

# Lab 6: CASUAL INFERENCE

*Yuan Feng, Roderick Whang, Sebastián Soriano Pérez*

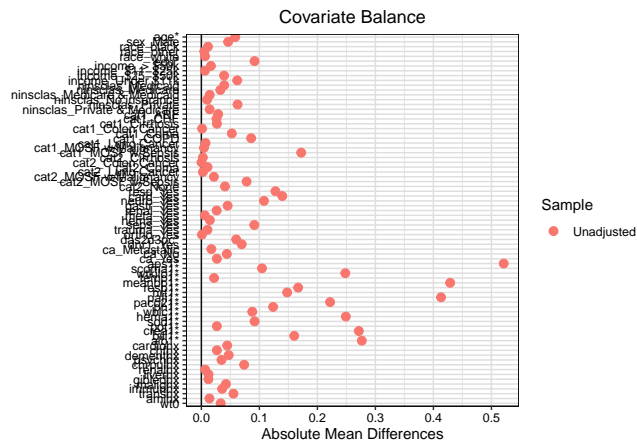
*11/15/2019*

## Answer 1

The distribution of several variables is not well ballanced across groups: cat1\_MOSF w/Sepsis, resp\_Yes, card\_Yes, neuro\_Yes, aps1, scoma1, wtkilo1, meanbp1, resp1, hrt1, paf1, paco21, ph1, hema1, crea1, bili1, alb1.

	Type	Diff.Un
age	Contin.	-0.0646862
sex_Male	Binary	0.0461617
race_black	Binary	-0.0113540
race_other	Binary	0.0050352
race_white	Binary	0.0063188
edu	Contin.	0.0910366
income_> \$50k	Binary	0.0164539
income_\$11-\$25k	Binary	0.0061712
income_\$25-\$50k	Binary	0.0391396
income_Under \$11k	Binary	-0.0617647
ninsclas_Medicaid	Binary	-0.0394813
ninsclas_Medicare	Binary	-0.0327111
ninsclas_Medicare & Medicaid	Binary	-0.0143656
ninsclas_No insurance	Binary	0.0098915
ninsclas_Private	Binary	0.0623893
ninsclas_Private & Medicare	Binary	0.0142773
cat1_ARF	Binary	-0.0290179
cat1_CHF	Binary	0.0261381
cat1_Cirrhosis	Binary	-0.0268460
cat1_Colon Cancer	Binary	-0.0012318
cat1_Coma	Binary	-0.0525311
cat1_COPD	Binary	-0.0858059
cat1_Lung Cancer	Binary	-0.0072854
cat1_MOSF w/Malignancy	Binary	0.0044761
cat1_MOSF w/Sepsis	Binary	0.1721039
cat2_Cirrhosis	Binary	-0.0025669
cat2_Colon Cancer	Binary	0.0001763
cat2_Coma	Binary	-0.0105552
cat2_Lung Cancer	Binary	-0.0027452
cat2_MOSF w/Malignancy	Binary	-0.0215987
cat2_MOSF w/Sepsis	Binary	0.0779737
cat2_None	Binary	-0.0406840
resp_Yes	Binary	-0.1276883

	Type	Diff.Un
card_Yes	Binary	0.1394948
neuro_Yes	Binary	-0.1078969
gastr_Yes	Binary	0.0453068
renal_Yes	Binary	0.0263688
meta_Yes	Binary	-0.0058546
hema_Yes	Binary	-0.0146493
seps_Yes	Binary	0.0912342
trauma_Yes	Binary	0.0104988
ortho_Yes	Binary	0.0009867
das2d3pc	Contin.	0.0654344
dnr1_Yes	Binary	-0.0695531
ca_Metastatic	Binary	-0.0171817
ca_No	Binary	0.0439190
ca_Yes	Binary	-0.0267373
aps1	Contin.	0.4837124
scoma1	Contin.	-0.1160363
wtkilo1	Contin.	0.2639761
temp1	Contin.	-0.0208689
meanbp1	Contin.	-0.4868502
resp1	Contin.	-0.1641331
hrt1	Contin.	0.1460095
pafi1	Contin.	-0.4566445
paco21	Contin.	-0.2880367
ph1	Contin.	-0.1162506
wblc1	Contin.	0.0798686
hema1	Contin.	-0.2954146
sod1	Contin.	-0.0926689
pot1	Contin.	-0.0274104
crea1	Contin.	0.2678085
bili1	Contin.	0.1329124
alb1	Contin.	-0.2009522
cardiohx	Binary	0.0445391
chfhx	Binary	0.0267570
dementhx	Binary	-0.0471661
psychhx	Binary	-0.0347531
chrpulhx	Binary	-0.0737360
renalhx	Binary	0.0065748
liverhx	Binary	-0.0123558
gibledhx	Binary	-0.0121657
malighx	Binary	-0.0422679
immunhx	Binary	0.0357878
transhx	Binary	0.0553857
amihx	Binary	0.0139290
wt0	Binary	-0.0333671



## Answer 2

The ATT for the patients with `rhc = TRUE` is 0.07299949.

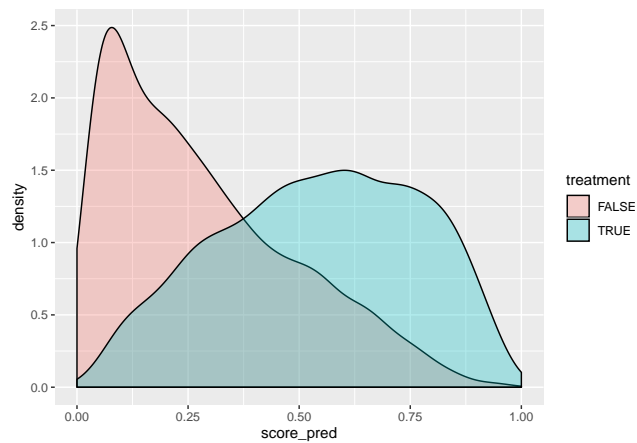
ATT
0.0729995

## Answer 3

The treated patients are worse off with RHC but we cannot trust this conclusion because the data is imbalanced.

## Answer 4

There are 103 observations that violate the overlap. We have discarded them from the dataset (`RHC_2`).



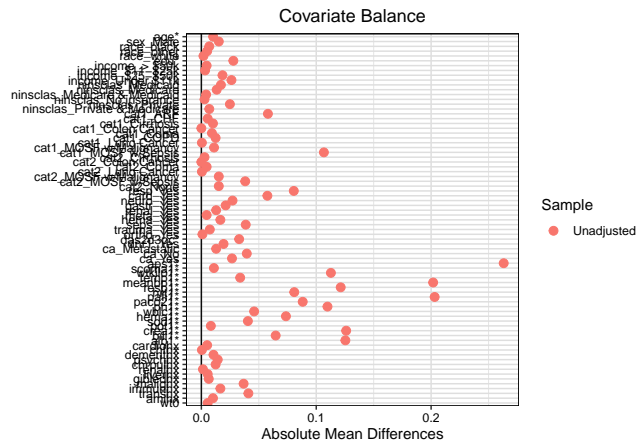
Overlap Violations	
1	94
2	9

## Answer 5

After using one-to-one, nearest neighbor matching on the estimated propensity scores, there are improvements overall but some covariates are still imbalanced, such as cat1, aps1, ph1, resp1, etc.

	Type	Diff.Un
age	Contin.	-0.0114173
sex_Male	Binary	0.0151724
race_black	Binary	-0.0068966
race_other	Binary	0.0050575
race_white	Binary	0.0018391
edu	Contin.	0.0278165
income_> \$50k	Binary	0.0045977
income_\$11-\$25k	Binary	0.0032184
income_\$25-\$50k	Binary	0.0183908
income_Under \$11k	Binary	-0.0262069
ninsclas_Medicaid	Binary	-0.0170115
ninsclas_Medicare	Binary	-0.0133333
ninsclas_Medicare & Medicaid	Binary	-0.0041379
ninsclas_No insurance	Binary	0.0027586
ninsclas_Private	Binary	0.0248276
ninsclas_Private & Medicare	Binary	0.0068966
cat1_ARF	Binary	-0.0579310
cat1_CHF	Binary	-0.0055172
cat1_Cirrhosis	Binary	-0.0101149
cat1_Colon Cancer	Binary	0.0000000
cat1_Coma	Binary	-0.0091954
cat1_COPD	Binary	-0.0124138
cat1_Lung Cancer	Binary	-0.0004598
cat1_MOSF w/Malignancy	Binary	-0.0110345
cat1_MOSF w/Sepsis	Binary	0.1066667
cat2_Cirrhosis	Binary	-0.0027586
cat2_Colon Cancer	Binary	0.0000000
cat2_Coma	Binary	-0.0045977
cat2_Lung Cancer	Binary	-0.0004598
cat2_MOSF w/Malignancy	Binary	-0.0151724
cat2_MOSF w/Sepsis	Binary	0.0381609
cat2_None	Binary	-0.0151724
resp_Yes	Binary	-0.0804598
card_Yes	Binary	0.0574713
neuro_Yes	Binary	-0.0271264
gastr_Yes	Binary	0.0211494
renal_Yes	Binary	0.0128736
meta_Yes	Binary	-0.0045977
hema_Yes	Binary	-0.0165517
seps_Yes	Binary	0.0386207
trauma_Yes	Binary	0.0073563
ortho_Yes	Binary	0.0009195
das2d3pc	Contin.	0.0358382
dnr1_Yes	Binary	-0.0193103

	Type	Diff.Un
ca_Metastatic	Binary	-0.0128736
ca_No	Binary	0.0395402
ca_Yes	Binary	-0.0266667
aps1	Contin.	0.2471414
scoma1	Contin.	-0.0112851
wtkilo1	Contin.	0.1145554
temp1	Contin.	-0.0340210
meanbp1	Contin.	-0.2119686
resp1	Contin.	-0.1180803
hrt1	Contin.	0.0826272
pafi1	Contin.	-0.2056386
paco21	Contin.	-0.0858324
ph1	Contin.	-0.1023420
wblc1	Contin.	0.0453669
hema1	Contin.	-0.0815870
sod1	Contin.	-0.0413264
pot1	Contin.	0.0084276
crea1	Contin.	0.1399223
bili1	Contin.	0.0618434
alb1	Contin.	-0.0933203
cardiohx	Binary	0.0050575
chfhx	Binary	-0.0004598
dementhx	Binary	-0.0105747
psychhx	Binary	-0.0142529
chrpulhx	Binary	-0.0124138
renalhx	Binary	-0.0013793
liverhx	Binary	-0.0055172
gibledhx	Binary	-0.0064368
malighx	Binary	-0.0367816
immunhx	Binary	0.0165517
transhx	Binary	0.0409195
amihx	Binary	0.0101149
wt0	Binary	-0.0055172



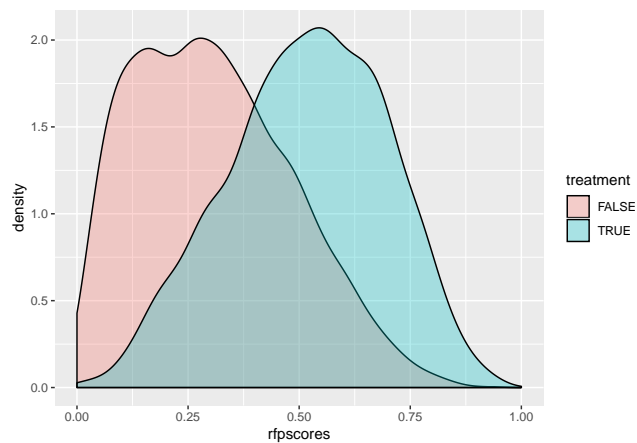
## Answer 6

The treated patients are worse off with RHC. The average causal effect Q using the matched sample obtained above is 0.07908; The 95% confidence interval is [0.0502, 0.1079].

X2.5.percent	Q	X97.5.percent
0.0502283	0.0790805	0.1079326

## Answer 7

There are 103 observations that violate the overlap once again. We have discarded them from the dataset (RHC\_3).



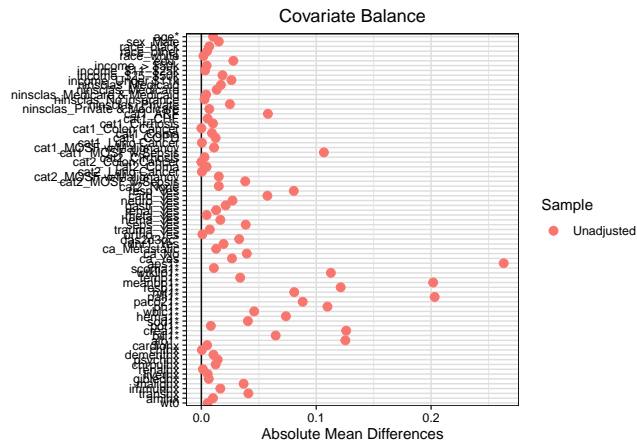
Overlap Violations	
1	94
2	9

## Answer 8

After using one-to-one, nearest neighbor matching on the new estimated propensity scores from question 7, some of covariates are still imbalanced, such as temp1, meanbp1, resp1, crea1, etc.

	Type	Diff.Un
age	Contin.	-0.0114173
sex_Male	Binary	0.0151724
race_black	Binary	-0.0068966
race_other	Binary	0.0050575
race_white	Binary	0.0018391
edu	Contin.	0.0278165
income_> \$50k	Binary	0.0045977
income_\$11-\$25k	Binary	0.0032184
income_\$25-\$50k	Binary	0.0183908
income_Under \$11k	Binary	-0.0262069
ninsclas_Medicaid	Binary	-0.0170115
ninsclas_Medicare	Binary	-0.0133333
ninsclas_Medicare & Medicaid	Binary	-0.0041379
ninsclas_No insurance	Binary	0.0027586
ninsclas_Private	Binary	0.0248276
ninsclas_Private & Medicare	Binary	0.0068966
cat1_ARF	Binary	-0.0579310
cat1_CHF	Binary	-0.0055172
cat1_Cirrhosis	Binary	-0.0101149
cat1_Colon Cancer	Binary	0.0000000
cat1_Coma	Binary	-0.0091954
cat1_COPD	Binary	-0.0124138
cat1_Lung Cancer	Binary	-0.0004598
cat1_MOSF w/Malignancy	Binary	-0.0110345
cat1_MOSF w/Sepsis	Binary	0.1066667
cat2_Cirrhosis	Binary	-0.0027586
cat2_Colon Cancer	Binary	0.0000000
cat2_Coma	Binary	-0.0045977
cat2_Lung Cancer	Binary	-0.0004598
cat2_MOSF w/Malignancy	Binary	-0.0151724
cat2_MOSF w/Sepsis	Binary	0.0381609
cat2_None	Binary	-0.0151724
resp_Yes	Binary	-0.0804598
card_Yes	Binary	0.0574713
neuro_Yes	Binary	-0.0271264
gastr_Yes	Binary	0.0211494
renal_Yes	Binary	0.0128736
meta_Yes	Binary	-0.0045977
hema_Yes	Binary	-0.0165517
seps_Yes	Binary	0.0386207
trauma_Yes	Binary	0.0073563
ortho_Yes	Binary	0.0009195
das2d3pc	Contin.	0.0358382
dnr1_Yes	Binary	-0.0193103

	Type	Diff.Un
ca_Metastatic	Binary	-0.0128736
ca_No	Binary	0.0395402
ca_Yes	Binary	-0.0266667
aps1	Contin.	0.2471414
scoma1	Contin.	-0.0112851
wtkilo1	Contin.	0.1145554
temp1	Contin.	-0.0340210
meanbp1	Contin.	-0.2119686
resp1	Contin.	-0.1180803
hrt1	Contin.	0.0826272
pafi1	Contin.	-0.2056386
paco21	Contin.	-0.0858324
ph1	Contin.	-0.1023420
wblc1	Contin.	0.0453669
hema1	Contin.	-0.0815870
sod1	Contin.	-0.0413264
pot1	Contin.	0.0084276
crea1	Contin.	0.1399223
bili1	Contin.	0.0618434
alb1	Contin.	-0.0933203
cardiohx	Binary	0.0050575
chfhx	Binary	-0.0004598
dementhx	Binary	-0.0105747
psychhx	Binary	-0.0142529
chrpulhx	Binary	-0.0124138
renalhx	Binary	-0.0013793
liverhx	Binary	-0.0055172
gibledhx	Binary	-0.0064368
malighx	Binary	-0.0367816
immunhx	Binary	0.0165517
transhx	Binary	0.0409195
amihx	Binary	0.0101149
wt0	Binary	-0.0055172





**Answer 9**

The average causal effect Q using the new matched sample obtained in question 8 is 0.08268; The 95% confidence interval is [0.05409,0.1113]. Treated patients are worse off with RHC.

X2.5.percent	Q	X97.5.percent
0.0501417	0.0790805	0.1080192

**Answer 10**

RHC have negative effect on patients as they have high odds of dying after RHC treatment.

X2.5.percent	Q	X97.5.percent
0.0511344	0.08	0.1088656