Paper 184-2007

# Using Propensity Scores to Adjust For Treatment Selection Bias

Scott Leslie, MedImpact Healthcare Systems, Inc., San Diego, CA
Patrick Thiebaud, MedImpact Healthcare Systems, Inc., San Diego, CA

## ABSTRACT

Estimating the effect of drug treatment on outcomes requires adjusting for many observed factors, particularly those influencing drug selection. This paper demonstrates the use of PROC LOGISTIC in creating propensity scores to address such potential treatment selection bias. In this example using a pharmacy claims database, this method evaluates the probability of patients being included in five treatment groups and calculates weights based on factors hypothesized to influence treatment selection. A propensity score-weighted regression model is then fitted to compare the outcome of adherence between groups and to study the possible predictors of adherence.

## INTRODUCTION

A limitation of observational studies is the lack of treatment assignment. This can lead to large differences in treatment groups that should be adjusted for to reduce selection bias and better clarify the effect of treatment. Regression adjustment, matching, and stratification using propensity scores are widely used techniques to compare groups, usually comparing a treatment group to a non treatment group. The purpose of this paper is to provide an example of using propensity score regression adjustment to balance five treatment groups.

## DESCRIPTION OF STUDY

In this study, we identified patients using lipid lowering and antihypertensive therapy from a large pharmacy claims database. Patients were categorized into five treatment groups with the main objective of comparing compliance and adherence rates. Compliance was measured as the proportion of days a medication was supplied over a 180 day period. Adherent patients were identified as those reaching a threshold of 80% compliance. Demographic variables (age, gender) and previous medication use was measured in a 6 month baseline period prior to treatment.

Table 1 describes the treatment groups. PROC GLM was used to compare groups. For most covariates, there were significant differences among treatment groups (p<.0001).

**Table 1. Unadjusted Demographic and Baseline Measures**

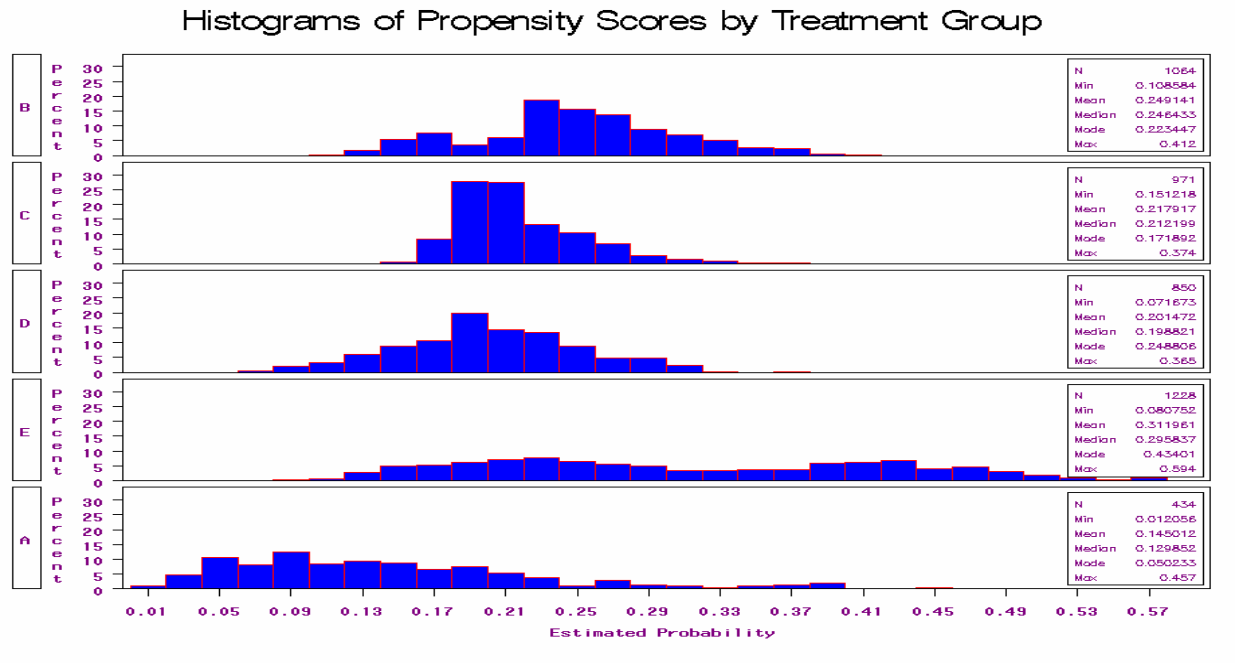
| UNADJUSTED VALUES | | A | B | C | D | E | p-value |
|---|---|---|---|---|---|---|---|
| Member Count | | 434 | 1,064 | 971 | 850 | 1,228 | |
| Age | Mean | 57.2 | 61.8 | 62.3 | 61.7 | 63.9 | <.0001 |
| | SD | 13.1 | 13.5 | 13.0 | 13.1 | 12.8 | |
| Female | % | 39.8% | 50.3% | 48.4% | 53.8% | 53.1% | <.0001 |
| HMO Business Type | % | 78.3% | 75.2% | 68.2% | 68.4% | 50.8% | <.0001 |
| Open Formulary | % | 73.0% | 53.1% | 53.7% | 47.5% | 49.3% | <.0001 |
| # of baseline drugs | Mean | 3.9 | 3.7 | 3.7 | 3.1 | 3.1 | <.0001 |
| | SD | 3.7 | 4.1 | 4.3 | 4.1 | 4.2 | |
| Maintenance Medication Refill % | % | 38.8% | 33.7% | 33.8% | 26.0% | 26.8% | <.0001 |
| ACE Inhibitor | % | 21.4% | 21.2% | 21.0% | 15.7% | 15.3% | <.0001 |
| Beta-blocker | % | 16.5% | 20.0% | 20.3% | 12.0% | 13.3% | <.0001 |
| Diuretic | % | 4.1% | 6.1% | 6.9% | 4.8% | 5.4% | 0.1813 |
| ARB | % | 18.8% | 9.1% | 10.5% | 6.2% | 7.8% | <.0001 |
| Coronary Vasodilator | % | 4.8% | 6.2% | 6.3% | 6.2% | 4.3% | 0.1348 |
| Digoxin | % | 1.1% | 1.0% | 2.1% | 1.7% | 1.7% | 0.2876 |
| Platelet Aggregation Inhibitor | % | 4.3% | 4.3% | 4.5% | 2.8% | 3.3% | 0.2398 |
| Diabetes | % | 14.9% | 16.1% | 16.5% | 12.5% | 15.0% | 0.1500 |
| Antidepressant | % | 13.1% | 12.6% | 11.6% | 9.6% | 9.6% | 0.0518 |

## CREATING PROPENSITY SCORES

The propensity score is the conditional probability of each patient receiving a particular treatment based on pre-treatment variables. Using the LOGISTIC procedure, propensity scores were calculated based on the 15 covariates listed in Table 1. The objective was to balance the treatment groups so to reduce bias of treatment selection and obtain better idea of treatment effect on the outcome of compliance. The generalized logit function is specified in the LINK option to contrast the reference group to the other four groups and the RSQUARE option assesses the amount of variation explained by the independent variables. The propensity score is output to data set named ps.

```
proc logistic data=ccb;
model cohort = age female pre_drgcnt hmo f_open ace betab diur arb nitrate
antiplatelet cvddig diabetes antidepressant maintrefillratio
/link=glogit rsquare
output out = ps pred = ps;
run;
```

After creating the propensity scores, an evaluation of the distributions by treatment groups checks for sizeable overlap among the groups demonstrating that the groups are comparable.




Next a propensity score weight, also referred to as the inverse probability of treatment weight (IPTW), is calculated as the inverse of the propensity score.

```
data ps_weight;
set ps;
ps_weight=1/ps;
if cohort =_level_;
run;
```

As of now the weights are based on the entire study group and would give more weight to the smaller treatment groups. A SQL procedure creates a weight that reflects the sample size for each of the treatment groups.

```
proc sql;
create table ps_weight_adj as
select *, (count(*)/4547)*ps_weight as ps_weight_adj
from ps_weight
group by cohort;
quit;
```

Table 2 describes the treatment groups based on the propensity score weights. The effect of the propensity score weights was to balance the groups. Comparisons among treatment groups found no significant differences, with exception to the diuretic covariate.

**Table 2. Propensity Score Weighted Demographic and Baseline Measures**

| PROPENSITY SCORE WEIGHTED VALUES | | A | B | C | D | E | p-value |
|---|---|---|---|---|---|---|---|
| Member Count | | 434 | 1,064 | 971 | 850 | 1,228 | 4,547 |
| Age | Mean | 62.6 | 61.6 | 62.1 | 62.0 | 61.6 | 0.6203 |
| | SD | 13.8 | 13.5 | 13.1 | 13.0 | 13.2 | |
| Female | % | 52.7% | 49.9% | 50.3% | 49.6% | 50.9% | 0.8539 |
| HMO Business Type | % | 67.5% | 66.7% | 66.0% | 66.4% | 67.0% | 0.9769 |
| Open Formulary | % | 54.5% | 53.6% | 52.6% | 53.1% | 53.5% | 0.9730 |
| # of baseline drugs | Mean | 3.8 | 3.5 | 3.5 | 3.5 | 3.6 | 0.6102 |
| | SD | 3.7 | 4.0 | 4.2 | 4.3 | 4.3 | |
| Maintenance Medication Refill % | % | 33.9% | 30.8% | 30.8% | 31.5% | 30.9% | 0.6252 |
| ACE Inhibitor | % | 19.0% | 18.6% | 18.7% | 19.5% | 19.0% | 0.9882 |
| Beta-blocker | % | 20.5% | 16.2% | 16.4% | 16.8% | 16.8% | 0.3324 |
| Diuretic | % | 10.5% | 5.7% | 5.6% | 5.7% | 5.9% | 0.0035 |
| ARB | % | 9.4% | 9.8% | 9.3% | 9.7% | 9.7% | 0.9941 |
| Coronary Vasodilator | % | 6.4% | 5.6% | 5.6% | 5.9% | 5.3% | 0.9310 |
| Digoxin | % | 1.5% | 1.5% | 1.5% | 1.5% | 1.4% | 0.9999 |
| Platelet Aggregation Inhibitor | % | 3.9% | 3.8% | 3.8% | 3.7% | 3.7% | 0.9999 |
| Diabetes | % | 18.1% | 15.3% | 15.1% | 14.9% | 16.0% | 0.5930 |
| Antidepressant | % | 12.6% | 11.3% | 11.0% | 11.0% | 11.6% | 0.9166 |

## PROPENSITY SCORE WEIGHTED OUTCOME MODEL

Next, a propensity score-weighted logistic regression model was fitted to compare the outcome of adherence. Table 3 shows the comparisons of treatment groups. A second model was run to assess the diuretic covariate.

```
proc logistic data=ps_weight_adj;
class cohort (ref="A") / param=reference ;
model pdc_80 (event="1") = cohort / rsquare clodds=wald lackfit
weight ps_weight_adj / normalize;
format cohort cohortreg.;
run;
```

**Table 3. Propensity Score Weighted Outcome Models**

| Effect | Odds Ratio | 95% Confidence Limits | p-value | Effect | Odds Ratio | 95% Confidence Limits | p-value |
|---|---|---|---|---|---|---|---|
| B vs. A | 0.823 | 0.658 - 1.029 | 0.0869 | B vs. A | 0.815 | 0.652 - 1.019 | 0.0726 |
| C vs. A | 0.706 | 0.563 - 0.886 | 0.0026 | C vs. A | 0.699 | 0.557 - 0.877 | 0.0020 |
| D vs. A | 0.867 | 0.688 - 1.092 | 0.2254 | D vs. A | 0.858 | 0.681- 1.082 | 0.1959 |
| E vs. A | 0.726 | 0.583 - 0.903 | 0.0040 | E vs. A | 0.719 | 0.577 - 0.895 | 0.0031 |
| | | | | Diuretic | 0.818 | 0.641 - 1.044 | 0.1061 |

## CONCLUSION

This paper presents a scenario where compliance outcomes are compared among five treatments groups using an inverse propensity score weighted logistic regression model. This is an additional method that can be used in conjunction with other regression adjustment techniques, such as propensity score matching, propensity score subclassification, and multivariable logistic regression, to reduce bias and better describe the effect of treatment.

## REFERENCES

Rosenbaum, P.R. and Rubin D.B. 1983. “The Central Role of the Propensity Score in Observational Studies for Causal Effects”, *Biometrika*, 70, 41-55.

D’Agostino, R. 1998. “Tutorial on Biostatistics: Propensity Score Methods for Bias Reduction in the comparison of a treatment to a non-randomized control group”. *Statistics in Medicine* 17, 2265-2281.

Kurth, T., et al. 2006. “Results of Multivariate Logistic Regression, Propensity Matching, Propensity Adjustment, and Propensity-based Weighting under Conditions of Nonuniform Effect”. *American Journal of Epidemiology* 206: 163:262-270.

Pasta, David J. 2000. “Using Propensity Scores to Adjust for Group Differences: Examples Comparing Alternative Surgical Methods”. *Proceedings of the Twenty-Fifth Annual SAS Users Group International Conference,* Indianapolis, IN, 261-25.

SAS Institute Inc. 2004. “SAS Procedures: The LOGISTIC Procedure”. *SAS OnlineDoc® 9.1.3.* Cary, NC: SAS Institute Inc.
http://support.sas.com/documentation/onlinedoc/91pdf/sasdoc_91/stat_ug_7313.pdf

## ACKNOWLEDGMENTS

The author would like to thank Eunice Chang for her consultation and review.

## CONTACT INFORMATION

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

Scott Leslie
MedImpact Healthcare Systems, Inc.
10680 Treena Street
San Diego, CA 92131
Work Phone: 858-790-6685
Fax: 858-689-1799
Email: scott.leslie@medimpact.com