 36-3833
Email:	Stephanie.fechtner@bayer.com
Web:	www.bayer.com