www.chrp.org/love

Explaining Post-Propensity Score Covariate Balance

Thomas E. Love, Ph. D.
Case Western Reserve University
Center for Health Care Research & Policy

ENAR – April 1, 2003

Center for Health Care Research & Policy

CWRU

Using Propensity Scores to Balance Covariates

Propensity Score= Pr (exposure|covariates)

- Goal: Compare exposed and unexposed groups with similar baseline covariates X.
- Uses of Propensity Scores
  - Matching
  - Stratification (Subclassification)
  - Regression Adjustment
  - Weighting

2

Building the Propensity Score Model

- Include all covariates that are important when selecting treatments, especially if they are also related to outcome.
- No point in parsimony here – we want to sop up as much "signal" as possible.
- Not a prediction model – quality of model should be solely judged on covariate balance after further adjustment using the PS.

3

Aspirin Use and Mortality

- 6174 adults having stress echocardiography for evaluation of known or suspected coronary disease.
- Treatment: 2310 (37%) were taking aspirin.
- Outcome: all-cause mortality (median follow-up 3y)
  - 4.5% of both aspirin and non-aspirin patients died
  - Unadjusted Hazard Ratio: 1.08 (0.85, 1.39)
- PS Model – logistic regression using 34 covariates
  - Demographics, Clinical Hx, Medication Use
  - Cardiovascular assessment and Exercise capacity

Gum PA et al. (2001) *JAMA*, 1187-1194.

4

Baseline Characteristics By Aspirin Use (before matching)

Variable	Aspirin (n = 2310)	No Aspirin (n = 3864)	P value
Men	77.0	56.1	< .001
Age, years	62 (11)	56 (12)	< .001
Resting systolic BP, mm Hg	141 (21)	138 (20)	< .001
Peak exercise cap., men (METs)	8.6 (2.4)	9.1 (2.6)	< .001
Clinical history: diabetes	16.8	11.2	< .001
hypertension	53.0	40.6	< .001
congestive heart failure	5.5	4.6	.12
Beta-blocker use	35.1	14.2	< .001
ACE inhibitor	13.0	11.4	< .001

- Aspirin user covariates indicate higher mortality risk.

5

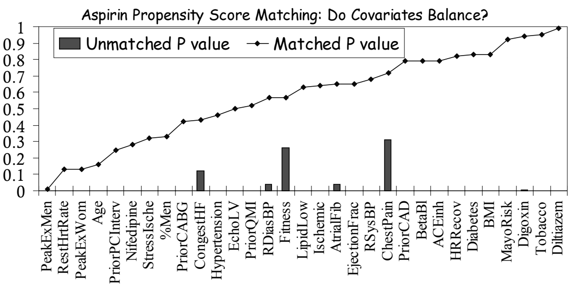
Baseline Characteristics By Aspirin Use [%] (after matching)

Variable	Aspirin (n = 1351)	No Aspirin (n = 1351)	P value
Men	70.4	72.1	.33
Age, years	60 (11)	61 (11)	.16
Resting systolic BP, mm Hg	141 (21)	141 (21)	.68
Peak exercise cap., men (METs)	8.7 (2.5)	8.3 (2.5)	**.01**
Clinical history: diabetes	15.0	15.3	.83
hypertension	50.3	51.7	.46
congestive heart failure	5.8	6.6	.43
Medication use: Beta-blocker	26.1	26.5	.79
ACE inhibitor	15.5	15.8	.79

- Matched users & non-users: similar covariate means.

6

Are The Covariates Balanced? "P values Plot"



7

Standardized Differences Measure Imbalance Well

- Appropriate summaries of Covariate Balance for both Continuous and Categorical Variables

$$d = \frac{100\left(\bar{x}_{\text{Exposed}} - \bar{x}_{\text{Control}}\right)}{\sqrt{\frac{s^2_{\text{Exposed}} + s^2_{\text{Control}}}{2}}} \quad \text{for continuous variables}$$

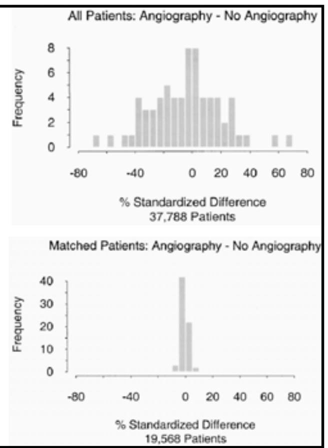
$$d = \frac{100\left(p_{\text{Exposed}} - p_{\text{Control}}\right)}{\sqrt{\frac{p_E(1-p_E) + p_C(1-p_C)}{2}}} \quad \text{for binary variables}$$

8

<u>Before</u> Matching

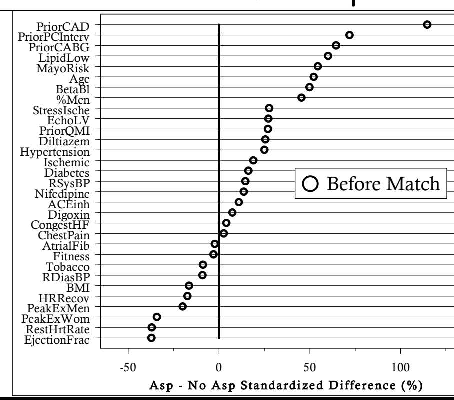
Standardized Differences (%) in Covariate Means

<u>After</u> Matching



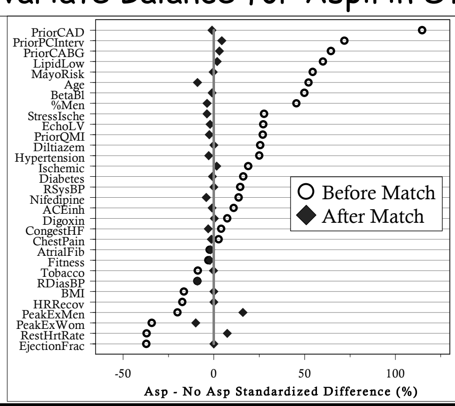
Normand et al. (2001) p. 395

Covariate Balance for Aspirin Study



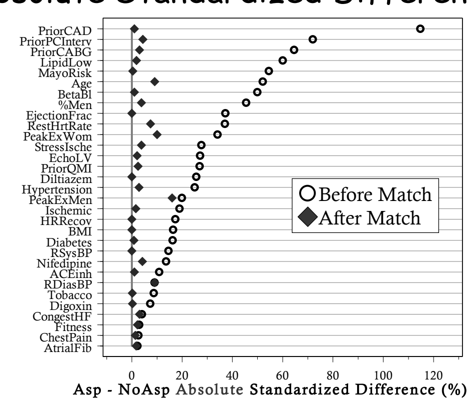
10

Covariate Balance for Aspirin Study



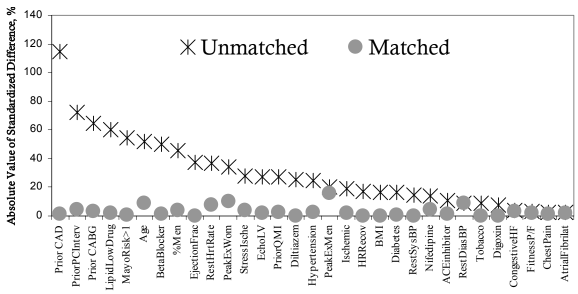
11

Absolute Standardized Differences



12

Covariate Balance for Aspirin Study



Absolute Value of Standardized Difference plotted for 31 covariates (in Excel).

13

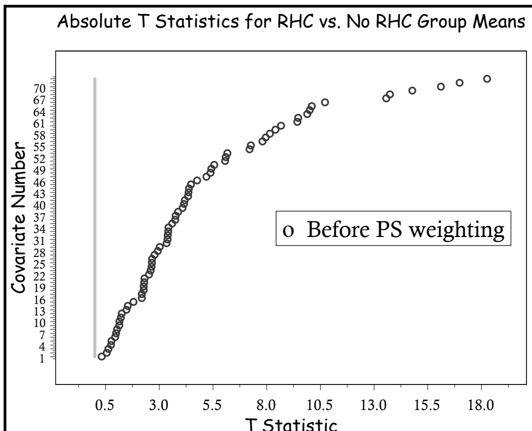
RHC Covariate Balance Before / After Propensity Score Weighting

	Before Weighting			After Weighting		
Variable	No RHC	RHC	t	No RHC	RHC	t
Age	61.76	60.74	-2.28	61.25	61.15	-0.19
Sex	0.46	0.41	-3.42	0.44	0.43	-0.85
Edu	11.56	11.85	3.35	11.68	11.71	0.39
COPD	0.11	0.02	-13.6	0.07	0.06	-1.10
Album	3.16	2.97	-8.15	3.08	3.15	0.69
Potass	4.07	4.04	-0.99	4.15	3.97	-4.10

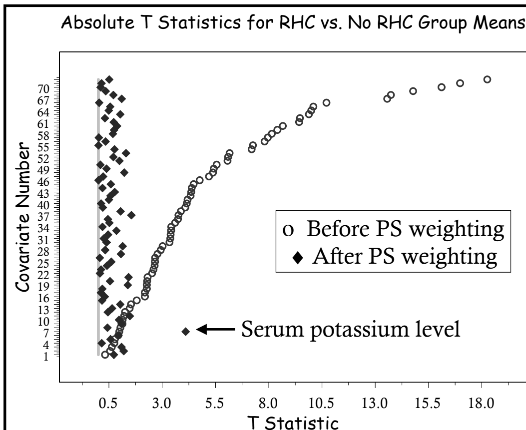
- 2184 RHC patients and 3551 controls
- 57 of 72 covariates (all with |t| > 2.0) used in PS

See Connors et al. (1996) and Hirano & Imbens (2001)

14

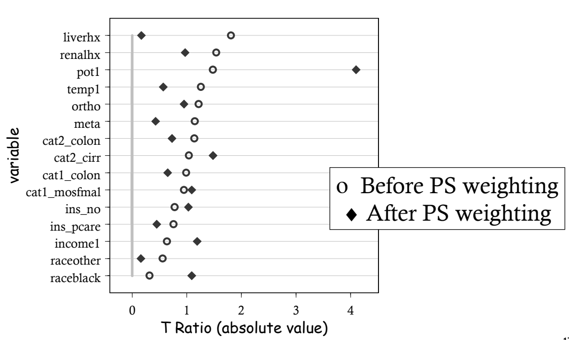


15



16

Balance on Covariates Not in The New Propensity Model



17

Choices To Be Made

- How should I summarize balance within a covariate?
  - What if I have both continuous and categorical covariates?
  - Standardized differences
  - Significance test results
- How can I best present the summarized results across covariates?
- What are the key messages to get across?

18

International Biometric Society - Eastern North American Region - Tampa, Florida. April 1, 2003 at 3:45 PM

Explaining Post-Propensity Score Covariate Balance

Thomas E. Love, Ph.D.
Director, Biostatistics and Evaluation Unit, Center for Health Care Research & Policy
Case Western Reserve University – MetroHealth Medical Center

Abstract

In observational studies of treatment effects, we regularly adjust for the possible effects of selection bias through methods such as propensity scores. In presenting the results of these studies, it is crucial to convince skeptics that, after adjustment, baseline covariates are balanced (comparable) between treated and untreated groups. Publications in the health care literature typically display this balance with massive tables of summary statistics. We develop more efficient and effective means of communication through graphical illustrations of the impact of propensity score matching, stratification or adjustment. The results should aid researchers to present the results of these sometimes complicated analyses more effectively to clinicians and other consumers.

A Few Useful References

- Special Issue of *Health Services Research and Outcomes Methodology* on Causal Inference (December 2001) Volume 2, Issues 3&4. Several relevant articles within, including the article by Hirano and Imbens that is cited in the presentation. PDF files available FREE online at http://www.kluweronline.com/issn/1387-3741/current
- Connors AF, Speroff T, Dawson NV, Thomas C, et al. (1996) The effectiveness of right heart catheterization in the initial care of critically ill patients (with Editorial), *Journal of the American Medical Association* **276**: 889-897, 915-918
- D'Agostino RB Jr. (1998) Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group, *Statistics in Medicine* **17**: 2265-2281
- Gum PA, Thamilarasan M, et al. (2001) Aspirin Use and All-Cause Mortality Among Patients Being Evaluated for Known or Suspected Coronary Artery Disease - A Propensity Analysis, *Journal of the American Medical Association* **286**: 1187-1194
- Joffe MM and Rosenbaum PR (1999) Propensity Scores, *American Journal of Epidemiology* **150**: 327-333
- Normand, S-L Landrum, MB et al. (2001) Validating recommendations for coronary angiography following acute myocardial infarction in the elderly: A matched analysis using propensity scores, *Journal of Clinical Epidemiology* **54**: 387-398
- Rosenbaum PR (2002) *Observational Studies*, 2nd Edition. New York: Springer-Verlag.
- Rosenbaum PR and Rubin DB (1984) Reducing bias in observational studies using subclassification on the propensity score, *Journal of the American Statistical Association* **79**: 516-524
- Rubin DB (1997) Estimating Causal Effects from Large Data Sets Using Propensity Scores, *Annals of Internal Medicine* **127**: 757-763.
- I also strongly recommend the work of William Cleveland, Edward Tufte and Howard Wainer on visualizing data and informative graphics.