

October 11, 2016

Considerations in the Use of Propensity Scores in Observational Studies

Lawrence Rasouliyan, Estel Plana, Jaume Aguado
RTI Health Solutions, Barcelona, Spain

PhUSE 2016 / Real World Evidence Stream

The power of **knowledge.**
The value of **understanding.**

Background

- Randomized Controlled Trials = Gold standard for treatment comparison
- Observational studies are increasingly being used to estimate the effects of treatments, exposures, and interventions on outcomes.
- Treated and untreated patients often have systematic differences in their distributions of underlying characteristics.
- Traditionally, multivariable models have been used.
- Propensity scores can be used to minimize the effects of bias in the estimate of the treatment effect.
 - Serves as a “balancing” score for measured confounders
 - Outcomes are compared taking into account this balancing to reduce potential bias of confounders

Propensity Score: Definition

- Probability of receiving Treatment A (versus Treatment B) given the patient's underlying characteristics.
- Usually determined by logistic regression:

Propensity Score: Definition

- Probability of receiving Treatment A (versus Treatment B) given the patient's underlying characteristics.
- Usually determined by logistic regression:

$$\ln\left(\frac{P_A}{1 - P_A}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots$$

Propensity Score: Definition

- Probability of receiving Treatment A (versus Treatment B) given the patient's underlying characteristics.
- Usually determined by logistic regression:

$$\ln\left(\frac{P_A}{1 - P_A}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots$$

- Solving for P_A , we get the propensity score:

$$P_A = \frac{\exp[\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots]}{1 + \exp[\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots]}$$

Propensity Score: Definition

- Probability of receiving Treatment A (versus Treatment B) given the patient's underlying characteristics.
- Usually determined by logistic regression:

$$\ln\left(\frac{P_A}{1 - P_A}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots$$

- Solving for P_A , we get the propensity score:

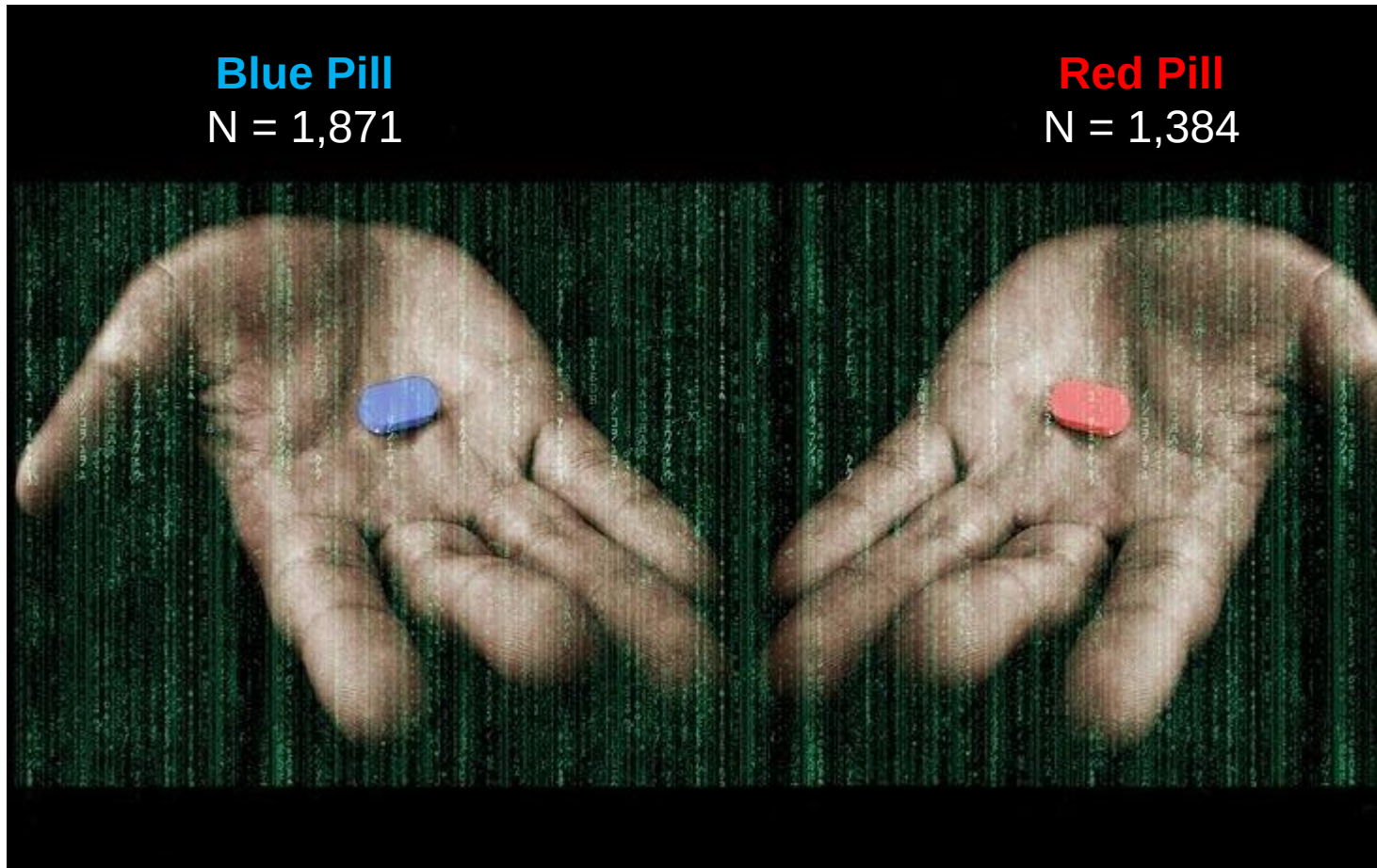
$$P_A = \frac{\exp[\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots]}{1 + \exp[\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots]}$$

- Then, we account for this propensity score to make treatment comparison.

Illustrative Example: Data Source

- Prospective 8-year observational study in oncology
- Outcomes of interest:
 - Overall survival (OS) = Time to death
 - Progression-free survival (PFS) = Time to progression or death
 - Overall response rate (ORR) = Response to treatment
- Data simulated to mimic general attributes of actual results
- Disease indication: “Cancer”
- Two possible therapies...

Two Possible Therapies



Which pill is “better” for outcomes of interest?

Patient Characteristics

Characteristic	Blue Pill N = 1871	Red Pill N = 1384	P value
Age group, n (%)			
≤ 50	59 (3.2%)	234 (16.9%)	<0.001
51 to 60	576 (30.8%)	684 (49.4%)	
61 to 70	925 (49.4%)	412 (29.8%)	
> 70	311 (16.6%)	54 (3.9%)	
Sex, n (%)			
Male	834 (44.6%)	593 (42.8%)	0.326
Female	1037 (55.4%)	791 (57.2%)	
Geographic region, n (%)			
North America	846 (45.2%)	553 (40.0%)	<0.001
Europe	705 (37.7%)	702 (50.7%)	
Asia	320 (17.1%)	129 (9.3%)	
Performance status, n (%)			
0	1060 (56.7%)	1001 (72.3%)	<0.001
1	688 (36.8%)	352 (25.4%)	
≥ 2	123 (6.6%)	31 (2.2%)	

Patient Characteristics

Characteristic	Blue Pill N = 1871	Red Pill N = 1384	P value
Age group, n (%)			
≤ 50	59 (3.2%)	234 (16.9%)	<0.001
51 to 60	576 (30.8%)	684 (49.4%)	
61 to 70	925 (49.4%)	412 (29.8%)	
> 70	311 (16.6%)	54 (3.9%)	
Sex, n (%)			
Male	834 (44.6%)	593 (42.8%)	0.326
Female	1037 (55.4%)	791 (57.2%)	
Geographic region, n (%)			
North America	846 (45.2%)	553 (40.0%)	<0.001
Europe	705 (37.7%)	702 (50.7%)	
Asia	320 (17.1%)	129 (9.3%)	
Performance status, n (%)			
0	1060 (56.7%)	1001 (72.3%)	<0.001
1	688 (36.8%)	352 (25.4%)	
≥ 2	123 (6.6%)	31 (2.2%)	

Red Pill Patients

Younger

From Europe

Better
Performance

Patient Characteristics (Continued)

Characteristic	Blue Pill N = 1871	Red Pill N = 1384	P value
Prognosis risk category, n (%)			
Low risk	297 (15.9%)	228 (16.5%)	<0.001
Medium risk	684 (36.6%)	656 (47.4%)	
High risk	890 (47.6%)	500 (36.1%)	
Disease stage, n (%)			
III	1210 (64.7%)	851 (61.5%)	0.063
IV	661 (35.3%)	533 (38.5%)	
Extranodal sites, n (%)			
≥ 2	591 (31.6%)	417 (30.1%)	0.374
< 2	1280 (68.4%)	967 (69.9%)	
LDH level, n (%)			
Elevated	498 (26.6%)	290 (21.0%)	<0.001
Normal	1373 (73.4%)	1094 (79.0%)	
Bone marrow involvement, n (%)			
Yes	1165 (62.3%)	900 (65.0%)	0.106
No	706 (37.7%)	484 (35.0%)	
Center type, n (%)			
Academic	400 (21.4%)	197 (14.2%)	<0.001
Community	1471 (78.6%)	1187 (85.8%)	

Red Pill Patients

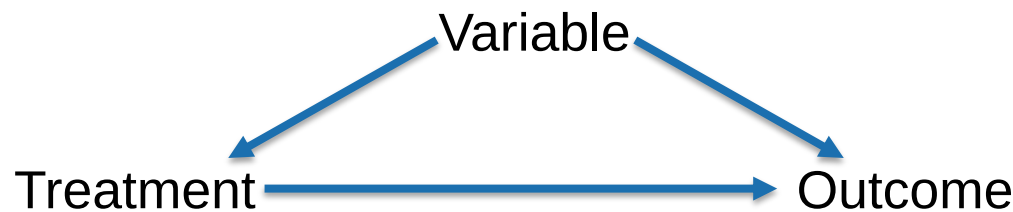
Better Prognosis Risk

Normal LDH Level

From Community Center

Specifying the Propensity Score Model

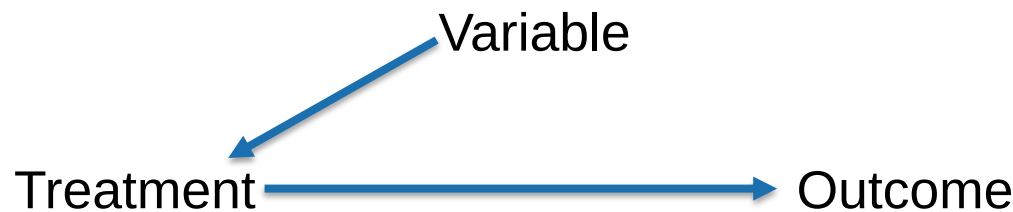
- Choosing main effects is an “art” and science
- Several factors to consider (e.g., strength of association, clinical plausibility, multicollinearity, effect modification)
- Remember the purpose: Control for confounding variables



- Simulation studies have shown:
 - Confounding variables should always be included

Specifying the Propensity Score Model

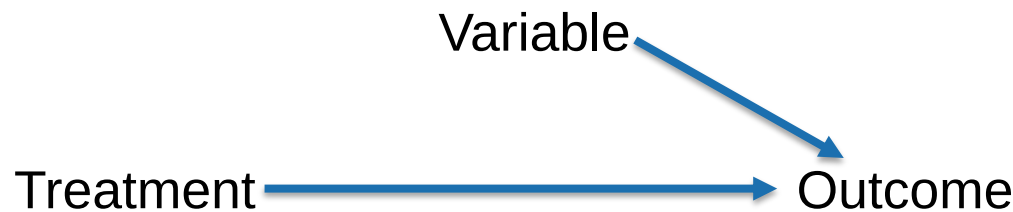
- Choosing main effects is an “art” and science
- Several factors to consider (e.g., strength of association, clinical plausibility, multicollinearity, effect modification)
- Remember the purpose: Control for confounding variables



- Simulation studies have shown:
 - Confounding variables should always be included
 - Variables associated with only the treatment decrease precision

Specifying the Propensity Score Model

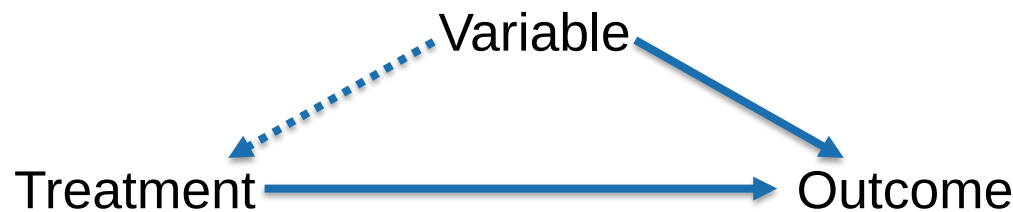
- Choosing main effects is an “art” and science
- Several factors to consider (e.g., strength of association, clinical plausibility, multicollinearity, effect modification)
- Remember the purpose: Control for confounding variables



- Simulation studies have shown:
 - Confounding variables should always be included
 - Variables associated with only the treatment decrease precision
 - Variables associated with only the outcome increase precision

Specifying the Propensity Score Model

- Choosing main effects is an “art” and science
- Several factors to consider (e.g., strength of association, clinical plausibility, multicollinearity, effect modification)
- Remember the purpose: Control for confounding variables



- Simulation studies have shown:
 - Confounding variables should always be included
 - Variables associated with only the treatment decrease precision
 - Variables associated with only the outcome increase precision
- Our strategy: **Include all variables associated with outcome**

Selection of Patient Characteristics for Model

- Look at variables one at a time (univariable models)
- Determine which variables are associated with any outcome

```
* Univariable Cox Regression Model Predicting OS;
```

```
proc phreg data=ad01;  
  class region (ref="North America");  
  model osmo*osevt(0) = region;  
run;
```

```
* Univariable Cox Regression Model Predicting PFS;
```

```
proc phreg data=ad01;  
  class region (ref="North America");  
  model pfsmo*pfsevt(0) = region;  
run;
```

```
* Univariable Logistic Regression Model Predicting ORR;
```

```
proc logistic data=ad01;  
  class region (ref="North America");  
  model orr (event = "1") = region;  
run;
```

Selection of Patient Characteristics for Model

Variable	OS		PFS		ORR	
	β	P Value	β	P Value	β	P Value
Age group						
≤ 50	Ref	Ref	Ref	Ref	Ref	Ref
51 to 60	0.291	0.141	0.038	0.728	0.411	0.001
61 to 70	0.922	<0.001	0.511	<0.001	-0.189	0.070
> 70	1.057	<0.001	0.393	0.001	-0.806	<0.001
Sex						
Male	0.049	0.536	0.002	0.970	0.021	0.734
Female	Ref	Ref	Ref	Ref	Ref	Ref
Geographic region						
North America	Ref	Ref	Ref	Ref	Ref	Ref
Europe	-0.041	0.628	-0.058	0.307	0.093	0.274
Asia	0.167	0.144	0.138	0.076	-0.298	0.004
Performance status						
0	Ref	Ref	Ref	Ref	Ref	Ref
1	0.059	0.491	0.092	0.104	0.019	0.859
≥ 2	0.679	<0.001	0.423	<0.001	-0.334	0.036
Disease stage						
III	Ref	Ref	Ref	Ref	Ref	Ref
IV	0.349	<0.001	0.335	<0.001	-0.217	0.001
Extranodal sites						
≥ 2	0.209	0.011	0.152	0.006	-0.201	0.001
< 2	Ref	Ref	Ref	Ref	Ref	Ref
LDH level						
Elevated	0.673	<0.001	0.335	<0.001	-0.417	<0.001
Normal	Ref	Ref	Ref	Ref	Ref	Ref
Bone marrow involvement						
Yes	-0.019	0.813	0.001	0.986	-0.018	0.780
No	Ref	Ref	Ref	Ref	Ref	Ref
Center type						
Academic	0.187	0.051	0.129	0.047	0.068	0.405
Community	Ref	Ref	Ref	Ref	Ref	Ref

Selection of Patient Characteristics for Model

Variable	OS		PFS		ORR	
	β	P Value	β	P Value	β	P Value
Age group						
≤ 50	Ref	Ref	Ref	Ref	Ref	Ref
51 to 60	0.291	0.141	0.038	0.728	0.411	0.001
61 to 70	0.922	<0.001	0.511	<0.001	-0.189	0.070
> 70	1.057	<0.001	0.393	0.001	-0.806	<0.001
Sex						
Male	0.049	0.536	0.002	0.970	0.021	0.734
Female	Ref	Ref	Ref	Ref	Ref	Ref
Geographic region						
North America	Ref	Ref	Ref	Ref	Ref	Ref
Europe	-0.041	0.628	-0.058	0.307	0.093	0.274
Asia	0.167	0.144	0.138	0.076	-0.298	0.004
Performance status						
0	Ref	Ref	Ref	Ref	Ref	Ref
1	0.059	0.491	0.092	0.104	0.019	0.859
≥ 2	0.679	<0.001	0.423	<0.001	-0.334	0.036
Disease stage						
III	Ref	Ref	Ref	Ref	Ref	Ref
IV	0.349	<0.001	0.335	<0.001	-0.217	0.001
Extranodal sites						
≥ 2	0.209	0.011	0.152	0.006	-0.201	0.001
< 2	Ref	Ref	Ref	Ref	Ref	Ref
LDH level						
Elevated	0.673	<0.001	0.335	<0.001	-0.417	<0.001
Normal	Ref	Ref	Ref	Ref	Ref	Ref
Bone marrow involvement						
Yes	-0.019	0.813	0.001	0.986	-0.018	0.780
No	Ref	Ref	Ref	Ref	Ref	Ref
Center type						
Academic	0.187	0.051	0.129	0.047	0.068	0.405
Community	Ref	Ref	Ref	Ref	Ref	Ref

Selected Variables

- Age group
- Geographic region
- Performance status
- Disease stage
- Extranodal sites
- LDH level

Propensity Score Model

* PS Model: Multivariable Logistic Regression Predicting Receipt of BP;

```
proc logistic data=ad01;
  class agegrp (ref="<=50") region (ref="North America")
    perfstat (ref="0") stage (ref="III")
    extranod (ref=">=2") ldhlevel (ref="Elevated") / param=ref;
  model tx (event="BP") = agegrp region perfstat stage extranod ldhlevel;
  output out=ad02 pred=propensity;
run;
```

Propensity Score Model

* PS Model: Multivariable Logistic Regression Predicting Receipt of BP;

```
proc logistic data=ad01;
  class agegrp (ref="<=50") region (ref="North America")
    perfstat (ref="0") stage (ref="III")
    extranod (ref=">=2") ldhlevel (ref="Elevated") / param=ref;
  model tx (event="BP") = agegrp region perfstat stage extranod ldhlevel;
  output out=ad02 pred=propensity;
run;
```

Model receipt
of the Blue Pill

Propensity Score Model

* PS Model: Multivariable Logistic Regression Predicting Receipt of BP;

```
proc logistic data=ad01;  
  class agegrp (ref="≤50") region (ref="North America")  
    perfstat (ref="0") stage (ref="III")  
    extranod (ref="≥2") ldhlevel (ref="Elevated") / param=ref;  
  model tx (event="BP") = agegrp region perfstat stage extranod ldhlevel;  
  output out=ad02 pred=propensity;  
run;
```

Model receipt
of the Blue Pill

As a function of the
6 selected
variables

Propensity Score Model

* PS Model: Multivariable Logistic Regression Predicting Receipt of BP;

```
proc logistic data=ad01;  
  class agegrp (ref="≤50") region (ref="North America")  
    perfstat (ref="0") stage (ref="III")  
    extranod (ref="≥2") ldhlevel (ref="Elevated") / param=ref;  
  model tx (event="BP") = agegrp region perfstat stage extranod ldhlevel;  
  output out=ad02 pred=propensity;  
run;
```

Model receipt
of the Blue Pill

Output the
results into a
data set named
AD02

As a function of the
6 selected
variables

Propensity Score Model

* PS Model: Multivariable Logistic Regression Predicting Receipt of BP;

```
proc logistic data=ad01;  
  class agegrp (ref="≤50") region (ref="North America")  
    perfstat (ref="0") stage (ref="III")  
    extranod (ref="≥2") ldhlevel (ref="Elevated") / param=ref;  
  model tx (event="BP") = agegrp region perfstat stage extranod ldhlevel;  
  output out=ad02 pred=propensity;  
run;
```

Model receipt
of the Blue Pill

Output the
results into a
data set named
AD02

With the variable
PROPENSITY for
each patient

As a function of the
6 selected
variables

Propensity Score Model

* PS Model: Multivariable Logistic Regression Predicting Receipt of BP;

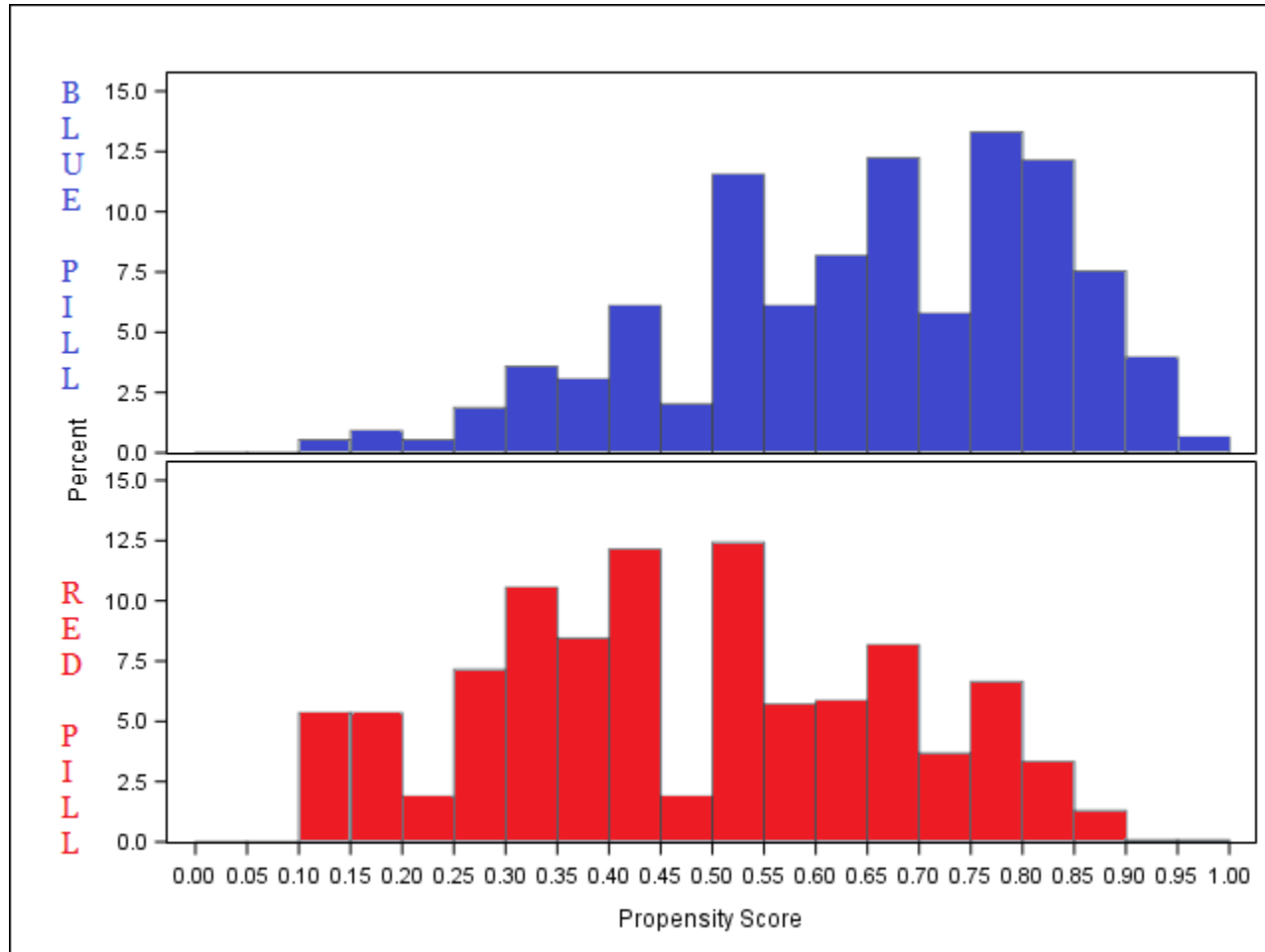
```
proc logistic data=ad01;  
  class agegrp (ref="<=50") region (ref="North America")  
    perfstat (ref="0") stage (ref="III")  
    extranod (ref=">=2") ldhlevel (ref="Elevated") / param=ref;  
  model tx (event="BP") = agegrp region perfstat stage extranod ldhlevel;  
  output out=ad02 pred=propensity;  
run;
```

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-1.1882	0.1860	40.7886	<.0001
agegrp	51 to 60	1	1.1653	0.1600	53.0770	<.0001
agegrp	61 to 70	1	2.1720	0.1610	181.9532	<.0001
agegrp	>=70	1	3.1191	0.2114	217.6932	<.0001
region	Asia	1	0.5041	0.1274	15.6471	<.0001
region	Europe	1	-0.4297	0.0835	26.4601	<.0001
perfstat	1	1	0.6027	0.0856	49.5625	<.0001
perfstat	>=2	1	1.2876	0.2182	34.8199	<.0001
stage	IV	1	0.1179	0.0811	2.1121	0.1461
extranod	<2	1	-0.0882	0.0849	1.0814	0.2984
ldhlevel	Normal	1	-0.3877	0.0925	17.5754	<.0001

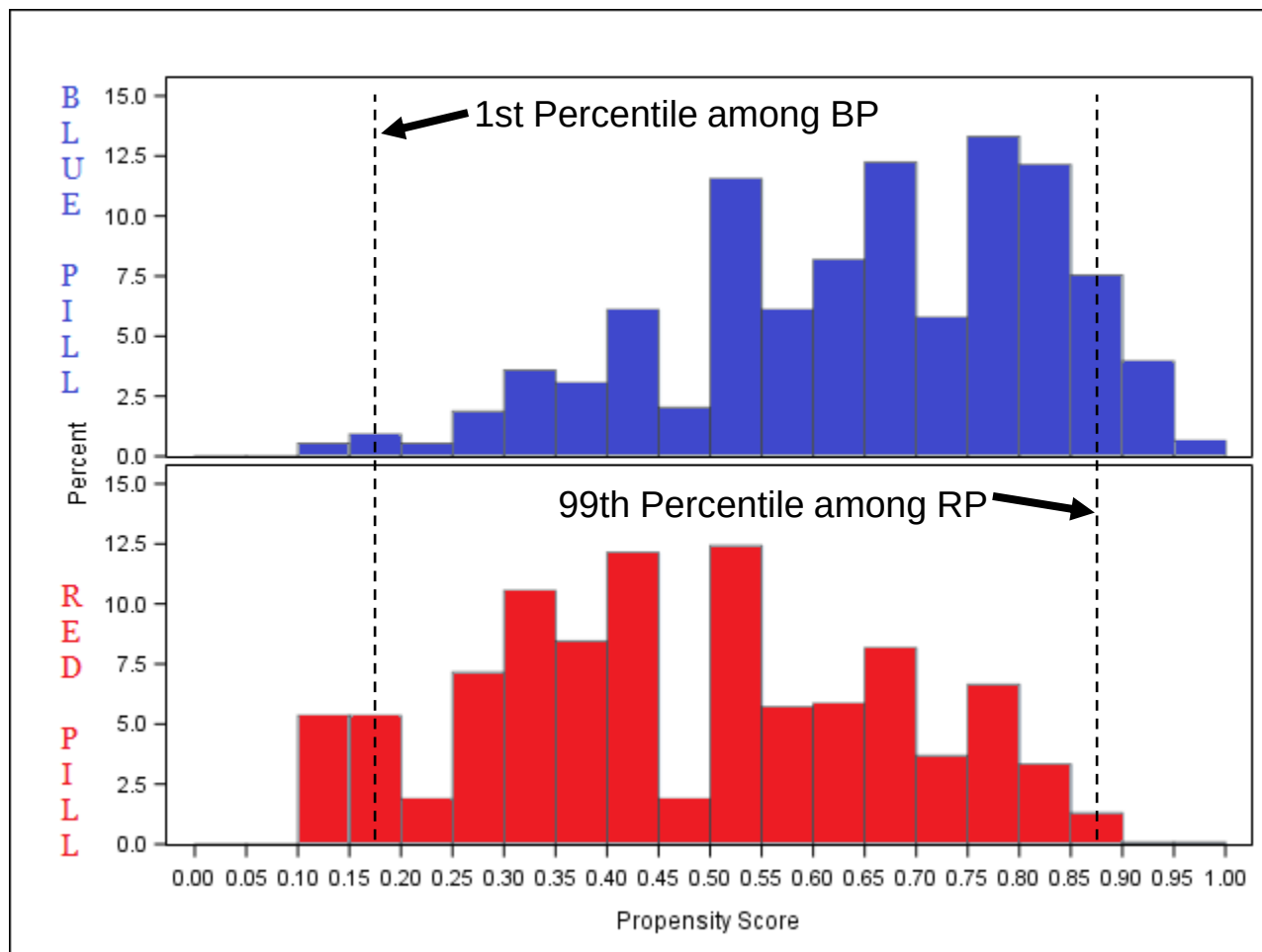
Propensity Score Trimming

- Propensity score distribution through stacked histograms



Propensity Score Trimming

- Propensity score distribution through stacked histograms



- 317 patients excluded from further analysis
- PS-trimmed cohort = 2,945 patients

Application of the Propensity Score

- We will cover 4 methods of applying the propensity score to estimate the adjusted treatment effect:
 - Stratification
 - Matching
 - Inverse Probability Treatment Weighting (IPTW)
 - Covariate adjustment
- Note about covariate balance

Application of the Propensity Score

- We will cover 4 methods of applying the propensity score to estimate the adjusted treatment effect:
 - Stratification
 - Matching
 - Inverse Probability Treatment Weighting (IPTW)
 - Covariate adjustment
- Note about covariate balance
- Continuous variables
- Categorical variables

$$d = \frac{(X_{BP} - X_{RP})}{\sqrt{\frac{s_{BP}^2 - s_{RP}^2}{2}}}$$

$$d = \frac{(p_{BP} - p_{RP})}{\sqrt{\frac{p_{BP}(1 - p_{BP}) + p_{RP}(1 - p_{RP})}{2}}}$$

Stratification

- Rank propensity scores into quintiles.
- Patients within any particular quintile are similar in characteristics.
- Treatment effect within a particular quintile should not be influenced by measured differences between BP and RP patients.

```
* Categorize PS-Trimmed Cohort into PS Quintiles;  
proc rank data=ad02 out=strat01 groups=5;  
  where (pstrimcohort eq 1);  
  var propensity;  
  ranks psrank;  
run;  
  
data strat02;  
  set strat01;  
  psquin = (psrank + 1);  
  label psquin = "Propensity Score Quintile (1 to 5)";  
run;
```

Stratification: Relative Treatment Effect

- Model each outcome as a function of treatment and propensity score quintile:

```
* PS Stratification;  
* Relative Treatment Effect (Hazard Ratio): OS;  
proc phreg data=strat02;  
    class tx (ref="BP") psquin;  
    model osmo*osevt(0) = tx psquin / rl;  
run;  
  
* Relative Treatment Effect (Hazard Ratio): PFS;  
proc phreg data=strat02;  
    class tx (ref="BP") psquin;  
    model pfsmo*pfsevt(0) = tx psquin / rl;  
run;  
  
* Relative Treatment Effect (Odds Ratio): ORR;  
proc logistic data=strat02;  
    class tx (ref="BP") psquin / param=ref;  
    model orr (event = "1") = tx psquin / rl;  
run;
```

Stratification: Relative Treatment Effect

- Model each outcome as a function of treatment and propensity score quintile:

```
* PS Stratification;
* Relative Treatment Effect (Hazard Ratio): OS;
proc phreg data=strat02;
  class tx (ref="BP") psquin;
  model osmo*osevt(0) = tx psquin / rl;
run;

* Relative Treatment Effect (Hazard Ratio): PFS;
proc phreg data=strat02;
  class tx (ref="BP") psquin;
  model pfsmo*pfsevt(0) = tx psquin / rl;
run;

* Relative Treatment Effect (Odds Ratio): ORR;
proc logistic data=strat02;
  class tx (ref="BP") psquin / param=ref;
  model orr (event = "1") = tx psquin / rl;
run;
```

Matching

- Each RP patient is matched to a BP patient with similar propensity score
- Various matching algorithms (we used 1:1 “greedy” matching)
- In our example:
 - 960 RP patients matched to 960 BP patients = 1,920 patients total
 - Additional 1,025 patients excluded
- Account for matched pairs in treatment effect.

Matching: Relative Treatment Effect

- Model each outcome adjusting for matched pairs:

- * PS Matching Accounting for Clustering Within Matched Pairs;

- * Relative Treatment Effect (Hazard Ratio): OS;

```
proc phreg data=ad_matching covs(aggregate);
  id matchid;
  class tx (ref="BP");
  model osmo*osevt(0) = tx / rl;
run;
```

- * Relative Treatment Effect (Hazard Ratio): PFS;

```
proc phreg data=ad_matching covs(aggregate);
  id matchid;
  class tx (ref="BP");
  model pfsmo*pfsevt(0) = tx / rl;
run;
```

- * PS Matching Accounting for Clustering Within Matched Pairs;

- * Relative Treatment Effect (Odds Ratio): ORR;

```
proc genmod data=ad_matching desc;
  class matchid;
  model orr = tx / dist=bin link=logit;
  repeated subject=matchid / type=un corrw covb;
  estimate 'RP vs. BP' tx 1 /exp;
run;
```

Matching: Relative Treatment Effect

- Model each outcome adjusting for matched pairs:

* PS Matching Accounting for Clustering Within Matched Pairs;

* Relative Treatment Effect (Hazard Ratio): OS;

```
proc phreg data=ad_matching covs(aggregate);
```

```
  id matchid;
```

```
  class tx (ref="BP");
```

```
  model osmo*osevt(0) = tx / rl;
```

```
run;
```

* Relative Treatment Effect (Hazard Ratio): PFS;

```
proc phreg data=ad_matching covs(aggregate);
```

```
  id matchid;
```

```
  class tx (ref="BP");
```

```
  model pfsmo*pfsevt(0) = tx / rl;
```

```
run;
```

* PS Matching Accounting for Clustering Within Matched Pairs;

* Relative Treatment Effect (Odds Ratio): ORR;

```
proc genmod data=ad_matching desc;
```

```
  class matchid;
```

```
  model orr = tx / dist=bin link=logit;
```

```
  repeated subject=matchid / type=un corrw covb;
```

```
  estimate 'RP vs. BP' tx 1 /exp;
```

```
run;
```

Matching: Relative Treatment Effect

- Model each outcome adjusting for matched pairs:

* PS Matching Accounting for Clustering Within Matched Pairs;

* Relative Treatment Effect (Hazard Ratio): OS;

```
proc phreg data=ad_matching covs(aggregate);
```

```
  id matchid;
```

```
  class tx (ref="BP");
```

```
  model osmo*osevt(0) = tx / rl;
```

```
run;
```

* Relative Treatment Effect (Hazard Ratio): PFS;

```
proc phreg data=ad_matching covs(aggregate);
```

```
  id matchid;
```

```
  class tx (ref="BP");
```

```
  model pfsmo*pfsevt(0) = tx / rl;
```

```
run;
```

* PS Matching Accounting for Clustering Within Matched Pairs;

* Relative Treatment Effect (Odds Ratio): ORR;

```
proc genmod data=ad_matching desc;
```

```
  class matchid;
```

```
  model orr = tx / dist=bin link=logit;
```

```
  repeated subject=matchid / type=un corrw covb;
```

```
  estimate 'RP vs. BP' tx 1 /exp;
```

```
run;
```

- Each patient weighted by the inverse of the probability of receiving the treatment that he or she actually received:
 - BP patients: $WEIGHT = 1 / PS$
 - RP patients: $WEIGHT = 1 / (1 - PS)$
- Weights are often “stabilized” as follows:
 - BP Patients: $STWEIGHT = P_{BP} / PS$
 - RP Patients: $STWEIGHT = (1 - P_{BP}) / (1 - PS)$

IPTW: Relative Treatment Effect

- Perform the appropriate regression while adjusting for stabilized weights:

```
* IPTW with Stabilized Weights;  
* Relative Treatment Effect (Hazard Ratio): OS;  
proc phreg data=ad_iptw;  
    class tx (ref="BP");  
    model osmo*osevt(0) = tx / rl;  
    weight stweight / normalize;  
run;  
  
* Relative Treatment Effect (Hazard Ratio): PFS;  
proc phreg data=ad_iptw;  
    class tx (ref="BP");  
    model pfsmo*pfsevt(0) = tx / rl;  
    weight stweight / normalize;  
run;  
  
* Relative Treatment Effect (Odds Ratio): ORR;  
proc logistic data=ad_iptw;  
    class tx (ref="BP") / param=ref;  
    model orr (event="1") = tx / rl;  
    weight stweight / normalize;  
run;
```

IPTW: Relative Treatment Effect

- Perform the appropriate regression while adjusting for stabilized weights:

```
* IPTW with Stabilized Weights;  
* Relative Treatment Effect (Hazard Ratio): OS;  
proc phreg data=ad_iptw;  
  class tx (ref="BP");  
  model osmo*osevt(0) = tx / rl;  
  weight stweight / normalize;  
run;  
  
* Relative Treatment Effect (Hazard Ratio): PFS;  
proc phreg data=ad_iptw;  
  class tx (ref="BP");  
  model pfsmo*pfsevt(0) = tx / rl;  
  weight stweight / normalize;  
run;  
  
* Relative Treatment Effect (Odds Ratio): ORR;  
proc logistic data=ad_iptw;  
  class tx (ref="BP") / param=ref;  
  model orr (event="1") = tx / rl;  
  weight stweight / normalize;  
run;
```

Covariate adjustment

- Most elementary approach
- Simply include propensity score as a continuous covariate in the model.

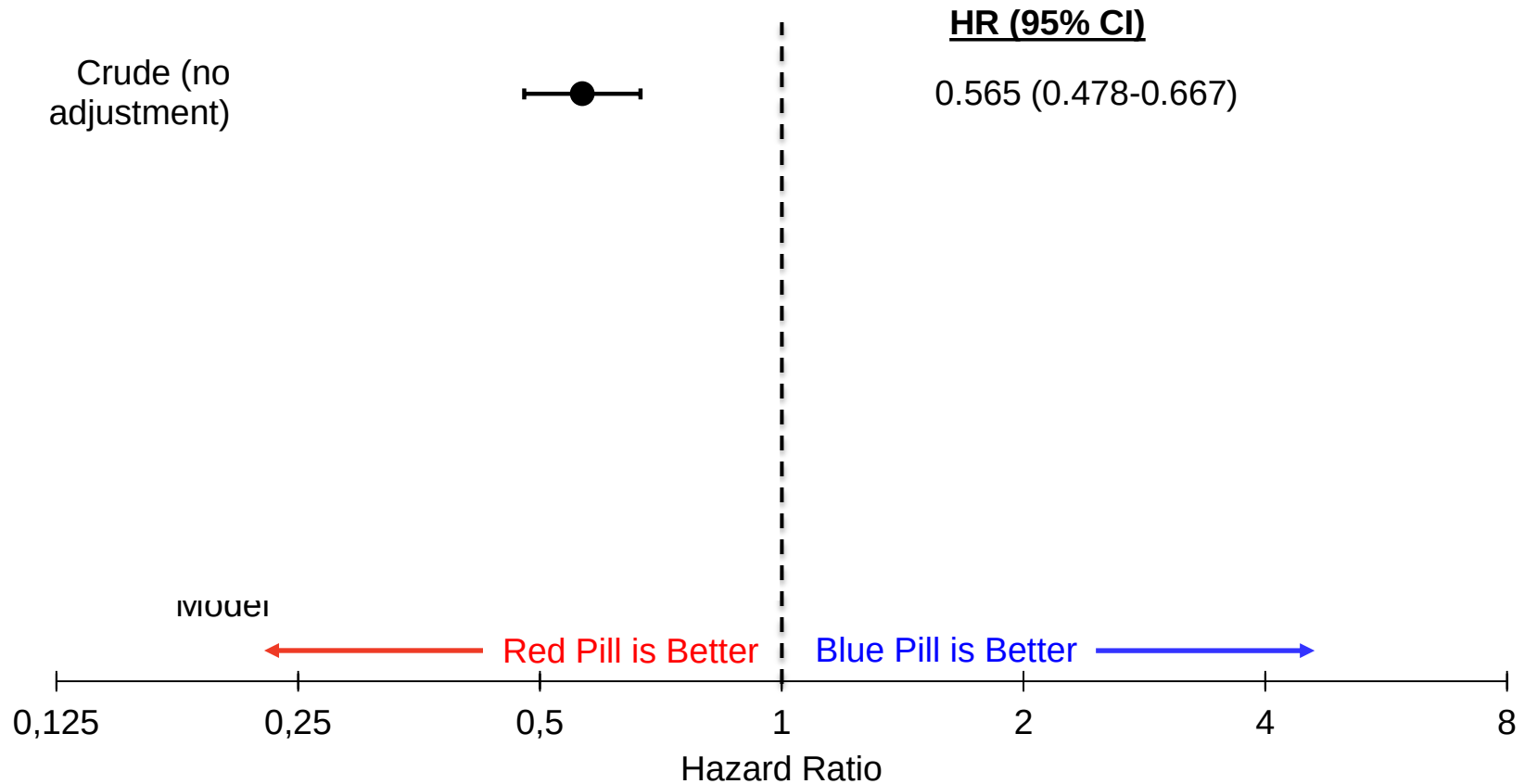
```
* Covariate Adjustment of PS;  
* Relative Treatment Effect: OS;  
proc phreg data=ad02;  
  class tx (ref="BP");  
  model osmo*osevt(0) = tx propensity / rl;  
run;  
  
* Relative Treatment Effect: PFS;  
proc phreg data=ad02;  
  class tx (ref="BP");  
  model pfsmo*pfsevt(0) = tx propensity / rl;  
run;  
  
* Relative Treatment Effect: ORR;  
proc logistic data=ad_matching;  
  class tx (ref="BP") / param=ref;  
  model orr (event = "1") = tx propensity / rl;  
run;
```

Covariate adjustment

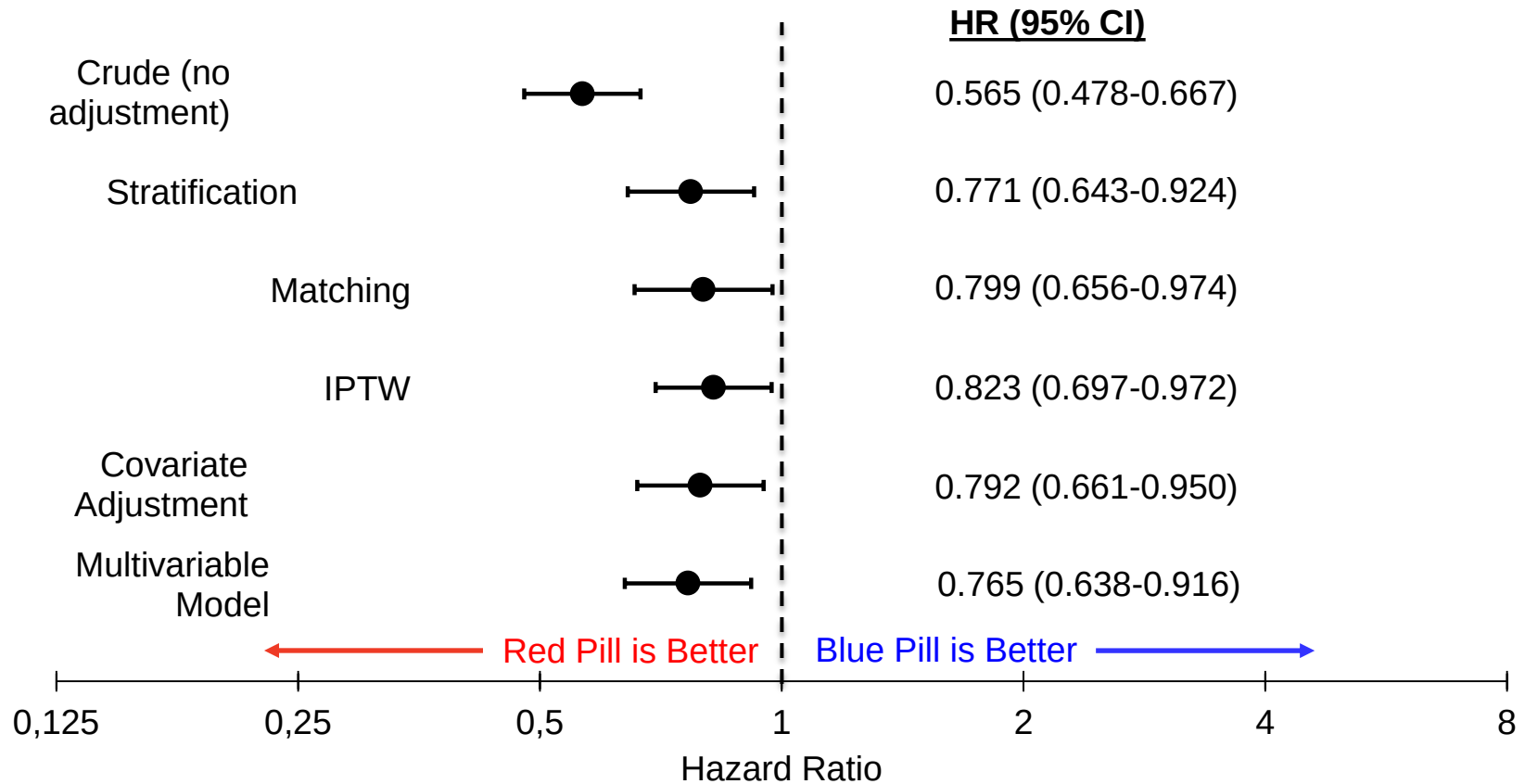
- Most elementary approach
- Simply include propensity score as a continuous covariate in the model.

```
* Covariate Adjustment of PS;  
* Relative Treatment Effect: OS;  
proc phreg data=ad02;  
  class tx (ref="BP");  
  model osmo*osevt(0) = tx propensity / rl;  
run;  
  
* Relative Treatment Effect: PFS;  
proc phreg data=ad02;  
  class tx (ref="BP");  
  model pfsmo*pfsevt(0) = tx propensity / rl;  
run;  
  
* Relative Treatment Effect: ORR;  
proc logistic data=ad_matching;  
  class tx (ref="BP") / param=ref;  
  model orr (event = "1") = tx propensity / rl;  
run;
```

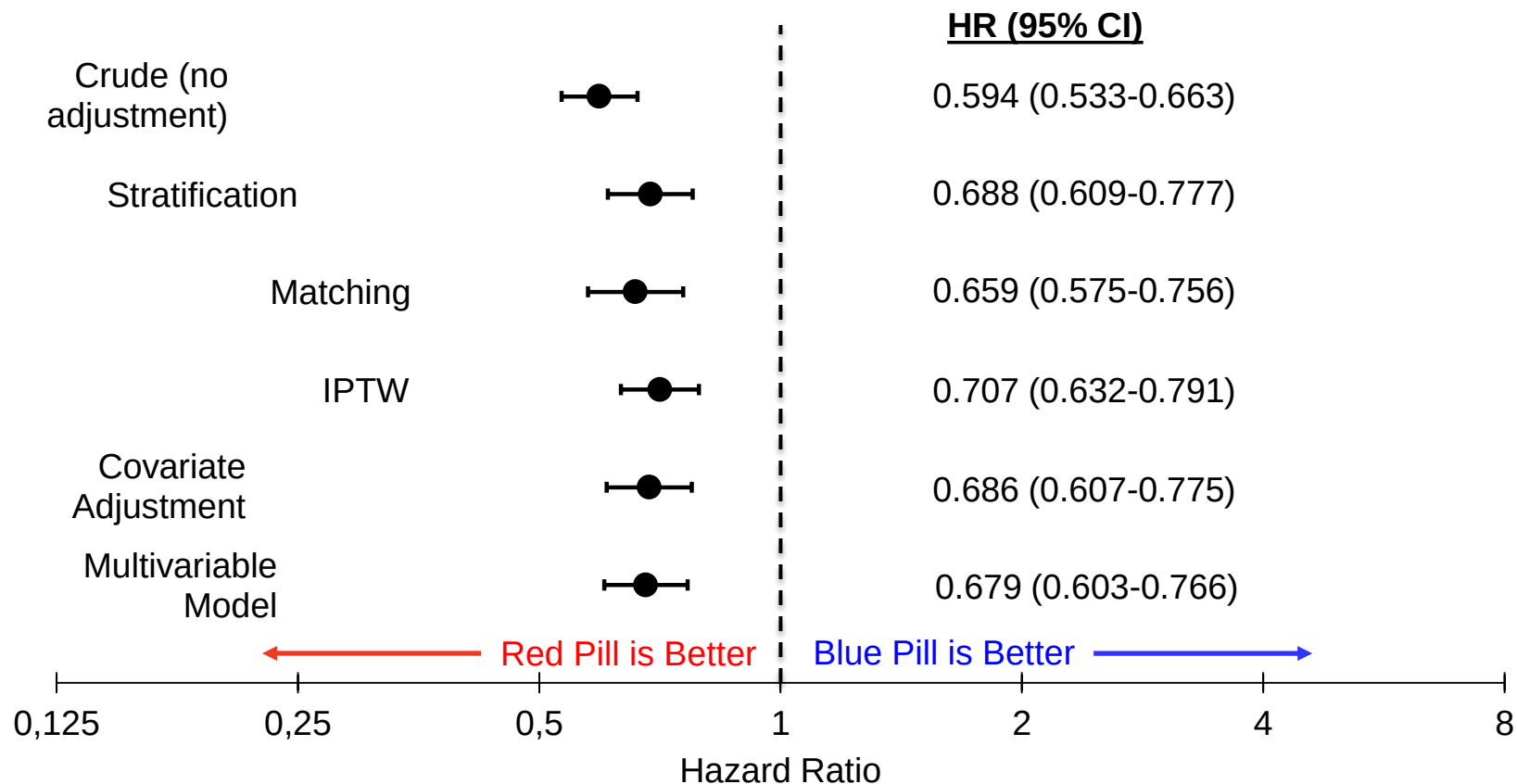
Results: OS



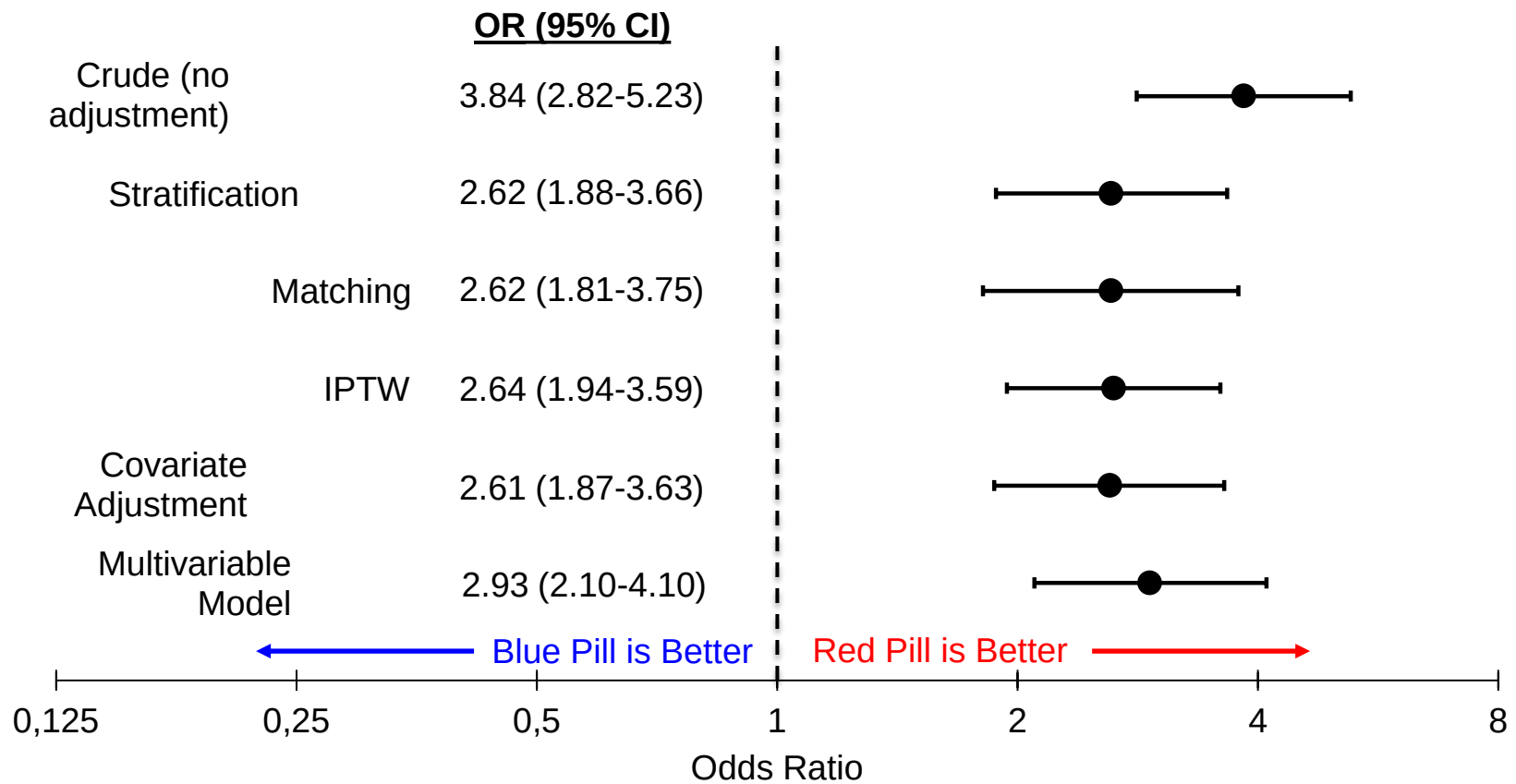
Results: OS



Results: PFS



Results: ORR



Conclusions

- Propensity scores are a powerful tool to deal with confounding when comparing treatments in observational studies, especially when there are few outcomes.
- Propensity score methods can adjust only for measured confounding variables.
- In our simulation:
 - Propensity score methods attenuated the crude observed treatment effect across all outcomes (all measures of association closer to the null).
 - Stratification, matching, and covariate adjustment yielded similar results for OS and PFS, while IPTW yielded point estimates closer to the null.
 - For ORR, similar odds ratios were obtained for all propensity score methods.
 - Multivariable model gave similar results to the propensity score results.

Conclusions

Choose the Red Pill! 😊



Contact Information

Lawrence Rasouliyan

+34 933 624 297

lrasouliyan@rti.org

Estel Plana

+34 933 622 832

eplana@rti.org

Jaume Aguado

+34 933 624 251

jaguado@rti.org