Interim Presentation

author: Dylan Sun & Zijiang Yang date: autosize: true {r, message=FALSE, warning=FALSE, echo = FALSE} library(readxl) library(knitr) library(data.table) library(corrplot) library(broom) library(stargazer) library(ggplot2) library(gridExtra) prostate <- read_excel("~/Downloads/Prostate SBRT Sexual Function Data.xlsx", skip = 7) prostate <- data.table(prostate) prostate <- prostate[complete.cases(prostate)] # remove last row which is empty # Code ADT as numeric; Y = 1, N = 0 prostate[ADT == "Y", ADT := 1] prostate[ADT == "N", ADT := 0] prostate[, ADT := as.numeric(ADT)] prostate[, centered_age := Age - mean(Age)]

Objectives

- Determining the erectile function rates following SBRT
- Determining predictors of erectile function preservation after SBRT
- If you undergo SBRT, how likely is it that you will lose erectile function?

Potential Predictors

Negatively Associated: - Age, Gleason Score, T-stage, PSA, ADT, BMI

Positively Associated: - HRQOL, Erectile Function at Baseline

We expect HRQOL and Erectile Function at Baseline to be highly correlated, because erectile function is included in the questionnaire.

How many people lose function?

- At baseline: 158 with function, 167 without (0.49)

- At two years: 102 with function, 223 without (0.31)

- A seemingly large number of people lose function in just two years

{r echo=FALSE} nf_nf <- dim(prostate['Erectile Function at Baseline' == 0 & 'Erectile Function at 2 years' == 0])[1] nf_hf <- dim(prostate['Erectile Function at Baseline' == 0 & 'Erectile Function at 2 years' == 1])[1] hf_nf <- dim(prostate['Erectile Function at Baseline' == 1 & 'Erectile Function at 2 years' == 0])[1] hf_hf <- dim(prostate['Erectile Function at Baseline' == 1 & 'Erectile Function at 2 years' == 1])[1] counts <- data.table(never_functional = nf_nf, gain_function = nf_hf, loss_function = hf_nf, retain_function = hf_hf) kable(counts)

Questions and Concerns

- BMI Outlier

```
{r echo=FALSE} bmi_hist <- ggplot(data = prostate, aes(x = BMI))
+    geom_histogram() bmi_hist
```

Subject 67 has a BMI of 112. This is theoretically possible.

Multicollinearity

```
{r, echo=FALSE, fig.height=10, fig.width=17} correlations <-
cor(prostate[, .(Age, 'Gleason Score', 'T-Stage Group', PSA,
HRQOL, ADT, BMI, Erec_base = 'Erectile Function at Baseline',
Erec_two = 'Erectile Function at 2 years')]) corrplot(correlations,
method = "circle", tl.cex = 1.5, number.cex = 0.6)
```

Model

$logit(EF2) = \beta_{intercept} + \beta_{age}Age + \beta_{Gleason}Gleason$
$+ \beta_{Tstage}TStage + \beta_{PSA}PSA + \beta_{HRQOL}HRQOL +$
$\beta_{ADT}ADT + \beta_{BMI}BMI + \beta_{EFbase}EFBase$

Results

"{r echo=FALSE} prostate_logreg <- glm(Erectile Function at 2 years~
centered_age +Gleason Score+T-Stage Group+ PSA + HRQOL + ADT + BMI
+Erectile Function at Baseline', data = prostate, family = binomial(link =
"logit")) # summary(prostate_logreg)

edit results table

prostate_logreg_results <- data.table(tidy(prostate_logreg)) prostate_logreg_results[,
lower := exp(estimate - std.error)] prostate_logreg_results[, OR :=
exp(estimate)] prostate_logreg_results[, upper := exp(estimate + std.error)]
kable(prostate_logreg_results[, .(term, lower, OR, upper, p.value)], digits = 2)
"'

Results

Error rate $= \dfrac{\sum |predicted - actual|}{total}$ {r echo=FALSE} prostate <-
prostate[, predicted_prob := predict(prostate_logreg, type =
"response")] prostate <- prostate[, predicted := ifelse(predicted_prob
>= 0.5, 1, 0)] error_rate <- prostate[, sum(abs(predicted -
'Erectile Function at 2 years'))]/dim(prostate)[1] error_rate

Conclusions

- Age, HRQOL, and erectile function are predictive of erectile function at two years
- PSA is somewhat protective
- Holding everything else constant, each one-year increase in baseline age increases the odds of losing erectile function by 6%.
- Holding everything else constant, HRQOL score is protective of erectile function.

Prediction

```
{r, echo = FALSE} example_prostate <- data.table(centered_age
= 0, "Gleason Score" =  7, "T-Stage Group" = 0, "PSA" = 7.7,
HRQOL = 56, ADT = 0, BMI = 29, "Erectile Function at Baseline" =
1) example_prostate <- example_prostate[, probability_of_function
:= predict(prostate_logreg, example_prostate, type = "response")]
example_prostate <- example_prostate[, predicted_function :=
ifelse(probability_of_function >= 0.5, 1, 0)] example_prostate
<- rbind(example_prostate, prostate[c(143, 179, 199, 231, 149),
.(centered_age, 'Gleason Score', 'T-Stage Group', PSA, HRQOL,
ADT, BMI, 'Erectile Function at Baseline', probability_of_function
= predicted_prob, predicted_function = predicted)])[, .(cen_age
= centered_age, 'Gleason Score', T_Stage = 'T-Stage Group', PSA,
HRQOL, ADT, BMI, Erec_base = 'Erectile Function at Baseline',
prob_func = probability_of_function, pred_func = predicted_function)]
kable(example_prostate, digits = 2)
```