Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Yizhao Xi Date

Assess Balance of Covariates after Propensity Score as Covariate Adjustment: SAS® macro Development and Application

By

Yizhao Xi

Master of Science in Public Health

Biostatistics and Bioinformatics

Yuan Liu, PhD
(Thesis Advisor)

Azhar Nizam, MS
(Reader)

Assess Balance of Covariates after Propensity Score as Covariate Adjustment: SAS® macro Development and Application

By

Yizhao Xi

B.S.
Tianjin University
2016

Thesis Committee Chair: Yuan Liu, PhD
Reader: Azhar Nizam, MS

An abstract of
A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Biostatistics and Bioinformatics
2018

Table of Contents

Abstract

Assess Balance of Covariates after Propensity Score as Covariate Adjustment: SAS® macro Development and Application

By Yizhao Xi

Propensity score (PS) is a method used to reduce selection bias or confounding effects in observational studies, and it is defined as the conditional probability of treatment assignment given observed baseline covariates. Covariate adjustment using propensity score is one of the popular method applied in public health studies. It involves regression adjustment where outcome is regressed on a treatment indicator and on estimated propensity score. Whenever a covariate adjustment is performed, the balance between two groups should be evaluated. Weighted standardized differences have been introduced to assess the quality of balancing between treated and untreated subjects after propensity score adjustment. Imbalances in the regression adjustment should be adjusted for further study when analyzing outcomes.

However, there are few goodness-of-fit tests for this method in practical application due to the lack of a user-friendly statistical tool or software packages. In this study, a SAS® macro is developed for performing the balance diagnosis using weighted standardized difference with PS as covariate adjustment, and compared that with standardized difference without PS as covariate adjustment. A list of covariates for both categorical and continuous baseline covariates will be analyzed at one time. The macro will create two RTF files by ODS style, and one is a summary table and the other one is a graph that display the Display the covariate balance improvement after PS adjustment by the weighted standardized difference from the standardized difference before PS adjustment. Additionally, a case study was utilized to illustrate the application of the SAS® macro and summary all the findings and reports. The results indicate that the weighted standardized difference is a feasible and practical measurement to assess the balance of covariate adjustment using propensity score. With this SAS® macro development, covariate adjustment using propensity score can be more easily applied in practice.

1. Introduction

There is increasing interest in using observational studies to estimate the effects of treatments and exposure of interest where patients are not assigned to exposure or treatment under the investigators' control. Since treated subjects may differ systematically from untreated subjects, the effect of treatment on health outcomes could be influenced by treatment-selection bias.[1] Differences in outcome may be incorrectly attributed to the exposure where there are important differences in confounder between each exposure groups.[2]

Propensity score (PS) was defined by Rosenbaum and Rubin[3] as the conditional probability of treatment assignment given observed baseline covariates. In other words, for a sets of subjects who have the same propensity score, treated and untreated ones will have similar distribution of baseline covariates. [4] From an individuals' point of view, the allocation of treatment could be considered as random. It assumes that two patients with the same propensity score would have an equal estimated probability of treatment. Thus, even though propensity score is not able to control for unknown confounders, it estimates the sensitivity of the model to unknown confounders and alerts investigators to focus on the selection bias.[1]

Four PS methods are commonly applied in the public health area[5]: matching on propensity score, stratification, inverse probability of treatment weighting (IPTW) using propensity score and covariate adjustment after propensity score. Matching entails generating matched subjects of treated and untreated cohorts with similar propensity score, then comparing the outcomes between two subjects. [6]It results in good balance, but information can be lost due to relatively smaller sample size after completing matching. In stratification on propensity score, patients are stratified into mutually exclusively strata based on estimated propensity score, and a pooled stratum-specific estimates of outcome is calculated.

[4]This method maintains a larger sample size than the matching approach, but also results in more heterogeneity within each stratum. IPTW method creates a sample by weighs based on propensity score where distribution of baseline covariates and treatment assignment are independent. But it gives imprecise estimates of the treatment effect when significant confounding exists. [7]In covariate adjustment using propensity scores, the outcome is regressed on an indicator variable denoting treatment status and estimated propensity scores.[4, 8] Compared with other three methods, it doesn't separate the study design from study analysis; however, the relationship between propensity score and the outcome must be specified correctly.

The true propensity score is unknown in an observational study, but can be estimated using logistic regression based on study data when only two treatment groups are under consideration. It models the treatment as the dependent variable and potential confounders as the independent variables. The estimated propensity score is the predicted probability of treatment derived from the fitted regression model.[9] It is of vital important to examine whether the propensity score model has been adequately specified. The principal idea is to diagnose whether treated and untreated subjects with similar propensity scores shared similar distributions of baseline covariates. Appropriate methods for assessing balance have been developed the in the context of matching and stratification including standardized differences, comparing higher-order moments, five number summary statistics, and graphic methods.[10-12] Several studies have described that the standardized differences can be used to compare the balance of continuous and binary variables between treated and untreated groups, and that it is not influenced by study sample size. [13]Austin defined a new balance checking method named weighted conditional standardized difference by extending the current standardized difference in the context of covariate adjustment.[14] It is defines as the average standardized absolute difference between treated and untreated groups that share similar propensity score.

Although covariate adjustment using propensity scores appears frequently in the public health literatures, there are few goodness-of-fit test for this method in practical application due to the lack of existing SAS® macro or package in other programming software. In this study, a SAS® macro has been developed to perform the balance diagnosis after covariate adjustment using propensity score by weighted standardized difference, and to assess the balance improvement by comparing it with the standardized difference without the propensity score covariate adjustment. A case study is utilized to illustrate the application of the SAS® macro and a summary of the findings is reported.

2. Method

Adequately specification of the model is crucial to the application of covariate adjustment. The systematic differences between treated and untreated subjects would be eliminated by conditioning on propensity score. The standardized difference is employed to examine the degree to which systematic differences have been removed based on the propensity score.[4, 15] For continuous covariate, the standardized difference is defined as

$$\mathrm{d} = \frac{(\bar{x}_{trt} - \bar{x}_{untrt})}{\sqrt{\frac{s^2_{trt} + s^2_{untrt}}{2}}} \quad (1)$$

where $\bar{x}_{trt} - \bar{x}_{untrt}$ denotes the difference in mean baseline covariate value between two groups and s^2_{trt}, s^2_{untrt} are the sample variance of the baseline covariates for each group. For categorical variable, it is defined as

$$\mathrm{d} = \frac{(\hat{p}_{trt} - \hat{p}_{untrt})}{\sqrt{\frac{\hat{p}_{trt}(1-\hat{p}_{trt}) + \hat{p}_{untrt}(1-\hat{p}_{untrt})}{2}}} \quad (2)$$

where $\hat{p}_{trt}$ and $\hat{p}_{untrt}$ denote the prevalence of the categorical covariates in two groups respectively.

Weighted standardized difference denotes as integrating the standardized difference in order to avoid having positive and negative difference cancel each other out.[14] It integrated over the distribution of estimated propensity score which is not uniform within range. In other words, weighted standardized difference reflects the average standardized difference between treated and untreated subjects based on same propensity score.

Once the propensity score estimates are obtained from the logistic regression, the weighted standardized difference for each covariate is calculated using a regression model where covariate X is treated as the outcome, and let T denote the treatment selection (T=1 for treated subjects, T=0 for untreated subjects), Z be the estimated propensity score from the logistic regression model. The interaction term between the treatment indicator and the propensity score is utilized to differentiate the mean difference of covariates between the two groups and given propensity score.[14] If X is a continuous covariate, a linear regression model is adapted as equation 3, otherwise we use a multinomial logistic regression model as equation 4.

$$\mathrm{X} = \beta_0 + \beta_1 T + \beta_2 Z + \beta_3 T * Z \tag{3}$$

$$\mathrm{logit}(\Pr(\mathrm{X} = 1)) = \beta_0 + \beta_1 T + \beta_2 Z + \beta_3 T * Z \tag{4}$$

For a given propensity score (Z), the predicted value for a continuous baseline covariate or the estimated logit for a categorical baseline covariate can be obtained as: $\beta_0 + \beta_1 + (\beta_2 + \beta_3) * Z$ for treated subjects and $\beta_0 + \beta_2 * Z$ for untreated subjects

Thus, in the SAS@ macro, we first evaluated the standardized difference at the individual level by fitting the interaction model (equation 3 or 4) to obtain the estimated probability of X at different treatment level after controlling for propensity score Z, and then fit the formula of standardized difference for each subject. Weighted standardized differences of each covariate are obtained by finding average of the absolute standardized difference across all subjects.

3. SAS macro

SAS@ macro %WEIGHTED_STD was developed to run tests for checking the balance of covariate adjustment using propensity score. The macro is able to process a list of categorical and continuous baseline covariates without interruption, and the final reports are in RTF format and include a table that lists distribution of each covariates between two treatment groups and standardized difference before PS adjustment and weighted standardized difference after PS adjustment and a graph shows the covariate balance improvement before and after PS covariate adjustment A complete copy of %WEIGHTED_STD is included in the Appendix

3.1 Standardized difference calculation

Standardized difference of each level of character baseline covariates without considering covariate adjustment was calculated as the column percentage from the FREQ procedure. For each level of covariates, the proportions of treated and untreated subjects are equal to the column percentage of each treatment group. The frequencies of groups are also displayed in the report table. For numeric baseline covariates, mean of treated and untreated subjects are generated using the class statement in PROC MEANS, and standard deviations are also obtained in this step. Based on the equations 2 and 3, standardized difference before covariate adjustment was obtained in both character and numeric covariate.

3.2 Weighted standardized difference calculation

3.2.1 Categorical covariates

Weighted standardized differences after covariate adjustment are based on the linear or multinomial logistic regression model described in equation 1. The flow chart below presents the calculation steps.

First, for a given character covariate X, the number of levels in X were determined using the FREQ procedure and X was fitted in a multinomial regression model using PROC LOGISTIC. Taking a three-level baseline covariate as an example, the estimated probability of each level in treated and untreated subjects can be obtained by applying the treatment and propensity score information of each observation into the multinomial regression model (equation 1). They correspond to the two equations below:

$$\log\left(\frac{\Pr\left(Level1\right)}{\Pr\left(Reference\ Level\right)}\right) = \beta_{10} + \beta_{11} * TRT + \beta_{12} * PS + \beta_{13} TRT * PS \quad (5)$$

$$\log\left(\frac{\Pr\left(Level2\right)}{\Pr\left(Reference\ Level\right)}\right) = \beta_{20} + \beta_{21} * TRT + \beta_{22} * PS + \beta_{23} TRT * PS \quad (6)$$

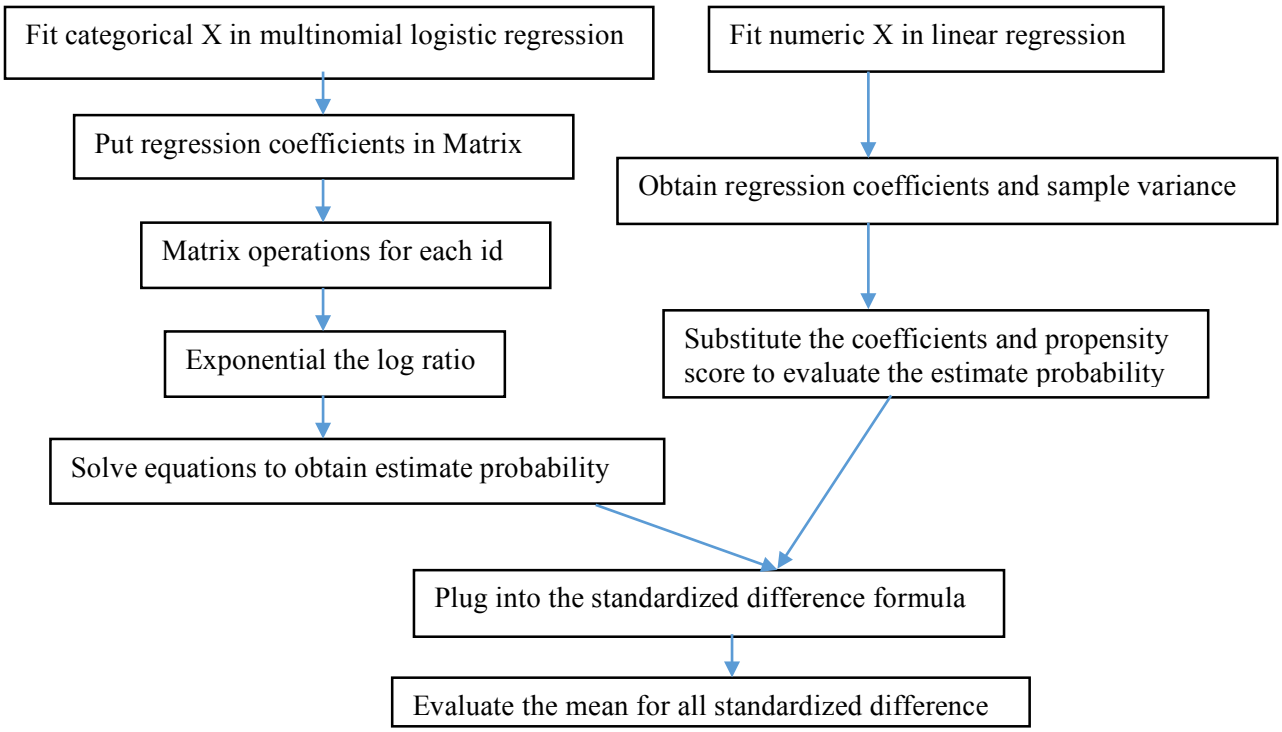
Second, all β coefficients gained from regression step were put into a (n-1) x4 matrix by IML procedure where n denotes the number of level in the covariate and 4 is the number of β coefficients. Since the estimated probability of the originally assigned treatment group could be obtained in the PROC LOGISTIC step, we only need to evaluate the estimated probability of the covariate in the other treatment. Thus, a 4 x m matrix was generated where m refers to the number of observations in the study. Each row in the matrix implies the intercept, indicator of treatment, propensity score and interaction term of propensity score and treatment indicator respectively. The matrix operation for each id is shown below:

$$\begin{pmatrix} \beta_{10}+\beta_{21}+\beta_{31}+\beta_{41} \\ \vdots \\ \beta_{(n-1)0}+\beta_{(n-1)1}+\beta_{(n-1)2}+\beta_{(n-1)3} \end{pmatrix} \begin{pmatrix} 1 \\ trt \\ PS \\ trt*PS \end{pmatrix} = \begin{pmatrix} \log\left(\frac{\Pr\left(Level1\right)}{\Pr\left(Reference\ Level\right)}\right) \\ \vdots \\ \log\left(\frac{\Pr\left(Level\ n-1\right)}{\Pr\left(Reference\ Level\right)}\right) \end{pmatrix}$$

Then, the ratio of probability choosing one category over the probability choosing the reference category can be obtained by exponentiating the linear equation 5 and 6. And the summation of estimate probability of each category is 1. Therefore, estimate probability of correspond treatment group can be obtained from solvable equations. At last, after applying the estimate probability into standardized difference formula (equation 3), the average of the standardized difference across all subjects is expressed as the weighted standardized difference.

3.2.2 Numeric covariates

Similarly, fit a given numeric covariate into a linear regression by PROC GENMOD to gain the estimated regression coefficients for treatment indicator, estimated propensity score and interaction term and square root of the estimate residual variance. Then, substitute the coefficients and propensity score for each subject ID into the regression model to obtain the estimate probability. Each observation in the study would have two estimate probability, one is for the one in the treated group and the other is for in the untreated group. Finally, after applying estimate probability into the standardized difference formula, weighted standardized difference is determined across all subjects by the mean of above quantity.



3.3 Parameter interpretation

SAS@ macro %WEIGHT_STD requires the user to provide the name of the datasets with patient id, the estimated propensity score, treatment variable and all covariates of interest. The treatment variable should have two non-missing values and code as 1 and 0 respectively. Each categorical covariate required to

have at least two non-missing value. If the value of standardized differences exceed criteria set by user, it would be bold in latter report table. The macro would create two RTF files by ODS style, one is the summary table and the other one is a graph to display the difference between standardized differences and weighted standardized differences. The two files could be generated separately and would not interact each other. This macro was written and tested using SAS@ version 9.4. I describe all the macro variables below:

DATASET is the name of the data set to be analyzed and it is required.

TRT is the variable name for comparison groups and the number of groups should be exactly two.

CLIST is a list of character variables and not include treatment variable, separated by a space.

NLIST is a list of numeric variables and not include treatment variable, separated by a space.

CRITERIA is the threshold of standardized difference can be used to indicated sufficient balance. The default value is 0.1.

DOC is the option for creating or suppress the RTF file. The default value is T.

OUTPATH is the path for output table and graph to be stored.

FNAME is the file name for output table and file for output graph would be plot &FNAME.

ORIENTATION is the value of PORTRAIT or LANDSCAPE to indicate the page layout of the report. The default value is PORTRAIT.

PLOT is the option for creating or suppress the graph. The default value is T.

DEBUG sets to T if running in debug mode. Work datasets will not be deleted in debug mode. This is useful if you are editing the code or want to further manipulate the resulting data sets. The default value is F.

4. Case Study

4.1 Background

Mantle cell lymphoma is a distinct subtype of B-cell non-Hodgkin lymphoma, and it comprises about 6% of all NHL cases. The incidence of MCL has been increasing steadily over the past several decades. Its age-adjusted incidence rate is about 0.51 to 0.55 per 100,000 person-years. The symptoms of MCL includes swelling, night sweats, fever and weight loss.[16] While multiple risk demographic factors are associated with outcome in a number of malignancies**,** few studies have examined the impact of race and socioeconomic. This study targets at assessing treatment disparities and outcomes by races and socioeconomic status in patients with MCL.

4.2 Data source

The National Cancer Database (NCDB) is a nationwide oncology outcomes database. It collects clinical data from over 1,500 Commission on Cancer (CoC) approved hospitals in the United States each year and sponsored by the American Cancer Society and the American College of Surgeons. The database features approximately 70% newly diagnosed cancer cases in the US, it used to explore trends in cancer care and serve as the basis for quality improvement activities as well. [17] In this study, the 2014 NCDB Participant User File (PUF) for lymphoma was used to select patients.

The database was queried for patients diagnosed with lymphoma from 2004 to 2013. Patients with histology as mantle cell were eligible to enroll in the study. We also excluded patients with missing information on overall survival. Demographic and socioeconomic characteristics that were examined included race (Hispanic, non-Hispanic), sex, age at diagnosis, primary payer (private insured, government insured, or uninsured), facility type (community cancer program, comprehensive community cancer program, or academic/research program), medium income quartile and education level. Tumor characteristics that were examined included Charlson-Deyo Score (surrogate for patient comorbidities),

nodal status (internal, external), systemic symptoms, AJCC stage, sequence number and treatment started from diagnosis. The primary outcome was overall survival defined as weeks between date of diagnosis and death.

4.3 Statistical Analysis

Statistical analysis was conducted using SAS version 9.4 and SAS macros developed Biostatistics and Bioinformatics department of Rollins School of Public Health, Emory University. Descriptive statistic for each baseline covariate were reported in table 2. The univariate association between each covariate and race group were assessed using the Chi-square test for character covariates and ANOVA for numeric covariates. The primary clinical outcome is overall survival and it is defined as months from date of diagnosis to date of death or last follow up if alive. The significant level was set at 0.5. Cox proportional hazard models were utilized to assess the univariate and multivariate association between overall survival and covariates of interest. Multivariate models were fit by a backward variable selection method applying an $\alpha = 0.2$ removal criteria.

In order to reduce selection bias, the propensity score of each patient was estimated by %CALC_PS (V7) in which treatment assignment (White Non-Hispanic vs. Hispanic) was regressed on the covariates of interest in a logistic regression. The propensity score calculation macro excluded patients with missing value at any of the 12 covariates. %WEIGHT_STD was applied into the dataset with estimate propensity score, criterion as 0.2 was set up in the macro result in the high disparities in propensity score among two race groups. Standardized differences before covariates adjustment were reported in table 1 while weighted standardized differences for each level of covariates were displayed in rightmost column of table 1. After covariate balanced by propensity score, the PS adjusted Hazard ratio for race is obtained by Cox proportional hazard model where overall survival was regressed on the race and propensity score.

4.4 Results

After applying the inclusion and exclusion criteria, a total of 16,789 patients were enrolled in this study. Of 16,789 patients, 1,096 (6.5%) were Hispanic and 15,693 (93.5%) were non-Hispanic. Descriptive analysis for other socioeconomic factors is displayed in table 2 where 6,333 (39.3%) patients had private insurance, 568 (3.5%) had Medicaid and 8,807 (54.7%) had Medicare. Table 2 presents the univariate association between race group and baseline covariate. Briefly, Hispanic patients generally have better survival rate than White Non-Hispanic group. They tend to residence in zip codes with lower income and decreased percentage high school degree. Also, Medicaid or uninsured status, more systemic symptoms at diagnosis and increased treatment at academic/research program were associated with Hispanic race group. Univariate and multivariate association between race and overall survival are reported in table 4 and table 5 respectively. In a multivariate analysis, Hispanic ethnicity (1.16, (1.03-1.30), p-value 0.013), private insurance status, high education level and young age are all associated with improving overall survival rate. Kaplan-Meier curve by group is shown in figure 2, the 5-year survival for Hispanic group is 55.8% (52.2%, 59.3%) and for non-Hispanic is 49.9% (49.0%, 50.8%).

Standardized differences for baseline covariates are evaluated in the table 1. There are 5 of 12 covariates of interest had standardized difference exceeding 0.2 including facility type, insurance status, median income, education level (high school percent) and age at diagnosis. It indicated the imbalance of these baseline covariates among difference race group. In contrast, for weighted standardized differences, only one (education level) of these covariates exceeding 0.2. Since there is no uniformly threshold as a meaningful indicator for standardized difference, it is acceptable for this study that there is no potentially meaningful imbalance after covariate adjustment using propensity score.[4, 18] From figure 1, all points of weighted standardized difference displayed within the threshold line. It's clear to detail that covariate adjustment using propensity score helps balance baseline covariates. After covariate balance by propensity score, the adjusted hazard ratio for non-Hispanic versus Hispanic is 1.11 (p-value 0.0749).

5. Discussion

Methods for balance checking in propensity score matching and stratifying have been developed for use. But there is no ready-to-use tool for assessing balance in covariate adjustment using propensity score. Austin proposed the weighted standardized difference method to compare the difference for baseline covariate between treated and untreated subjects. It is an intuitive index to evaluate balance in observational study.

A standard macro to conduct balance diagnosis after covariate adjustment using propensity score with clean and readable table and graph allows for thorough testing of balance. %WEIGHT_STD runs a series of tests quickly and could be employed in baseline covariates with multiple levels as needed. This macro has better performance at automaticity, a list of covariates could be tested in one step and summary output will be generated directly and clearly. One of the limitations in the macro is, it could only be applied in a study with two treatment groups. Another is that there is no uniformly accepted threshold to determine the significant of standardized difference. An empirical guidance to use standardized difference is less than 0.1 could be viewed as an indication of negligible imbalance, and other applicator call less than 0.25 as sufficient balance.

In conclusion, an easy-to-use SAS@ macro was developed to assess the balance diagnosis for covariate adjustment using propensity score. It creates summary tables including the percentage and frequency about each covariate between treated and untreated subjects, standardized differences before covariate adjustment and weighted standardized differences afterward. This macro also generates a graph for displaying the comparison of two kinds of differences intuitively. According to results of case study, the weighted standardized differences is a practical measurement for evaluating the balance of covariate

adjustment using propensity score. Given this new handy tool, covariate adjustment using propensity score can be extensively applied in the further medical and public health research.

6. Reference

1. Pattanayak, C.W., D.B. Rubin, and E.R. Zell, *Propensity Score Methods for Creating Covariate Balance in Observational Studies.* Revista Española de Cardiología (English Edition), 2011. **64**(10): p. 897-903.
2. Shah, B.R., et al., *Propensity score methods gave similar results to traditional regression modeling in observational studies: a systematic review.* J Clin Epidemiol, 2005. **58**(6): p. 550-9.
3. Rosenbaum, P.R. and D.B. Rubin, *The central role of the propensity score in observational studies for causal effects.* Biometrika, 1983. **70**(1): p. 41-55.
4. Austin, P.C., *An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies.* Multivariate Behav Res, 2011. **46**(3): p. 399-424.
5. Austin, P.C. and M.M. Mamdani, *A comparison of propensity score methods: a case-study estimating the effectiveness of post-AMI statin use.* Statistics in medicine, 2006. **25**(12): p. 2084-2106.
6. Rosenbaum, P.R. and D.B. Rubin, *Constructing a control group using multivariate matched sampling methods that incorporate the propensity score.* The American Statistician, 1985. **39**(1): p. 33-38.
7. Lunceford, J.K. and M. Davidian, *Stratification and weighting via the propensity score in estimation of causal treatment effects: a comparative study.* Statistics in medicine, 2004. **23**(19): p. 2937-2960.
8. Elze, M.C., et al., *Comparison of Propensity Score Methods and Covariate Adjustment: Evaluation in 4 Cardiovascular Studies.* J Am Coll Cardiol, 2017. **69**(3): p. 345-357.
9. Austin, P.C., *Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples.* Stat Med, 2009. **28**(25): p. 3083-107.
10. Ali, M.S., et al., *Reporting of covariate selection and balance assessment in propensity score analysis is suboptimal: a systematic review.* J Clin Epidemiol, 2015. **68**(2): p. 112-21.

11. Austin, P.C., *The relative ability of different propensity score methods to balance measured covariates between treated and untreated subjects in observational studies.* Medical Decision Making, 2009. **29**(6): p. 661-677.
12. Morgan, S.L. and J.J. Todd, *A diagnostic routine for the detection of consequential heterogeneity of causal effects.* Sociological Methodology, 2008. **38**(1): p. 231-281.
13. Austin, P.C., *Using the standardized difference to compare the prevalence of a binary variable between two groups in observational research.* Communications in Statistics-Simulation and Computation, 2009. **38**(6): p. 1228-1234.
14. Austin, P.C., *Goodness-of-fit diagnostics for the propensity score model when estimating treatment effects using covariate adjustment with the propensity score.* Pharmacoepidemiol Drug Saf, 2008. **17**(12): p. 1202-17.
15. Flury, B.K. and H. Riedwyl, *Standard distance in univariate and multivariate analysis.* The American Statistician, 1986. **40**(3): p. 249-251.
16. Wang, Y. and S. Ma, *Racial differences in mantle cell lymphoma in the United States.* BMC cancer, 2014. **14**(1): p. 764.
17. Bilimoria, K.Y., et al., *The National Cancer Data Base: a powerful initiative to improve cancer care in the United States.* Ann Surg Oncol, 2008. **15**(3): p. 683-90.
18. Normand, S.-L.T., et al., *Validating recommendations for coronary angiography following acute myocardial infarction in the elderly: a matched analysis using propensity scores.* Journal of clinical epidemiology, 2001. **54**(4): p. 387-398.

7. Tables and Figure

Table 1 Balance diagnosis for covariate adjustment using propensity score

Covariate	Level	Statistics	Race Group		Standardized Difference	Weighted Standardized Difference
			Hispanic N=798	Non-hispanic N=11896		
Facility Type	Academic/Research Program	N (Col%)	56 (7.02)	1191 (10.01)	0.1074	0.1388
	Community Cancer Program	N (Col%)	245 (30.7)	5101 (42.88)	**0.2546**	0.1022
	Comprehensive Community Cancer Program	N (Col%)	374 (46.87)	4491 (37.75)	0.1853	0.1510
	Integrated Network Cancer Program	N (Col%)	123 (15.41)	1113 (9.36)	0.1847	0.1098
Sex	Female	N (Col%)	591 (74.06)	8537 (71.76)	0.0517	0.0081
	Male	N (Col%)	207 (25.94)	3359 (28.24)	0.0517	0.0081
Primary Payor	Medicaid	N (Col%)	77 (9.65)	236 (1.98)	**0.3320**	0.0421
	Medicare	N (Col%)	297 (37.22)	4863 (40.88)	0.0751	0.1084
	Not Insured	N (Col%)	97 (12.16)	350 (2.94)	**0.3542**	0.0634
	Private Insurance	N (Col%)	327 (40.98)	6447 (54.19)	**0.2670**	0.0888

Covariate	Level	Statistics	Race Group		Standardized Difference	Weighted Standardized Difference
			Hispanic N=798	Non-hispanic N=11896		
Median Income Quarti	$38,000-$47,999	N (Col%)	163 (20.43)	1598 (13.43)	0.1873	0.0686
	$48,000-$62,999	N (Col%)	198 (24.81)	2836 (23.84)	0.0227	0.1703
	$63,000 +	N (Col%)	181 (22.68)	3296 (27.71)	0.1159	0.1117
	<$38,000	N (Col%)	256 (32.08)	4166 (35.02)	0.0623	0.1347
Percent No High Scho	13-20%	N (Col%)	297 (37.22)	1314 (11.05)	**0.6425**	0.0885
	7.0-12.9%	N (Col%)	180 (22.56)	2846 (23.92)	0.0324	0.0620
	<7%	N (Col%)	175 (21.93)	4211 (35.4)	**0.3012**	0.1151
	>=21%	N (Col%)	146 (18.3)	3525 (29.63)	**0.2679**	**0.2538**
Charlson-Deyo Score	0	N (Col%)	630 (78.95)	9290 (78.09)	0.0208	0.0054
	1	N (Col%)	134 (16.79)	1972 (16.58)	0.0058	0.0198
	2+	N (Col%)	34 (4.26)	634 (5.33)	0.0500	0.0255
Nodal status	Extranodal	N (Col%)	666 (83.46)	10108 (84.97)	0.0415	0.0398
	Lymph node	N (Col%)	132 (16.54)	1788 (15.03)	0.0415	0.0398
Systemic Symptoms at	No	N (Col%)	441 (55.26)	7048 (59.25)	0.0806	0.0656
	Unknown	N (Col%)	233 (29.2)	3219 (27.06)	0.0476	0.0646
	Yes	N (Col%)	124 (15.54)	1629 (13.69)	0.0522	0.0210

Covariate	Level	Statistics	Race Group		Standardized Difference	Weighted Standardized Difference
			Hispanic N=798	Non-hispanic N=11896		
AJCC Analytic Stage	Stage I	N (Col%)	67 (8.4)	939 (7.89)	0.0184	0.0047
	Stage II	N (Col%)	72 (9.02)	859 (7.22)	0.0660	0.0513
	Stage III	N (Col%)	88 (11.03)	1648 (13.85)	0.0857	0.0347
	Stage IV	N (Col%)	487 (61.03)	7452 (62.64)	0.0333	0.0350
	Unknown	N (Col%)	84 (10.53)	998 (8.39)	0.0731	0.0454
Sequence Number	0-1	N (Col%)	42 (5.26)	711 (5.98)	0.0310	0.0114
	2+	N (Col%)	756 (94.74)	11185 (94.02)	0.0310	0.0114
Age at Diagnosis		Mean (Std)	64.75 (11.67)	67.22 (11.15)	**0.2184**	0.1447
Treatment started, Weeks from Diagnosis		Mean (Std)	4.39 (7.78)	3.97 (5.74)	0.0657	0.0057

Table 2 Descriptive Statistics for all variables of interest

Variable	Level	N (%) = 16789
Race Group	Hispanic	1096 (6.5)
	Non-hispanic	15693 (93.5)
Facility Type	Community Cancer Program	1653 (10.0)
	Comprehensive Community Cancer Program	6938 (41.8)
	Academic/Research Program	6425 (38.7)
	Integrated Network Cancer Program	1584 (9.5)
	Missing	189
Sex	Male	11928 (71.0)
	Female	4861 (29.0)
Primary Payor	Not Insured	398 (2.5)
	Private Insurance	6333 (39.3)
	Medicaid	568 (3.5)
	Medicare	8807 (54.7)
	Missing	683
Median Income Quartiles 2008-2012	Not Available	263
	<$38,000	2317 (14.0)
	$38,000-$47,999	3901 (23.6)
	$48,000-$62,999	4498 (27.2)
	$63,000 +	5810 (35.2)
Percent No High School Degree 2008-2012	Not Available	253
	>=21%	2190 (13.2)
	13-20%	3944 (23.9)
	7.0-12.9%	5688 (34.4)
	<7%	4714 (28.5)

Variable	Level	N (%) = 16789
Charlson-Deyo Score	0	13167 (78.4)
	1	2684 (16.0)
	2+	938 (5.6)
Nodal status	Lymph node	14057 (83.7)
	Extranodal	2732 (16.3)
Systemic Symptoms at Diagnosis	No	9771 (58.2)
	Yes	4187 (24.9)
	Unknown	2831 (16.9)
AJCC Analytic Stage Group	Stage I	1425 (8.5)
	Stage II	1209 (7.2)
	Stage III	2257 (13.4)
	Stage IV	10015 (59.7)
	Unknown	1883 (11.2)
Sequence Number	0-1	924 (5.5)
	2+	15865 (94.5)
os_censor	0	8738 (52.0)
	1	8051 (48.0)
Age at Diagnosis	Mean	67.33
	Median	68.00
	Minimum	20.00
	Maximum	90.00
	Std Dev	11.85
	Missing	0.00
Great Circle Distance (/50)	Mean	0.79
	Median	0.21
	Minimum	0.00
	Maximum	77.73
	Std Dev	2.87
	Missing	256.00

Variable	Level	N (%) = 16789
Treatment started, Weeks from Diagnosis	Mean	4.04
	Median	2.86
	Minimum	0.00
	Maximum	167.86
	Std Dev	5.88
	Missing	3162.00
Overall Survival (Months)	Mean	39.86
	Median	32.07
	Minimum	0.00
	Maximum	142.82
	Std Dev	32.91
	Missing	0.00

Table 3 Univariate Association with RACE

Covariate	Statistics	Level	Race Group: Hispanic N=1096	Race Group: Non-hispanic N=15693	Parametric P-value*
Facility Type	N (Col %)	Community Cancer Program	75 (7.02)	1578 (10.16)	**<.001**
	N (Col %)	Comprehensive Community Cancer Program	338 (31.65)	6600 (42.49)	
	N (Col %)	Academic/Research Program	501 (46.91)	5924 (38.14)	
	N (Col %)	Integrated Network Cancer Program	154 (14.42)	1430 (9.21)	
Sex	N (Col %)	Male	800 (72.99)	11128 (70.91)	0.142
	N (Col %)	Female	296 (27.01)	4565 (29.09)	
Primary Payor	N (Col %)	Not Insured	98 (9.42)	300 (1.99)	**<.001**
	N (Col %)	Private Insurance	386 (37.12)	5947 (39.47)	
	N (Col %)	Medicaid	125 (12.02)	443 (2.94)	
	N (Col %)	Medicare	431 (41.44)	8376 (55.6)	
Median Income Quartiles 2008-2012	N (Col %)	<$38,000	228 (21.09)	2089 (13.53)	**<.001**
	N (Col %)	$38,000-$47,999	246 (22.76)	3655 (23.66)	
	N (Col %)	$48,000-$62,999	258 (23.87)	4240 (27.45)	
	N (Col %)	$63,000 +	349 (32.28)	5461 (35.36)	

Covariate	Statistics	Level	Race Group: Hispanic N=1096	Race Group: Non-hispanic N=15693	Parametric P-value*
Percent No High School Degree 2008-2012	N (Col %)	>=21%	406 (37.52)	1784 (11.54)	**<.001**
	N (Col %)	13-20%	236 (21.81)	3708 (23.99)	
	N (Col %)	7.0-12.9%	246 (22.74)	5442 (35.21)	
	N (Col %)	<7%	194 (17.93)	4520 (29.25)	
Charlson-Deyo Score	N (Col %)	0	872 (79.56)	12295 (78.35)	0.300
	N (Col %)	1	174 (15.88)	2510 (15.99)	
	N (Col %)	2+	50 (4.56)	888 (5.66)	
Nodal status	N (Col %)	Lymph node	901 (82.21)	13156 (83.83)	0.159
	N (Col %)	Extranodal	195 (17.79)	2537 (16.17)	
Systemic Symptoms at Diagnosis	N (Col %)	No	578 (52.74)	9193 (58.58)	**<.001**
	N (Col %)	Yes	311 (28.38)	3876 (24.7)	
	N (Col %)	Unknown	207 (18.89)	2624 (16.72)	

Covariate	Statistics	Level	Race Group: Hispanic N=1096	Race Group: Non-hispanic N=15693	Parametric P-value*
AJCC Analytic Stage Group	N (Col %)	Stage I	98 (8.94)	1327 (8.46)	**0.016**
	N (Col %)	Stage II	89 (8.12)	1120 (7.14)	
	N (Col %)	Stage III	122 (11.13)	2135 (13.6)	
	N (Col %)	Stage IV	639 (58.3)	9376 (59.75)	
	N (Col %)	Unknown	148 (13.5)	1735 (11.06)	
Sequence Number	N (Col %)	0-1	49 (4.47)	875 (5.58)	0.121
	N (Col %)	2+	1047 (95.53)	14818 (94.42)	
Age at Diagnosis	N		1096	15693	**<.001**
	Mean		64.53	67.53	
	Median		65	68	
	Min		20	21	
	Max		90	90	
	Std Dev		12.64	11.76	

Covariate	Statistics	Level	Race Group: Hispanic N=1096	Race Group: Non-hispanic N=15693	Parametric P-value*
Treatment started, Weeks from Diagnosis	N		873	12754	**0.004**
	Mean		4.59	4	
	Median		2.86	2.86	
	Min		0	0	
	Max		167.86	147.29	
	Std Dev		8.1	5.7	

* The parametric p-value is calculated by ANOVA for numerical covariates and chi-square test for categorical covariates.

Table 4 Univariate association with overall survival

			Overall Survival (Months)		
Covariate	**Level**	**N**	**Hazard Ratio (95% CI)**	**HR P-value**	**Type3 P-value**
Race Group	Non-hispanic	15693	1.19 (1.08-1.31)	**<.001**	**<.001**
	Hispanic	1096	-	-	
Facility Type	Community Cancer Program	1653	1.57 (1.45-1.69)	**<.001**	**<.001**
	Comprehensive Community Cancer Program	6938	1.41 (1.35-1.49)	**<.001**	
	Integrated Network Cancer Program	1584	1.31 (1.21-1.42)	**<.001**	
	Academic/Research Program	6425	-	-	
Sex	Male	11928	1.04 (0.99-1.09)	0.146	0.146
	Female	4861	-	-	
Primary Payor	Not Insured	398	1.47 (1.25-1.72)	**<.001**	**<.001**
	Medicaid	568	1.65 (1.45-1.88)	**<.001**	
	Medicare	8807	2.45 (2.33-2.58)	**<.001**	
	Private Insurance	6333	-	-	
Median Income Quartiles 2008-2012	<$38,000	2317	1.47 (1.38-1.57)	**<.001**	**<.001**
	$38,000-$47,999	3901	1.26 (1.19-1.34)	**<.001**	
	$48,000-$62,999	4498	1.16 (1.09-1.22)	**<.001**	
	$63,000 +	5810	-	-	
Percent No High School Degree 2008-2012	>=21%	2190	1.35 (1.25-1.45)	**<.001**	**<.001**
	13-20%	3944	1.32 (1.24-1.40)	**<.001**	
	7.0-12.9%	5688	1.18 (1.11-1.25)	**<.001**	
	<7%	4714	-	-	
Charlson-Deyo Score	1	2684	1.69 (1.60-1.79)	**<.001**	**<.001**
	2+	938	2.78 (2.56-3.01)	**<.001**	
	0	13167	-	-	
Nodal status	Lymph node	14057	1.08 (1.02-1.15)	**0.011**	**0.011**
	Extranodal	2732	-	-	

			Overall Survival (Months)		
Covariate	**Level**	**N**	**Hazard Ratio (95% CI)**	**HR P-value**	**Type3 P-value**
Systemic Symptoms at Diagnosis	No	9771	0.70 (0.67-0.74)	**<.001**	**<.001**
	Unknown	2831	0.84 (0.78-0.90)	**<.001**	
	Yes	4187	-	**-**	
AJCC Analytic Stage Group	Stage II	1209	1.19 (1.05-1.34)	**0.005**	**<.001**
	Stage III	2257	1.49 (1.35-1.65)	**<.001**	
	Stage IV	10015	1.38 (1.26-1.50)	**<.001**	
	Unknown	1883	1.44 (1.30-1.60)	**<.001**	
	Stage I	1425	-	**-**	
Sequence Number	2+	15865	1.24 (1.13-1.36)	**<.001**	**<.001**
	0-1	924	-	**-**	
Age at Diagnosis		16789	1.06 (1.05-1.06)	**<.001**	**<.001**
Great Circle Distance (/50)		16533	0.97 (0.96-0.98)	**<.001**	**<.001**
Treatment started, Weeks from Diagnosis		13627	0.97 (0.96-0.98)	**<.001**	**<.001**

Table 5 Multivariable Survival Analysis of OS

		Overall Survival (Months)		
Covariate	**Level**	**Hazard Ratio (95% CI)**	**HR P-value**	**Type3 P-value**
Race Group	Non-hispanic	1.16 (1.03-1.30)	**0.013**	**0.013**
	Hispanic	-	-	
Facility Type	Community Cancer Program	1.15 (1.05-1.25)	**0.003**	**<.001**
	Comprehensive Community Cancer Program	1.13 (1.06-1.20)	**<.001**	
	Integrated Network Cancer Program	1.11 (1.01-1.21)	**0.031**	
	Academic/Research Program	-	-	
Sex	Male	1.15 (1.08-1.21)	**<.001**	**<.001**
	Female	-	-	
Primary Payor	Not Insured	1.51 (1.26-1.80)	**<.001**	**<.001**
	Medicaid	1.43 (1.23-1.66)	**<.001**	
	Medicare	1.10 (1.02-1.18)	**0.011**	
	Private Insurance	-	-	
Median Income Quartiles 2008-2012	<$38,000	1.16 (1.05-1.28)	**0.004**	**0.009**
	$38,000-$47,999	1.04 (0.96-1.13)	0.355	
	$48,000-$62,999	1.00 (0.93-1.07)	0.941	
	$63,000 +	-	-	
Percent No High School Degree 2008-2012	>=21%	1.19 (1.07-1.33)	**0.001**	**<.001**
	13-20%	1.21 (1.11-1.32)	**<.001**	
	7.0-12.9%	1.15 (1.07-1.23)	**<.001**	
	<7%	-	-	
Charlson-Deyo Score	1	1.38 (1.30-1.47)	**<.001**	**<.001**
	2+	1.93 (1.75-2.13)	**<.001**	
	0	-	-	
Nodal status	Lymph node	1.14 (1.05-1.22)	**<.001**	**<.001**
	Extranodal	-	-	

Covariate	Level	Overall Survival (Months)		
		Hazard Ratio (95% CI)	HR P-value	Type3 P-value
Systemic Symptoms at Diagnosis	No	0.71 (0.67-0.76)	**<.001**	**<.001**
	Unknown	0.82 (0.76-0.89)	**<.001**	
	Yes	-	-	
AJCC Analytic Stage Group	Stage II	1.16 (1.01-1.34)	**0.039**	**<.001**
	Stage III	1.48 (1.31-1.67)	**<.001**	
	Stage IV	1.62 (1.46-1.81)	**<.001**	
	Unknown	1.59 (1.40-1.82)	**<.001**	
	Stage I	-	-	
Sequence Number	2+	1.18 (1.07-1.31)	**0.002**	**0.002**
	0-1	-	-	
Age at Diagnosis		1.05 (1.05-1.06)	**<.001**	**<.001**
Treatment started, Weeks from Diagnosis		0.97 (0.97-0.98)	**<.001**	**<.001**

* Number of observations in the original data set = 16789. Number of observations used = 12694.
** Backward selection with an alpha level of removal of .20 was used. No variables were removed from the model.

Figure 1 Illustration of balance improvement before and after PS adjustment

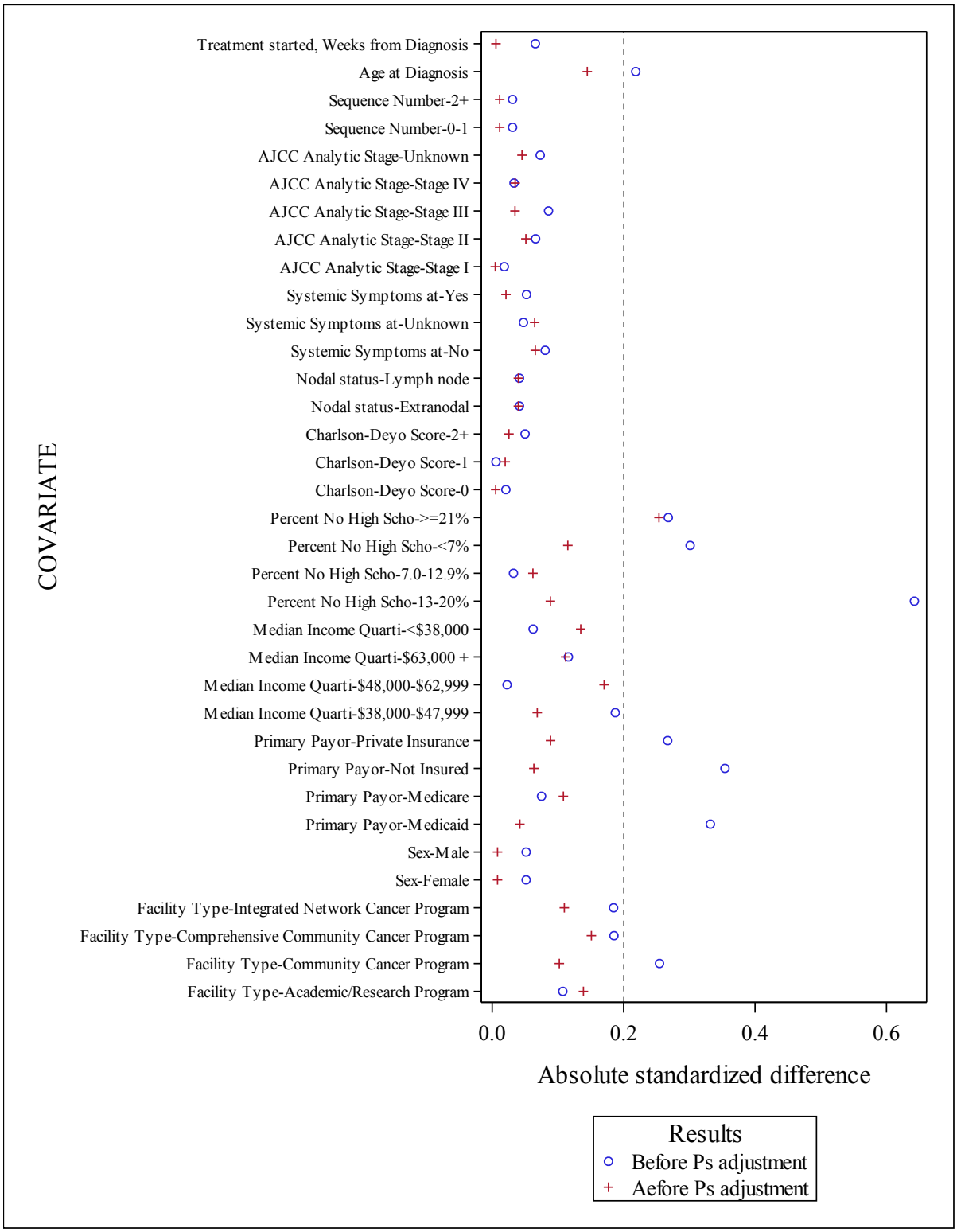
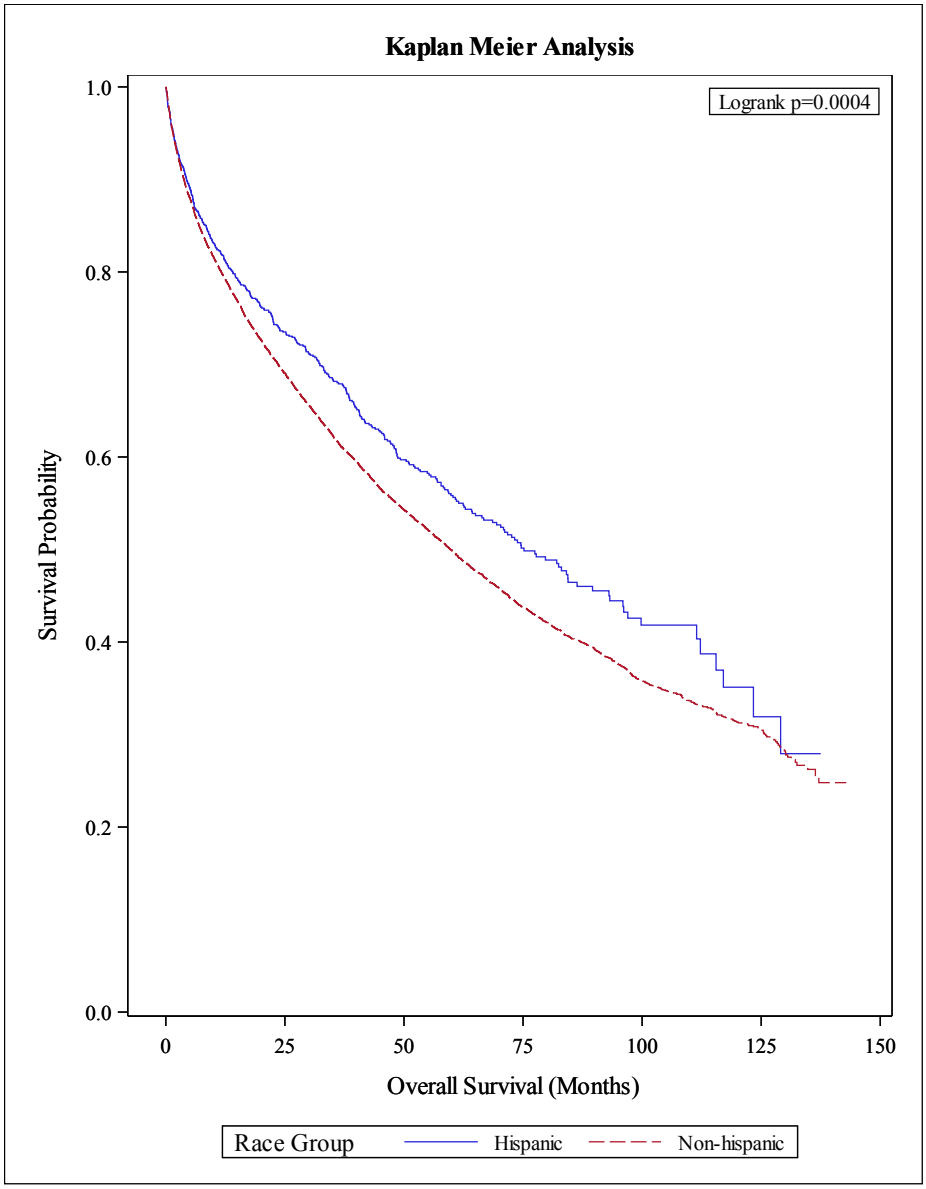


Figure 2 KM curve by race group



Race Group	No. of Subject	Event	Censored	Median Survival (95% CI)	12 Mo Survival	60 Mo Survival
Hispanic	1096	437 (40%)	659 (60%)	75.2 (65, 89.6)	81.8% (79.3%, 84.0%)	55.8% (52.2%, 59.3%
Non-hispanic	15693	7614 (49%)	8079 (51%)	59.7 (58, 61.6)	79.7% (79.0%, 80.3%)	49.9% (49.0%, 50.8%

8. Appendix

```
%MACRO WEIGHT_STD(DATASET=, TRT=, CLIST=, NLIST=, BY=, DOC=T, CRITERIA=0.1,
OUTPATH=,
     FNAME =, ORIENTATION = PORTRAIT, PLOT=T, DEBUG= F);

  %local OUTVAR NUM_OUT i STATL CVAR TVAR j __MACRO_ERR WORK_SETS COLABEL m
TABLESUM
     NONP_CTEST n COL OUT_CNT NVAR CTEST VV check pair_var;
   %if &CLIST ~= %STR() %then %let cvar_cnt = %sysfunc(countw(&CLIST));
   %else %let cvar_cnt = 0;
   %if &NLIST ~= %STR() %then %let nvar_cnt = %sysfunc(countw(&NLIST));
   %else %let nvar_cnt = 0;

   %let debug = %UPCASE(&debug);
   %let doc = %UPCASE(&doc);
   %let plot = %UPCASE(&plot);
   %let clist = %UPCASE(&CLIST);

   %let __Macro_Err= 0;

   /* Make sure that outcome variables are also not listed in CLIST */
     %do j = 1 %to &cvar_cnt;
         %if  %SCAN(&CLIST, &j) = &TRT %then %do;
             %put ERROR: Outcome &TRT cannot appear in CLIST .;
             %let __Macro_Err=1;
          %end;
     %end;
   /* Make sure that each categorical variable has at least two non-missing
values */
      %do i = 1 %to &cvar_cnt;
       PROC SQL noprint;
          select count(distinct %SCAN(&CLIST, &i)) into :check
          from &dataset
          where MISSING(%SCAN(&CLIST, &i)) = 0;
        QUIT;

        %if &check <= 1 %then %do;
          %put ERROR: The variable %SCAN(&CLIST, &i) has less than two non-
missing levels.  Please remove from CLIST.;
           %let __Macro_Err=1;
        %end;
      %end;

   /* Make sure that treatment has at exact two non-missing values */
       PROC SQL noprint;
          select count(distinct &trt) into :check2
          from &dataset
          where MISSING (&trt) = 0;
        QUIT;

        %if &check2 ne 2 %then %do;
          %put ERROR: The treatment variable should have exact two non-
missing levels. Please check.;
           %let __Macro_Err=1;
        %end;
```

```
%if &__Macro_Err. %then %do;
    data _null_;
       abort 3;
    run;
%end;

/* Get list of data sets in work library to avoid deletion later */
 ODS EXCLUDE members Directory;
 ODS OUTPUT Members(nowarn)=_DataSetList;
 PROC DATASETS lib=work memtype=data;
 QUIT;

 /* If there are data sets in the work library */
 %if %sysfunc(exist(_DataSetList)) %then %do;
    PROC SQL noprint;
       select Name
       into :work_sets separated by ' '
       from _DataSetList;
    quit;
 %end;
 %else %do;
    %let work_sets =;
 %end;

 /* Save current options */
 PROC OPTSAVE out=_options;
 RUN;

 /* Format missing values consistently */
 OPTIONS MISSING = " ";

 /* Get outcome categories & N for header row */
 PROC FREQ DATA=&DATASET noprint;
    TABLE &trt/PLOTS=NONE out=_onefreq;
    WHERE MISSING(&trt) = 0;
 RUN;

 data _onefreq;
   set _onefreq;
  /* Concanate outcome name and category */
     catelabel = catt(strip(vvalue(&trt)), " N=", ROUND(COUNT,1));
   outcat = "O" || strip(vvalue(&trt));
 run;

 proc sql noprint;
       select catelabel
       into :catelabel&m separated by "*"
       from _onefreq;
 quit;


 /* Get outcome variable names as they appear in the report data set */
 PROC TRANSPOSE DATA = _onefreq out=_tran (drop=_NAME_);
    id outcat;
 RUN;
```

```
   proc contents data=_tran out=_vname
noprint;run;

   proc sort data=_vname; by varNum;run;

   proc sql noprint ;
      select name into: categories separated by "  "
      from _vname where name not in ("_LABEL_");
   quit;

 * Character Variables;

 %IF &CLIST NE  %THEN %DO;

       %do N = 1 %to &cvar_cnt;

         %let cvar = %SCAN(&CLIST, &N);


* std without covariate adjustment;
    ODS SELECT NONE;
    ODS OUTPUT  "Cross-Tabular Freq Table" = _cfreq ;
    PROC FREQ DATA=&dataset;
      TABLE &CVAR*&trt/nopercent;
    RUN;
      ODS SELECT ALL;

 *get column percentage for each level;
    data _cfreqt;
      length outv $96.;
      set _cfreq;
      outv = strip(vvalue(&trt));
        /* Concanate outcome number and category */
      outcat = "O" || outv;
        /* Calculate the frequency of each treatment*/
        measure = catt(floor(Frequency), " (", round(ColPercent, 0.01), ")");
      /* Get rid of total rows that won't be used */
      where _TYPE_ not in ('01' '00' '10');
    run;

    data _cfreqt;
       set _cfreqt;
       where substr(outv, 1,1) not in ( " " ".");
    run;
    proc sort data=_cfreqt;by &CVAR;run;

      proc transpose data=_cfreqt out=_cfreqtt;
       var measure;
       by &CVAR;
       id outcat;
    run;

*get the proportion for two treatment in each level;
    ods select none; ods output Summary = _c_sd;
      proc means data=_cfreqt min max;
            var ColPercent;
```

```
            by &CVAR;
    run;
    ods select all;


      data _csdiff;
       set _c_sd;
         std_diff = abs(ColPercent_Min -
ColPercent_Max)/sqrt(0.5*(ColPercent_Min*(100-
ColPercent_Min)+ColPercent_Max*(100-ColPercent_Max)));
        run;

      data _cfreqtt;
        merge _cfreqtt _csdiff(keep=&CVAR std_diff);
        by &cvar;
      run;

    PROC CONTENTS DATA = _cfreqtt (drop=_NAME_  &CVAR std_diff) out=_cont
noprint;RUN;

    PROC SQL noprint;
      select name into :tvar separated by ' '
      from _cont;
    QUIT;

    data _cfreq&N;
       length covariate $256. statistic $256.;
       set _cfreqtt;
       Covariate= label(&CVAR);
       Level =  strip(vvalue(&CVAR));
         statistic= "N (Col%)";
       keep Covariate Level &tvar statistic std_diff;
      run;


     data _cfreq&N;
        set _cfreq&N;
        if substr(level, 1,1) not in ( " " ".");
        /* Order variable - keep original order */
        _order = &N;
     run;

**weighted standarized difference;

**how many distinct level in catagorical data;
ods select none;
ods output nlevels=_Level OneWayFreqs=_freq;
proc freq data=&DATASET nlevels order=formatted;
   tables &cvar/nocol norow nopercent;
run;
ods select all;


proc sql noprint;
 select distinct &cvar into: leveln separated by '' from _freq;
 select NLevels into: lev  SEPARATED by '' from _Level;      ***save number
of level to N and do dum1-dumN and if xxx="name" then dumn=1;
```

```
run;

%let mlevel= %eval(&lev-1);
***GET LEVEL NAME OF COVARIATE;
DATA _C_NAME;
 SET _FREQ;
 KEEP &CVAR;
RUN;


**MULTINOMIAL ;
      data _test;set &dataset;run;

    ods select none;
      proc datasets lib=work memtype=data;
            modify _test;
            attrib  &trt format=;
      run;
      ods select all;

    PROC FREQ DATA=_test noprint;
      TABLE &trt/PLOTS=NONE out=_trtfreq;
    RUN;

      proc sql noprint;
            select count (*) into: lastobs from _trtfreq;
            select distinct &trt into: trtnum1
                  from _trtfreq(obs=1);
            select distinct &trt into: trtnum2
                  from _trtfreq(firstobs=&lastobs) ;
      quit;

   data _newdat;
     set _test;
      if &trt=&trtnum1 then trt_1=0; else trt_1=1;
   run;

    ODS SELECT NONE;
      PROC LOGISTIC DATA =_newdat; *RORDER=DATA;
            CLASS &TRT/PARAM=GLM;
            MODEL &CVAR = TRT_1 PS TRT_1*PS /LINK =GLOGIT;
            output out = _prop predicted = ps predprobs=(individual)
xbeta=logit_ps;
        ODS OUTPUT  ParameterEstimates=_PEST;
      RUN;
      ODS SELECT ALL;

**Grab the coefficients of regression;
      DATA _P1_EST _P2_EST _P3_EST _P4_EST;
        SET _PEST;
            IF DF=0 THEN DELETE;
            IF VARIABLE="Intercept" THEN OUTPUT _P1_EST;
            IF VARIABLE="trt_1" THEN OUTPUT _P2_EST;
            IF VARIABLE="ps" THEN OUTPUT _P3_EST;
            IF VARIABLE="trt_1*ps" THEN OUTPUT _P4_EST;
      RUN;
      PROC SORT DATA=_PROP OUT=_PROP_S; BY PUF_CASE_ID;RUN;
```

```
*The other treatment for each patients;
      DATA _COVA;
       SET _PROP_S;
            INT=1;
            PS_=PS;
            IF trt_1=0 THEN ETRT=1;
            IF trt_1=1 THEN ETRT=0;
            EPS_TRT=PS_*ETRT;
            /*KEEP PUF_CASE_ID INT PS_ ETRT EPS_TRT;*/
            BY PUF_CASE_ID;
            IF FIRST.PUF_CASE_ID THEN OUTPUT _COVA;
    RUN;

      DATA _ID;
        SET _COVA;
        KEEP PUF_CASE_ID trt_1;
    RUN;

*Get the estimate treatment probability;
      PROC IML;

      USE _P1_EST;
      READ ALL VAR{ESTIMATE} INTO X1;
      CLOSE _P1_EST;

      USE _P2_EST;
      READ ALL VAR{ESTIMATE}INTO X2;
      CLOSE _P2_EST;

      USE _P3_EST;
      READ ALL VAR{ESTIMATE} INTO X3;
      CLOSE _P3_EST;

      USE _P4_EST;
      READ ALL VAR{ESTIMATE}INTO X4;
      CLOSE _P4_EST;

    X=X1||X2||X3||X4;

      USE _COVA;
      READ ALL VAR{INT}INTO Y1;
      READ ALL VAR{PS_}INTO Y3;
      READ ALL VAR{ETRT} INTO Y2;
      READ ALL VAR{EPS_TRT} INTO Y4;
      CLOSE _COVA;
      Y=Y1||Y2||Y3||Y4;
      TY=Y`;
      PROB=(X*TY)`;

      CREATE _MYDATA FROM PROB;
      APPEND FROM PROB;
      CLOSE _MYDATA;

      QUIT;
**the column of mydata represents log(probability of level1 / baseline);
```

```
***CORRESPOND EACH PROBABILITY TO ITS OWN
CASE_ID;***ESTIMATE PROBABILITY OF THREE LEVEL;
      DATA _ID_PROB;
            MERGE _ID _MYDATA;
            ARRAY COL[*] COL1-COL&mlevel; ***;
            ARRAY  a(&mlevel) P1-P&mlevel; ***;
                  DO i=1 TO &mlevel;
                  a(i)=EXP(COL[i]);
                  END;
            P&lev=1;
            TOTAL=SUM(OF P1-P&lev);

            ARRAY P[*] P1-P&lev;
            ARRAY b(&lev) PRO1-PRO&lev;
                   DO j=1 TO &lev;
                        b(j)=P[j]/Total;
                   END;

            DROP i j;
            KEEP PUF_CASE_ID TRT_1 PRO1-PRO&LEV;
      run;


***ESTIMATE PROBABILITY OF THREE LEVEL;
      DATA _TRS;
       SET _PROP;
       KEEP PUF_CASE_ID ps2 _LEVEL_ ;
      RUN;

      PROC SORT DATA=_TRS; BY PUF_CASE_ID;RUN;

      PROC TRANSPOSE DATA=_TRS OUT=_TTRS(DROP=_NAME_ _LABEL_);
            VAR ps2;
            BY PUF_CASE_ID;
      RUN;

      DATA _STDDIFF;
       MERGE _TTRS _ID_PROB;
       BY PUF_CASE_ID;
        ARRAY COL[*] COL1-COL&LEV;
        ARRAY PRO[*] PRO1-PRO&LEV;
        ARRAY c(&LEV) TRT1-TRT&LEV;
        ARRAY d(&LEV) CON1-CON&LEV;
         IF TRT_1=0 THEN DO;
           DO i=1 TO &LEV;
           d(i)=COL(i);
           c(i)=PRO(i);
             END;
         END;
         IF TRT_1=1 THEN DO;
           DO i=1 TO &LEV;
           d(i)=PRO(i);
           c(i)=COL(i);
             END;
         END;
       ARRAY TRT[*] TRT1-TRT&LEV;
       ARRAY CON[*] CON1-CON&LEV;
```

```
       ARRAY f(&LEV) STDDIFF1-STDDIFF&LEV;
         DO j=1 TO &LEV;
         f(j)=ABS( (TRT(j)-CON(j)) / SQRT( (TRT(j)*(1-TRT(j)) +CON(j)*(1-
CON(j))) / 2 ) );
         END;
       DROP i j;
       IF &LEV=2 THEN STDDIFF2=STDDIFF1;
      RUN;

      PROC MEANS DATA = _STDDIFF MEAN STD NOPRINT;
            VAR STDDIFF1-STDDIFF&lev;
            OUTPUT OUT=_AVG;
      RUN;


      DATA _AVG1;
       SET _AVG;
       KEEP _STAT_ STDDIFF1-STDDIFF&LEV;
      RUN;

      PROC TRANSPOSE DATA=_AVG1 OUT=_TAVG(KEEP=_NAME_ MEAN);
             VAR STDDIFF1-STDDIFF&LEV;
             ID  _STAT_;
      RUN;

      DATA _FINAL&N;
       LENGTH COVARIATE $20.;
       MERGE _C_NAME _TAVG;
            COVARIATE=LABEL(&CVAR);
            _ORDER=&N;
            RENAME MEAN=WEIGHTED_STD;
            LEVEL=strip(vvalue(&CVAR));
            KEEP COVARIATE LEVEL MEAN _ORDER;
      RUN;
%END;

    DATA _cfreq_all;
      SET _cfreq1-_cfreq&cvar_cnt;
      _order2 = _N_;
    RUN;
    /* Put back into original order */
      PROC SORT DATA = _cfreq_all;by  _order _order2;RUN;

    DATA _FINAL_all;
       SET _FINAL1-_FINAL&CVAR_CNT;
       _order2 = _N_;
     RUN;
       /* Put back into original order*/
       PROC SORT DATA = _FINAL_all;by  _order _order2;RUN;

       DATA _CVAR_STDALL;
     MERGE _CFREQ_ALL _FINAL_ALL; BY _ORDER _ORDER2;
     RUN;

%END;

*numeric;
```

```
%IF &NLIST NE  %THEN %DO;
       %do N = 1 %to &nvar_cnt;

         %let nvar = %SCAN(&NLIST, &N);

*standarized difference without covariate adjustment;

  PROC MEANS DATA=&DATASET noprint;
        var &NVAR;
          class &trt;
            output out=_summary (DROP= _FREQ_)mean=mean std=std ;
    RUN;

    data _n_sum;
      length covariate $100. measure $100. outv $100.;
        set _summary;
        covariate = label(mean);
        measure = catt(round(mean,0.01), " (",round(std,.01), ")" );
      /* Concanate outcome number and category */
            outv= strip(vvalue(&trt));
        outcat = "O" || outv;
            where _TYPE_ = 1;
    run;

    data _n_sum;
      set _n_sum;
        where substr(outv, 1,1) not in ( " " ".");
    run;

    proc transpose data=_n_sum out=_s_sumt;
       var measure ;
       id outcat;
         copy covariate;
    run;

   *mean and std^2 of each trt;
      proc sort data=_summary; by _type_ mean;run;
      data _numtest;
        set _summary;
        by _type_;
        std2= std*std;
        if first._type_ then output;
        if last._type_ then output;
    run;

 *standarized difference formula: x1bar-x2bar/sqrt(mean of s1^2 and s2^2);
    ods select none;
    ods output summary = _sd_;
      proc means data=_numtest range mean;
        var mean std2;
      run;
    ods select all;

   data _sdiff&N; length statistic $256.;
      merge _s_sumt _sd_;
        std_diff = mean_Range/sqrt(std2_Mean);
```

```
        _order=&N;
      statistic= "Mean (Std)";
      if _NAME_="" then delete;
      keep covariate statistic &tvar std_diff _order;
    run;

**weighted standardized diff;
    ODS SELECT NONE;
    PROC GENMOD DATA =_newdat;
           CLASS TRT_1;
           MODEL &NVAR = TRT_1 PS TRT_1*PS / DIST =NOR LINK = ID TYPE3;
           LSMEANS TRT_1 / DIFF;
    ODS OUTPUT PARAMETERESTIMATES = _NTEST; **coefficients estimator;
    ODS OUTPUT MODELFIT = _NTEST1;          **Criteria For Assessing
Goodness Of Fit;
    RUN;
    ODS SELECT ALL;

* Save parameter estimate into different datasets;

    DATA _TRT_ESTc (KEEP =  TRT0_EST) _PS_ESTc (KEEP = PS_EST)
     _TRTPS_ESTc (KEEP =  TRT0PS_EST) _INTRCPT_ESTc(KEEP = INTRCPT_EST);

    SET _NTEST;

    IF PARAMETER = "trt_1" AND DF = 1 THEN DO;
          TRT0_EST = ESTIMATE;
          OUTPUT _TRT_ESTc;
          END;

          IF PARAMETER = "ps" THEN DO;
          PS_EST = ESTIMATE;
          OUTPUT _PS_ESTc;
          END;

        IF PARAMETER = "ps*trt_1" AND DF=1 THEN DO;
          TRT0PS_EST = ESTIMATE;
          OUTPUT _TRTPS_ESTc;
          END;

        IF PARAMETER = "Intercept" THEN DO;
          INTRCPT_EST = ESTIMATE;
          OUTPUT _INTRCPT_ESTc;
          END;
    RUN;

**std error;
    DATA _NTEST1;

     SET _NTEST1;

           IF CRITERION = "Deviance";
         SIGMA = SQRT(ValueDF);

     KEEP SIGMA;
    RUN;
```

```
      DATA _N_EST;

       MERGE _NTEST1 _TRT_ESTc _PS_ESTc _TRTPS_ESTc _INTRCPT_ESTc;

       KEEP TRT0_EST PS_EST TRT0PS_EST INTRCPT_EST SIGMA;
      RUN;
 * Merge parameter estimates with analysis data to allow computation
 of predicted values for each patient. ;

* For each observation, compute the predicted value assuming each treatment
group *;
      DATA _NALL;

       MERGE &DATASET _N_EST;
       RETAIN TRT0_EST_N PS_EST_N TRT0PS_EST_N INTRCPT_EST_N SIGMA_N;
             IF _N_ = 1 THEN DO;
             TRT0_EST_N = TRT0_EST;
             PS_EST_N = PS_EST;
             TRT0PS_EST_N = TRT0PS_EST;
             INTRCPT_EST_N = INTRCPT_EST;
             SIGMA_N = SIGMA;
       END;

       PRED0 = INTRCPT_EST_N + TRT0_EST_N + PS_EST_N*ps +TRT0PS_EST_N*ps;
       PRED1 = INTRCPT_EST_N + PS_EST_N*PS;

* Compute the standardized difference for binary covariates *;
       TRTDIFF = TRT0_EST_N + TRT0PS_EST_N*PS;
       STDDIFF = ABS(TRT0_EST_N + TRT0PS_EST_N*PS) / SIGMA_N;
      RUN;

      PROC MEANS DATA = _NAll MEAN NOPRINT;
             VAR STDDIFF;
             OUTPUT OUT=_NAVG;
      RUN;

      DATA _NAVG;
       SET _NAVG;
       KEEP  _STAT_ STDDIFF;
      RUN;

      PROC TRANSPOSE DATA=_NAVG OUT=_NTAVG(KEEP=_NAME_ MEAN);
            VAR STDDIFF;
            ID  _STAT_;
      RUN;

      DATA _NFINAL&N;
            SET _NTAVG;
            _order=&N;
      RUN;
%END;

*merge all st difference without covariate adjustment;

      DATA _SDIFF_ALL;
        set _sdiff1-_sdiff&NVAR_CNT;
         /* Order variable - keep original order */
```

```
        _ORDER2 = _N_;
      RUN;

*Put back into original order;
      PROC SORT DATA = _SDIFF_ALL;
            BY  _ORDER _ORDER2;
      RUN;

*merge all weighted st difference;
      DATA _NFINAL_all;
        SET _NFINAL1-_NFINAL&NVAR_CNT;
         _ORDER2 = _N_;
         RENAME MEAN=WEIGHTED_STD;
      RUN;
      PROC SORT DATA= _NFINAL_all; BY _ORDER _ORDER2;RUN;


      DATA _NUM_STDALL;
        MERGE _SDIFF_ALL _NFINAL_all; BY _ORDER _ORDER2;
        DROP _NAME_;
      RUN;

%END;
   DATA _report;
      set _CVAR_STDALL _NUM_STDALL;
   RUN;

   DATA _NULL_;
    SET &DATASET (OBS=1);
      CALL SYMPUT("outLabel", put(label(&trt),$256.));
   RUN;

*generate the comparision plot;
   DATA _PLOT1;
      LENGTH Y_LB $256.;
      SET _REPORT;
      ID+1;
      IND=1;
      IF statistic= "Mean (Std)" THEN Y_LB=COVARIATE;
      ELSE Y_LB= catt( COVARIATE, "-", LEVEL);
      RENAME STD_DIFF= X_LB;
      KEEP Y_LB ID IND std_diff;
   RUN;

  DATA _PLOT2;
      LENGTH Y_LB $256.;
      SET _REPORT;
      ID+1;
      IND=2;
      IF statistic= "Mean (Std)" THEN Y_LB=COVARIATE;
      ELSE Y_LB= catt( COVARIATE, "-", LEVEL);
      RENAME WEIGHTED_STD= X_LB;
      KEEP Y_LB ID IND WEIGHTED_STD;
  RUN;

  PROC FORMAT;
      VALUE X 1="Before Ps adjustment"  2="Aefore Ps adjustment";
```

```
  run;

  DATA _PLOT;
      SET _PLOT1 _PLOT2;
      FORMAT IND X.;
      LABEL X_LB="Absolute standardized difference" Y_LB="COVARIATE"
            IND="Results";
  RUN;



   *---- table template -----;

   ODS PATH WORK.TEMPLAT(UPDATE) SASUSR.TEMPLAT(UPDATE)
SASHELP.TMPLMST(READ);

   PROC TEMPLATE;
   DEFINE STYLE STYLES.TABLES;
   NOTES "MY TABLE STYLE";
   PARENT=STYLES.MINIMAL;

     STYLE SYSTEMTITLE /FONT_SIZE = 12pt     FONT_FACE = "TIMES NEW ROMAN";

     STYLE HEADER /
           FONT_FACE = "TIMES NEW ROMAN"
            CELLPADDING=8
            JUST=C
            VJUST=C
            FONT_SIZE = 10pt
           FONT_WEIGHT = BOLD;

     STYLE TABLE /
            FRAME=HSIDES            /* outside borders: void, box,
above/below, vsides/hsides, lhs/rhs */
           RULES=GROUP              /* internal borders: none, all, cols,
rows, groups */
           CELLPADDING=6            /* the space between table cell contents
and the cell border */
            CELLSPACING=6           /* the space between table cells, allows
background to show */
            JUST=C
            FONT_SIZE = 10pt
           BORDERWIDTH = 0.5pt;  /* the width of the borders and rules */

     STYLE DATAEMPHASIS /
           FONT_FACE = "TIMES NEW ROMAN"
           FONT_SIZE = 10pt
           FONT_WEIGHT = BOLD;

     STYLE DATA /
           FONT_FACE = "TIMES NEW ROMAN"
           FONT_SIZE = 10pt;

     STYLE SYSTEMFOOTER /FONT_SIZE = 9pt FONT_FACE = "TIMES NEW ROMAN"
JUST=C;
   END;
```

```
   *------- build the table -----;

   OPTIONS ORIENTATION=&ORIENTATION MISSING = "" NODATE;
   %if &doc = T %then %do;
      ODS rtf style= tables FILE= "&OUTPATH.&FNAME &SYSDATE..doc";
   %end;

   PROC REPORT DATA=_report HEADLINE CENTER STYLE(REPORT)={JUST=CENTER}
SPLIT='~' nowd
     SPANROWS LS=256;
      COLUMNS Covariate level  statistic
             ("&outLabel" '_____________________________'( &categories))
std_diff WEIGHTED_STD;

       DEFINE Covariate/order order=data "Covariate" STYLE(COLUMN) = {JUST =
L};
       DEFINE level/DISPLAY "Level" STYLE(COLUMN) = {JUST = L};
         DEFINE statistic/DISPLAY "Statistics" STYLE(COLUMN) = {JUST = C};
         DEFINE std_diff/DISPLAY "Standardized Difference" STYLE(COLUMN) =
{JUST = C} format=6.4;
         DEFINE WEIGHTED_STD/DISPLAY "Weighted Standardized Difference"
STYLE(COLUMN) = {JUST = C} format=6.4;
        %LET I = 1;
        %DO %UNTIL (%SCAN(&categories, &I) =   );
             %LET col = %SCAN(&categories, &I);
              %LET colabel = %SCAN(%BQUOTE(&catelabel), &I, *);
              DEFINE &col/DISPLAY  "&colabel" STYLE(COLUMN) = {JUST = C}  ;
              %LET I = %EVAL(&I+1);
        %END;

         COMPUTE std_diff;
              IF std_diff > &criteria THEN CALL DEFINE("std_diff", "STYLE",
"STYLE=[FONT_WEIGHT=BOLD]");
            ENDCOMP;

         COMPUTE WEIGHTED_STD;
              IF WEIGHTED_STD > &criteria THEN CALL DEFINE("WEIGHTED_STD",
"STYLE", "STYLE=[FONT_WEIGHT=BOLD]");
            ENDCOMP;


      compute after covariate; line ''; endcomp;
      run;




   %if &doc = T %then %do;
      ODS RTF CLOSE;
   %end;


   %if &plot = T %then %do;
     ods rtf file="&OUTPATH.plot &FNAME &SYSDATE..doc";
```

```
      ODS GRAPHICS/WIDTH=700PX HEIGHT=900PX;
   %end
       Title "Comparision between standarized difference and weighted
standarized difference after covariate adjustment";
       PROC SGPLOT DATA=_PLOT;
           SCATTER X=X_LB Y=Y_LB/GROUP=IND;
           REFLINE &CRITERIA/AXIS=X LINEATTRS=(PATTERN=2);
           YAXIS VALUEATTRS=(SIZE=8);
       RUN;
       Title;

   %if &plot = T %then %do;
      ODS RTF CLOSE;
   %end;




 /* Reload original options that were in use before running the macro */
   PROC OPTLOAD data=_options;
   RUN;

   /* Only delete files if not in debug mode */
   %if &debug ~= T %then %do;

      /* If there are work data sets that should not be deleted */
      %if %sysevalf(%superq(work_sets)~=,boolean) %then %do;
         /* DELETE ALL TEMPORARY DATASETS that were created */
         proc datasets lib=work memtype=data noprint;
            save &work_sets;
         quit;
      %end;
      %else %do;
         proc datasets lib=work kill memtype=data noprint;
         quit;
      %end;
   %end;



%MEND;
```