

Creating Effect Size Plots (ggplot2)

Jacky Choi, MPH

jmc4005@med.cornell.edu

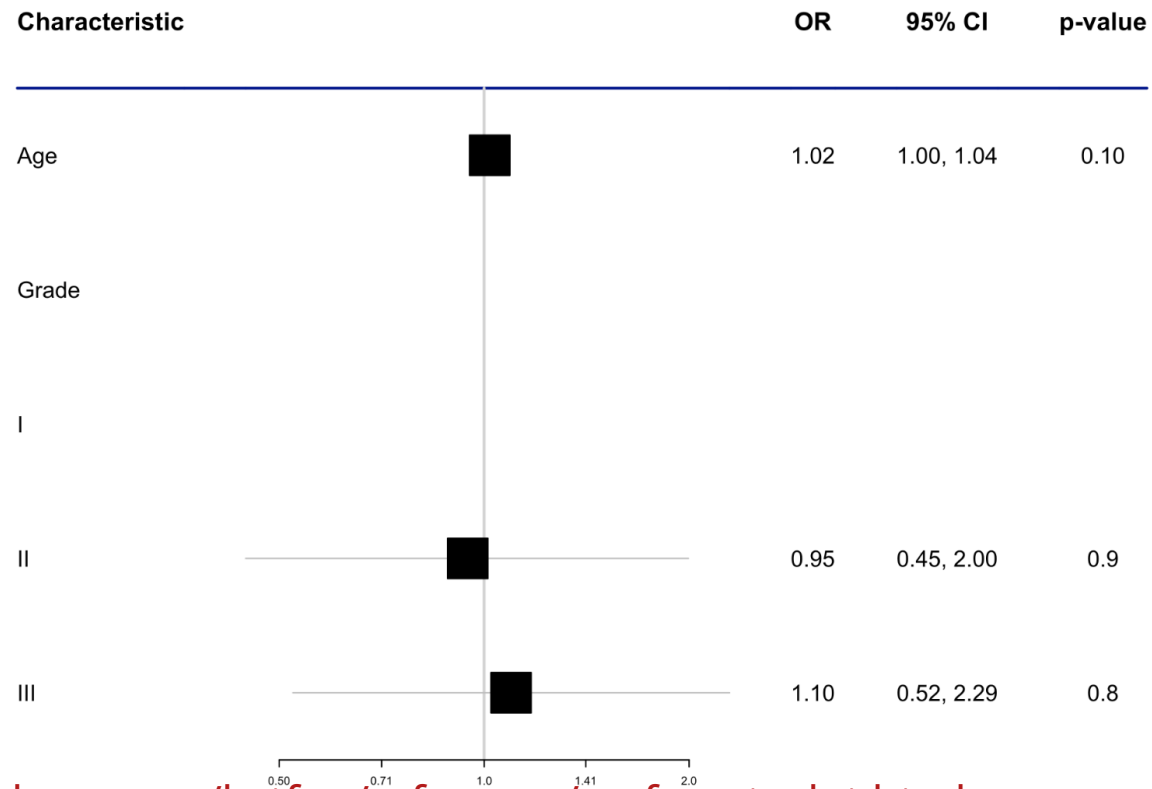
2022-06-23

About

- Effect size plots illustrate the impact of your coefficients in your regression model
- Easy to understand visually and good summary of your findings
- Using ggplot2 allows customizable edits and additions
- Can use `gtsummary` as a quick option
- Code here can be adapted for forest plots, but look into package `forestplot`

Example

Unadjusted results between several covariates and tumor response



1 https://www.danielsjoberg.com/bstfun/reference/as_forest_plot.html

Components

Table

Contains the estimate and CI, the outcomes and/or other variables of interest

- Optionally, cases/events in each substratum, p-values, etc.

Plot

Displays the actual plot associated with the estimate and CI

- Can customize heavily here

Let's use an example and start building the components!

Example

- Retrospective study examining the effect of hip arthroplasty on three separate primary outcomes (related return to OR, unplanned and related readmission, and prolonged LOS) after adjusting for other covariates
- After data cleaning, we obtained 146 patients who underwent hip arthroplasty and 332 patients who did not undergo the procedure

```
##          cohort    sex age  bmi race_new rtor readmit prolonged_los
## 1 No Hip Arthroplasty  Male  62 36.41   Asian    0     0           0
## 2 No Hip Arthroplasty  Male  50 26.64   White     0     0           0
## 3 No Hip Arthroplasty  Female 35 20.20   Asian    0     0           1
## 4 No Hip Arthroplasty  Female 54 31.58           0     0           0
## 5 Hip Arthroplasty    Male  43 30.57   White     0     0           1
## 6 No Hip Arthroplasty  Female 53 21.63   White     0     0           0
## 7 No Hip Arthroplasty  Female 21 20.25   White     0     0           1
## 8 No Hip Arthroplasty  Male  54 40.06   White     0     0           1
## 9 Hip Arthroplasty    Male  88 17.55   White     0     0           1
## 10 No Hip Arthroplasty  Female 49 28.48   Asian    0     0           0
```

Results

Table 3. Multivariable Logistic Reg Model - Related Return to OR

Characteristic	Estimate ¹	95% CI ¹	p-value
Cohort			
No Hip Arthroplasty	—	—	
Hip Arthroplasty	1.05	0.34, 3.00	>0.9
Sex			
Female	—	—	
Male	0.27	0.09, 0.72	0.014

Table 4. Multivariable Logistic Reg Model: Unplanned Related Readmission

Characteristic	Estimate ¹	95% CI ¹	p-value
Cohort			
No Hip Arthroplasty	—	—	
Hip Arthroplasty	1.05	0.46, 2.30	>0.9
Sex			
Female	—	—	
Male	1.02	0.52, 2.00	>0.9

Table 5. Multivariable Logistic Reg Model - Prolonged LOS

Characteristic	Estimate ¹	95% CI ¹	p-value
Cohort			
No Hip Arthroplasty	—	—	
Hip Arthroplasty	1.72	0.99, 2.99	0.052
Sex			
Female	—	—	
Male	0.87	0.54, 1.40	0.6

Table

```
#Set up table
table_df <- data.frame(labels = c("Outcome",
                                #Outcome 1
                                "Related Return to OR", "  No Hip Arthroplasty", "  Hip Arthroplasty",
                                #Outcome 2
                                "Unplanned Related Readmission", "  No Hip Arthroplasty ", "  Hip Arthroplasty ",
                                #Outcome 3
                                "Prolonged LOS", "  No Hip Arthroplasty  ", "  Hip Arthroplasty  "),
  eventnum = c("events/N",
              #Outcome 1
              NA, "21/332", "7/146",
              #Outcome 2,
              NA, "33/332", "11/146",
              #Outcome 3,
              NA, "71/332", "44/146"),
  or_ci = c("aOR (95% CI)",
           NA, "----", "1.05 (0.34, 3.00)",
           NA, "----", "1.05 (0.46, 2.30)",
           NA, "----", "1.72 (0.99, 2.99)"),
  color = c(NA,
           #Outcome 1
           "gray95", NA, NA,
           #Outcome 2
           "gray95", NA, NA,
           #Outcome 3
           "gray95", NA, NA))
```

Table

labels	eventnum	or_ci	color
Outcome	events/N	aOR (95% CI)	NA
Related Return to OR	NA	NA	gray95
No Hip Arthroplasty	21/332	---	NA
Hip Arthroplasty	7/146	1.05 (0.34, 3.00)	NA
Unplanned Related Readmission	NA	NA	gray95
No Hip Arthroplasty	33/332	---	NA
Hip Arthroplasty	11/146	1.05 (0.46, 2.30)	NA
Prolonged LOS	NA	NA	gray95
No Hip Arthroplasty	71/332	---	NA
Hip Arthroplasty	44/146	1.72 (0.99, 2.99)	NA

Plot

```
#Set up plot
plot_df <- data.frame(labels = c(table_df$labels),
                      or = c(NA,
                            #Outcome 1
                            NA, NA, 1.05,
                            #Outcome 2
                            NA, NA, 1.05,
                            #Outcome 3
                            NA, NA, 1.72),
                      or_low = c(NA,
                                 #Outcome 1
                                 NA, NA, 0.34,
                                 #Outcome 2
                                 NA, NA, 0.46,
                                 #Outcome 3
                                 NA, NA, 0.99),
                      or_high = c(NA,
                                  #Outcome 1
                                  NA, NA, 3.00,
                                  #Outcome 2
                                  NA, NA, 2.30,
                                  #Outcome 3
                                  NA, NA, 2.99))
```

labels	or	or_low	or_high
Outcome	NA	NA	NA
Related Return to OR	NA	NA	NA
No Hip Arthroplasty	NA	NA	NA
Hip Arthroplasty	1.05	0.34	3.00
Unplanned Related Readmission	NA	NA	NA
No Hip Arthroplasty	NA	NA	NA
Hip Arthroplasty	1.05	0.46	2.30
Prolonged LOS	NA	NA	NA
No Hip Arthroplasty	NA	NA	NA
Hip Arthroplasty	1.72	0.99	2.99

Join Table and Plot

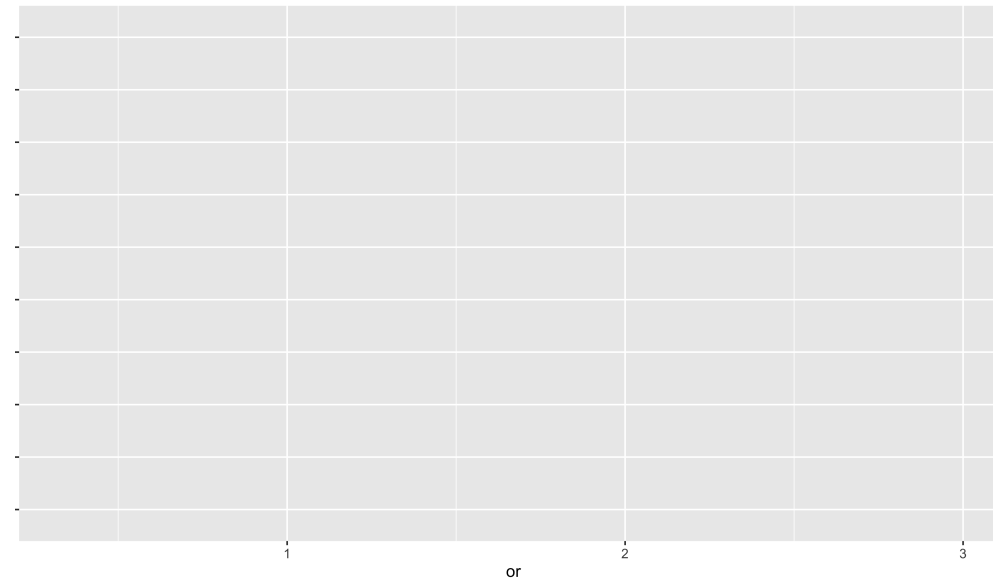
```
#Join table and plot
table_plot_df <- table_df %>%
  left_join(plot_df, "labels")

#Set labels in final plot
table_plot_df$labels <- factor(table_plot_df$labels, levels = table_plot_df$labels)
```

labels	eventnum	or_ci	color	or	or_low	or_high
Outcome	events/N	aOR (95% CI)	NA	NA	NA	NA
Related Return to OR	NA	NA	gray95	NA	NA	NA
No Hip Arthroplasty	21/332	---	NA	NA	NA	NA
Hip Arthroplasty	7/146	1.05 (0.34, 3.00)	NA	1.05	0.34	3.00
Unplanned Related Readmission	NA	NA	gray95	NA	NA	NA
No Hip Arthroplasty	33/332	---	NA	NA	NA	NA
Hip Arthroplasty	11/146	1.05 (0.46, 2.30)	NA	1.05	0.46	2.30
Prolonged LOS	NA	NA	gray95	NA	NA	NA
No Hip Arthroplasty	71/332	---	NA	NA	NA	NA
Hip Arthroplasty	44/146	1.72 (0.99, 2.99)	NA	1.72	0.99	2.99

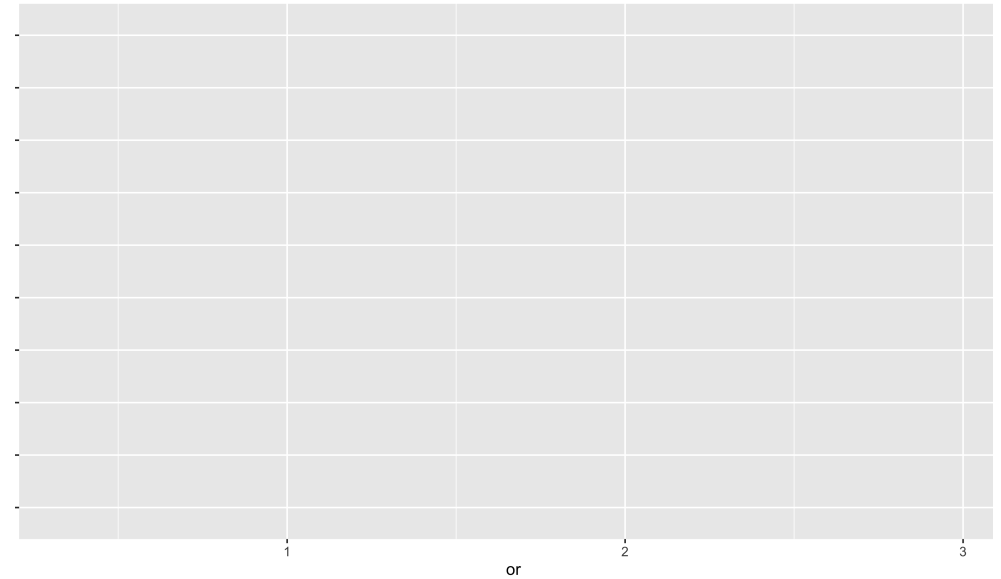
Plot - ggplot

```
#ggplot for plot  
ggplot(table_plot_df, aes(x = or, y = labels, xmin = or_low, xmax =  
  theme(axis.text.y = element_blank(),  
        axis.title.y = element_blank(),  
        plot.margin = margin(55, -10, 33, 0))
```



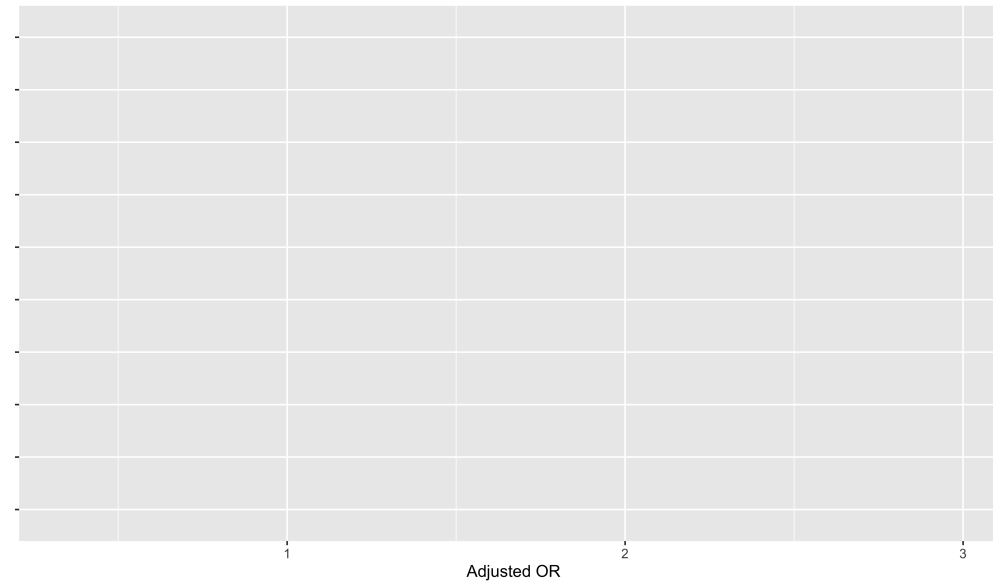
Plot - ggplot

```
#ggplot for plot  
ggplot(table_plot_df, aes(x = or, y = labels, xmin = or_low, xmax =  
  scale_y_discrete(limits = rev(table_plot_df$labels)) +  
  theme(axis.text.y = element_blank(),  
        axis.title.y = element_blank(),  
        plot.margin = margin(55, -10, 33, 0))
```



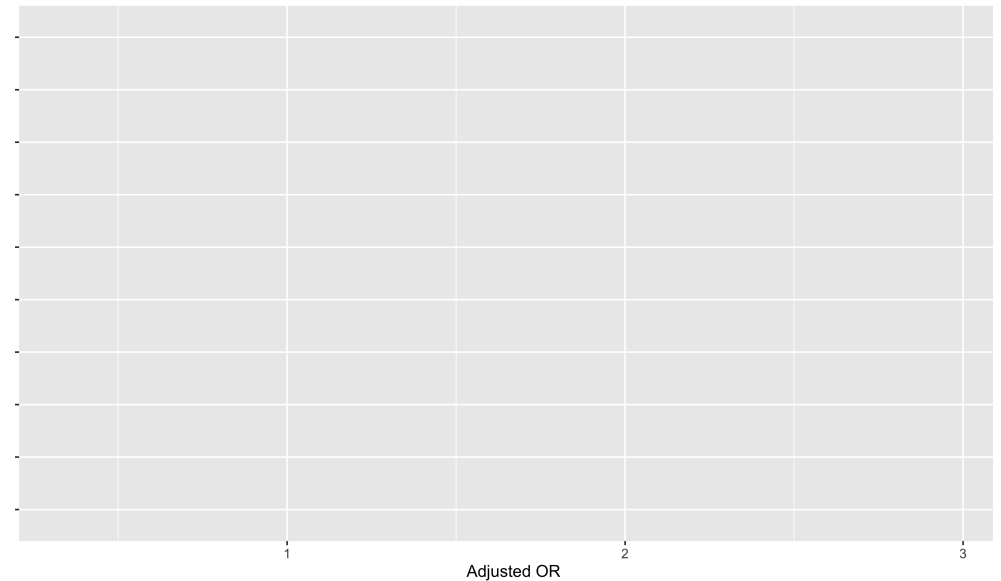
Plot - ggplot

```
#ggplot for plot  
ggplot(table_plot_df, aes(x = or, y = labels, xmin = or_low, xmax =  
  scale_y_discrete(limits = rev(table_plot_df$labels)) +  
  xlab("Adjusted OR") +  
  theme(axis.text.y = element_blank(),  
        axis.title.y = element_blank(),  
        plot.margin = margin(55, -10, 33, 0))
```



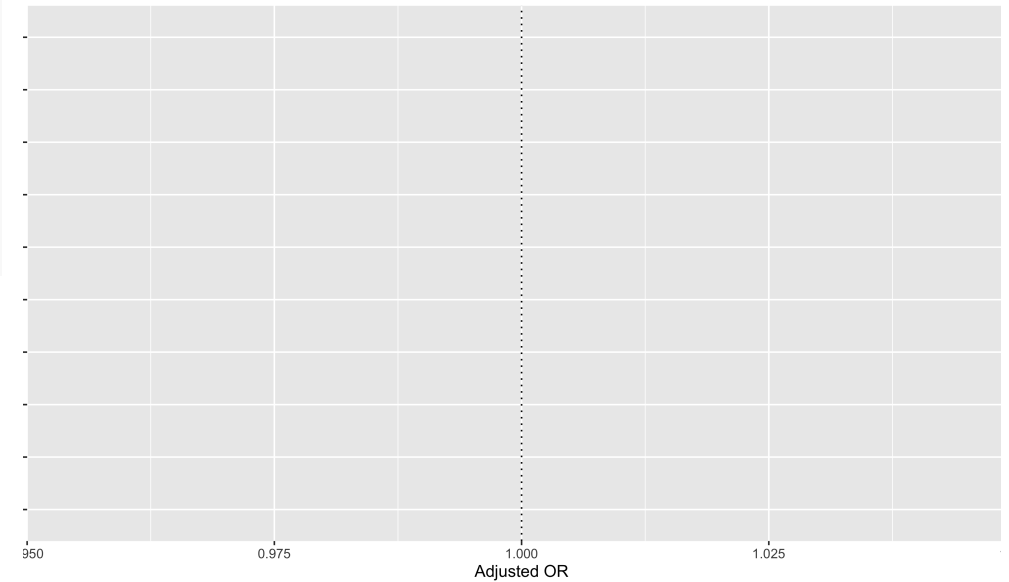
Plot - ggplot

```
#ggplot for plot  
ggplot(table_plot_df, aes(x = or, y = labels, xmin = or_low, xmax =  
  scale_y_discrete(limits = rev(table_plot_df$labels)) +  
  xlab("Adjusted OR") +  
  ylab("Adjusted OR with 95% Confidence Interval") +  
  theme(axis.text.y = element_blank(),  
        axis.title.y = element_blank(),  
        plot.margin = margin(55, -10, 33, 0))
```



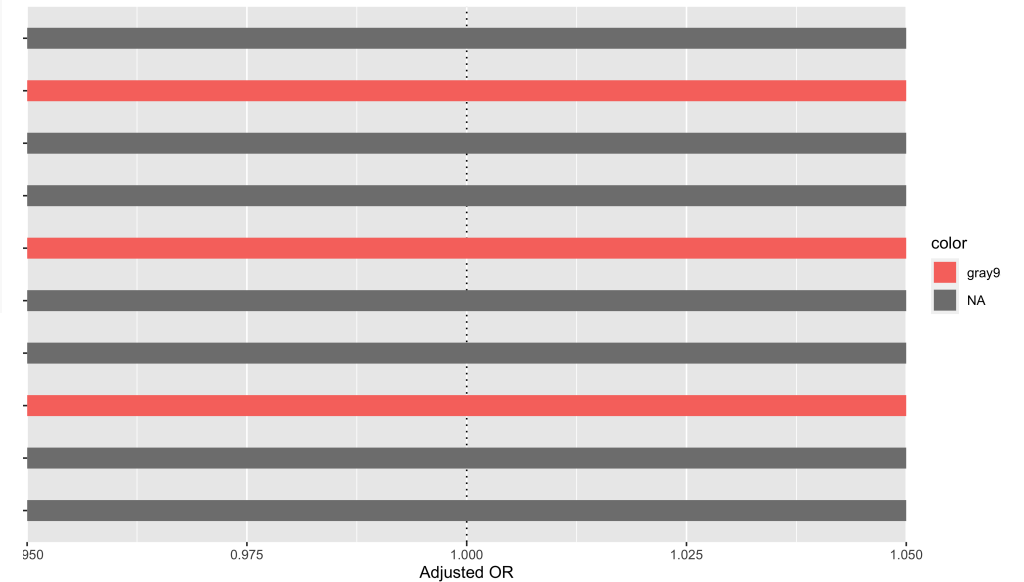
Plot - ggplot

```
#ggplot for plot  
ggplot(table_plot_df, aes(x = or, y = labels, xmin = or_low, xmax =  
  scale_y_discrete(limits = rev(table_plot_df$labels)) +  
  xlab("Adjusted OR") +  
  ylab("Adjusted OR with 95% Confidence Interval") +  
  geom_vline(xintercept = 1, linetype = 3) +  
  theme(axis.text.y = element_blank(),  
        axis.title.y = element_blank(),  
        plot.margin = margin(55, -10, 33, 0))
```



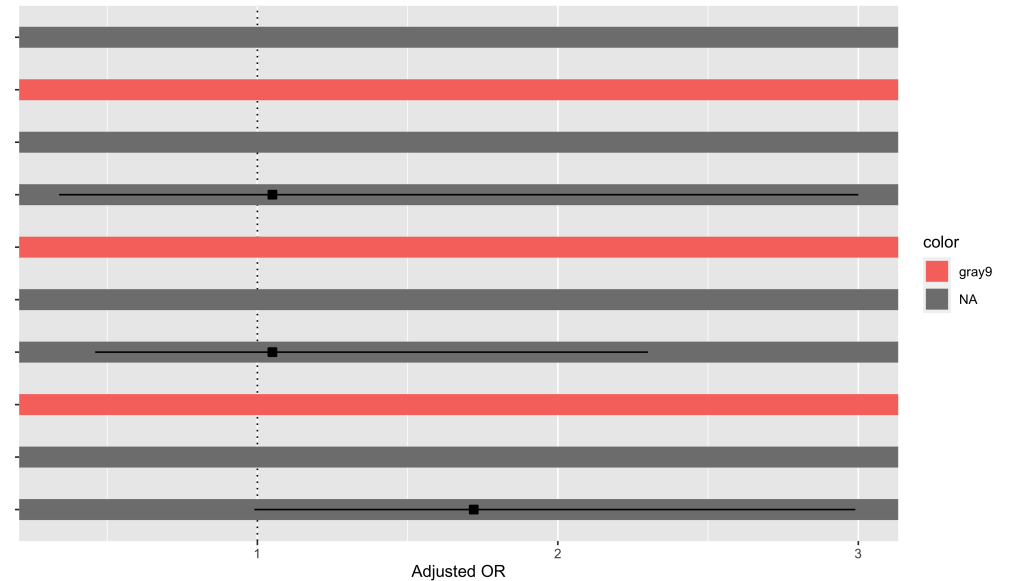
Plot - ggplot

```
#ggplot for plot  
ggplot(table_plot_df, aes(x = or, y = labels, xmin = or_low, xmax =  
  scale_y_discrete(limits = rev(table_plot_df$labels)) +  
  xlab("Adjusted OR") +  
  ylab("Adjusted OR with 95% Confidence Interval") +  
  geom_vline(xintercept = 1, linetype = 3) +  
  geom_hline(aes(yintercept = labels, colour = color), size = 6.5)  
  theme(axis.text.y = element_blank(),  
        axis.title.y = element_blank(),  
        plot.margin = margin(55, -10, 33, 0))
```



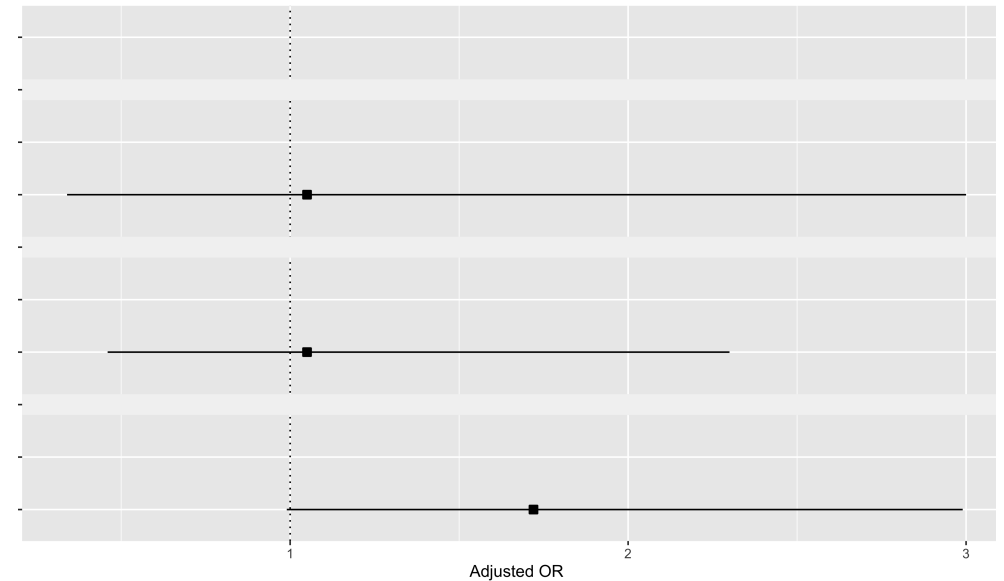
Plot - ggplot

```
#ggplot for plot  
ggplot(table_plot_df, aes(x = or, y = labels, xmin = or_low, xmax =  
  scale_y_discrete(limits = rev(table_plot_df$labels)) +  
  xlab("Adjusted OR") +  
  ylab("Adjusted OR with 95% Confidence Interval") +  
  geom_vline(xintercept = 1, linetype = 3) +  
  geom_hline(aes(yintercept = labels, colour = color), size = 6.5)  
  geom_pointrange(shape = 22, fill = "black") +  
  theme(axis.text.y = element_blank(),  
        axis.title.y = element_blank(),  
        plot.margin = margin(55, -10, 33, 0))
```



Plot - ggplot

```
#ggplot for plot  
ggplot(table_plot_df, aes(x = or, y = labels, xmin = or_low, xmax =  
  scale_y_discrete(limits = rev(table_plot_df$labels)) +  
  xlab("Adjusted OR") +  
  ylab("Adjusted OR with 95% Confidence Interval") +  
  geom_vline(xintercept = 1, linetype = 3) +  
  geom_hline(aes(yintercept = labels, colour = color), size = 6.5)  
  geom_pointrange(shape = 22, fill = "black") +  
  scale_colour_identity() +  
  theme(axis.text.y = element_blank(),  
        axis.title.y = element_blank(),  
        plot.margin = margin(55, -10, 33, 0))
```



Plot - ggplot

```
#ggplot for plot  
ggplot(table_plot_df, aes(x = or, y = labels, xmin = or_low, xmax =  
  scale_y_discrete(limits = rev(table_plot_df$labels)) +  
  xlab("Adjusted OR") +  
  ylab("Adjusted OR with 95% Confidence Interval") +  
  geom_vline(xintercept = 1, linetype = 3) +  
  geom_hline(aes(yintercept = labels, colour = color), size = 6.5)  
  geom_pointrange(shape = 22, fill = "black") +  
  scale_colour_identity() +  
  theme_classic() +  
  theme(axis.text.y = element_blank(),  
        axis.title.y = element_blank(),  
        plot.margin = margin(55, -10, 33, 0))
```

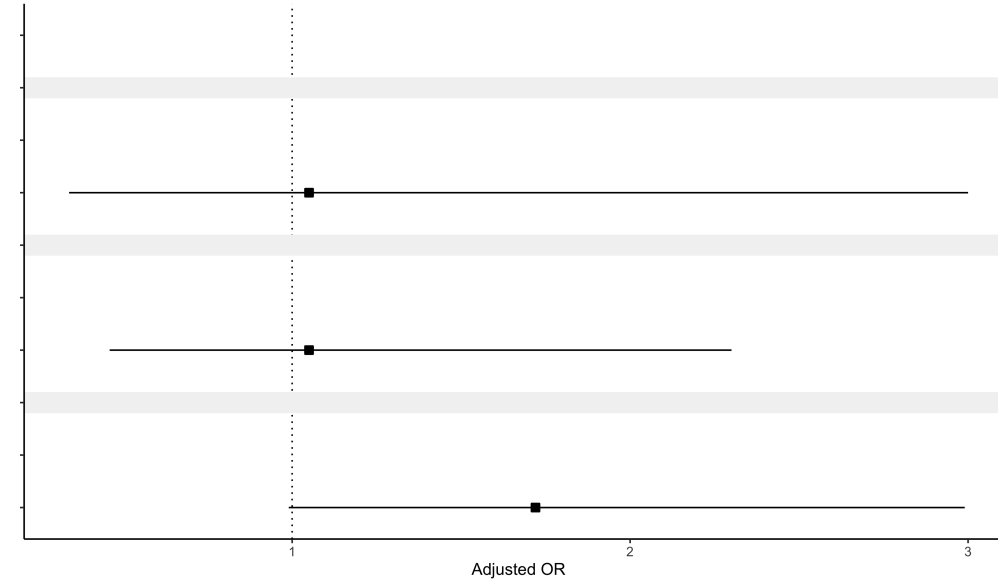


Table - ggplot

```
#ggplot for table  
ggplot(table_plot_df, aes(y = labels))
```

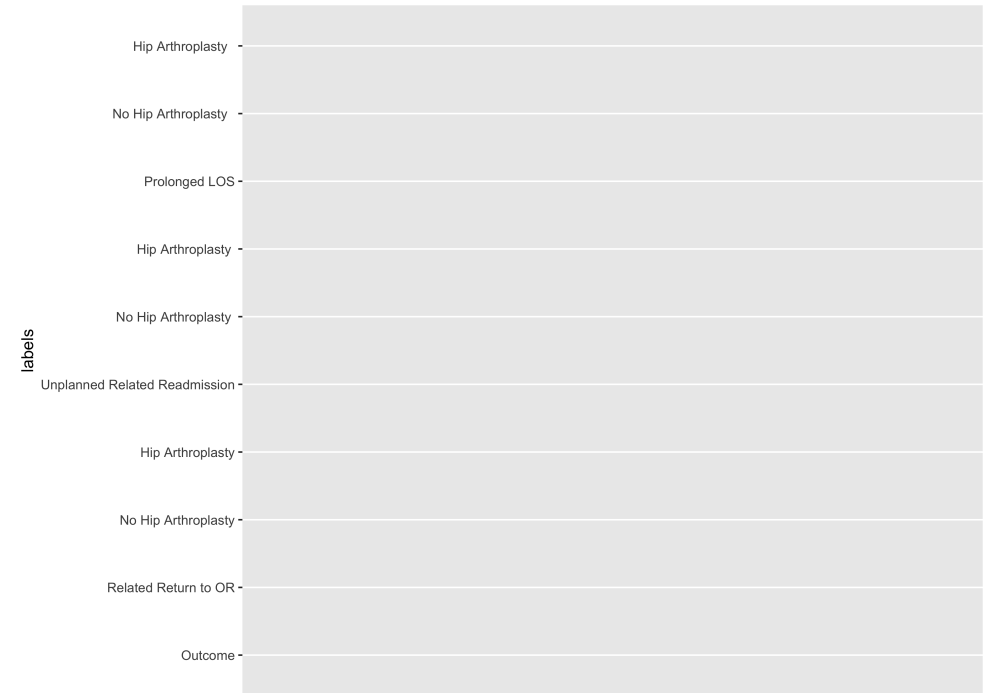


Table - ggplot

```
#ggplot for table  
ggplot(table_plot_df, aes(y = labels)) +  
  scale_y_discrete(limits = rev(table_plot_df$labels))
```

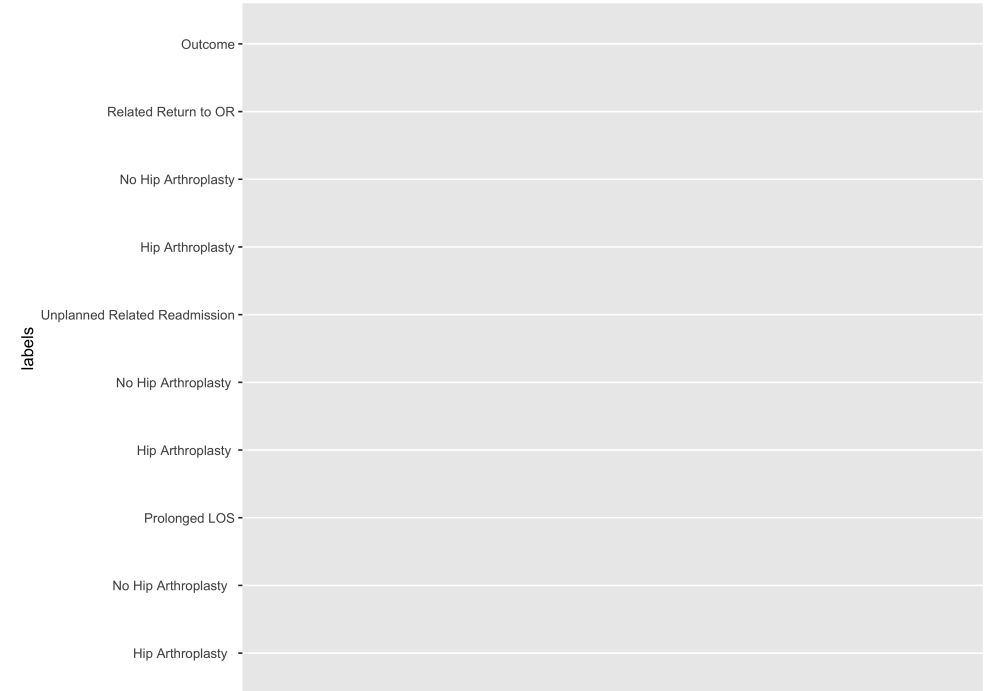


Table - ggplot

```
#ggplot for table  
ggplot(table_plot_df, aes(y = labels)) +  
  scale_y_discrete(limits = rev(table_plot_df$labels)) +  
  geom_text(aes(x = 0, label = labels), hjust = 0,  
             fontface = ifelse(table_plot_df$labels == "Outcome", "b", "n"),  
             size = ifelse(table_plot_df$labels == "Outcome", 3.5, 3))
```

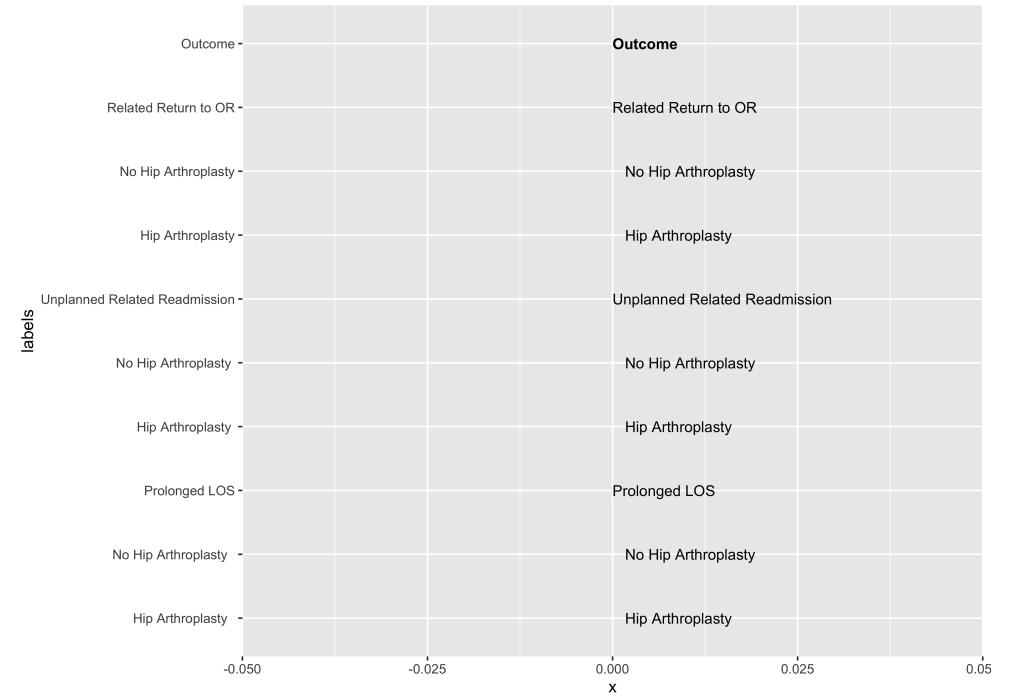


Table - ggplot

```
#ggplot for table  
ggplot(table_plot_df, aes(y = labels)) +  
  scale_y_discrete(limits = rev(table_plot_df$labels)) +  
  geom_text(aes(x = 0, label = labels), hjust = 0,  
            fontface = ifelse(table_plot_df$labels == "Outcome", "b",  
                               "n"),  
            size = ifelse(table_plot_df$labels == "Outcome", 3.5, 3),  
            color = "black") +  
  geom_text(aes(x = 4, label = eventnum),  
            fontface = ifelse(table_plot_df$eventnum == "events/N",  
                              "b", "n"),  
            size = ifelse(table_plot_df$eventnum == "events/N", 3.5,
```

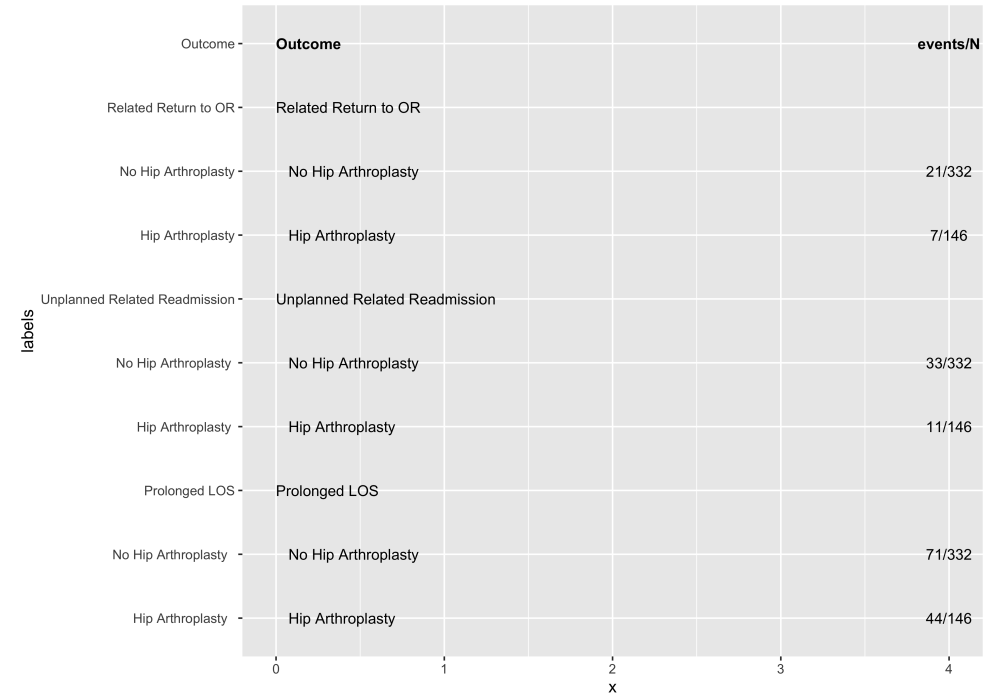


Table - ggplot

```
#ggplot for table
ggplot(table_plot_df, aes(y = labels)) +
  scale_y_discrete(limits = rev(table_plot_df$labels)) +
  geom_text(aes(x = 0, label = labels), hjust = 0,
            fontface = ifelse(table_plot_df$labels == "Outcome", "b",
                              "normal"),
            size = ifelse(table_plot_df$labels == "Outcome", 3.5, 3),
  geom_text(aes(x = 4, label = eventnum),
            fontface = ifelse(table_plot_df$eventnum == "events/N",
                              "normal", "b"),
            size = ifelse(table_plot_df$eventnum == "events/N", 3.5,
                          3),
  geom_text(aes(x = 7, label = or_ci), hjust = 1,
            fontface = ifelse(table_plot_df$or_ci == "aOR (95% CI)",
                              "normal", "b"),
            size = ifelse(table_plot_df$or_ci == "aOR (95% CI)", 3.5,
                          3)
```

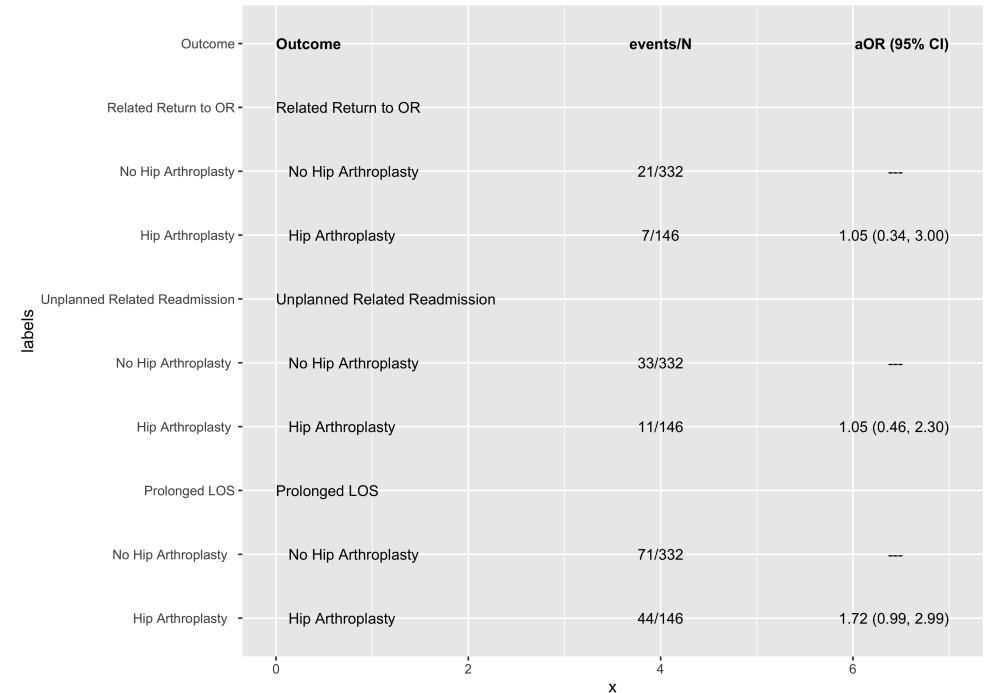


Table - ggplot

```
#ggplot for table
ggplot(table_plot_df, aes(y = labels)) +
  scale_y_discrete(limits = rev(table_plot_df$labels)) +
  geom_text(aes(x = 0, label = labels), hjust = 0,
            fontface = ifelse(table_plot_df$labels == "Outcome", "b",
                              "normal"),
            size = ifelse(table_plot_df$labels == "Outcome", 3.5, 3),
  geom_text(aes(x = 4, label = eventnum),
            fontface = ifelse(table_plot_df$eventnum == "events/N",
                              "normal", "b"),
            size = ifelse(table_plot_df$eventnum == "events/N", 3.5, 3),
  geom_text(aes(x = 7, label = or_ci), hjust = 1,
            fontface = ifelse(table_plot_df$or_ci == "aOR (95% CI)",
                              "normal", "b"),
            size = ifelse(table_plot_df$or_ci == "aOR (95% CI)", 3.5, 3),
  geom_hline(aes(yintercept = labels, colour = color), size = 6.5)
```

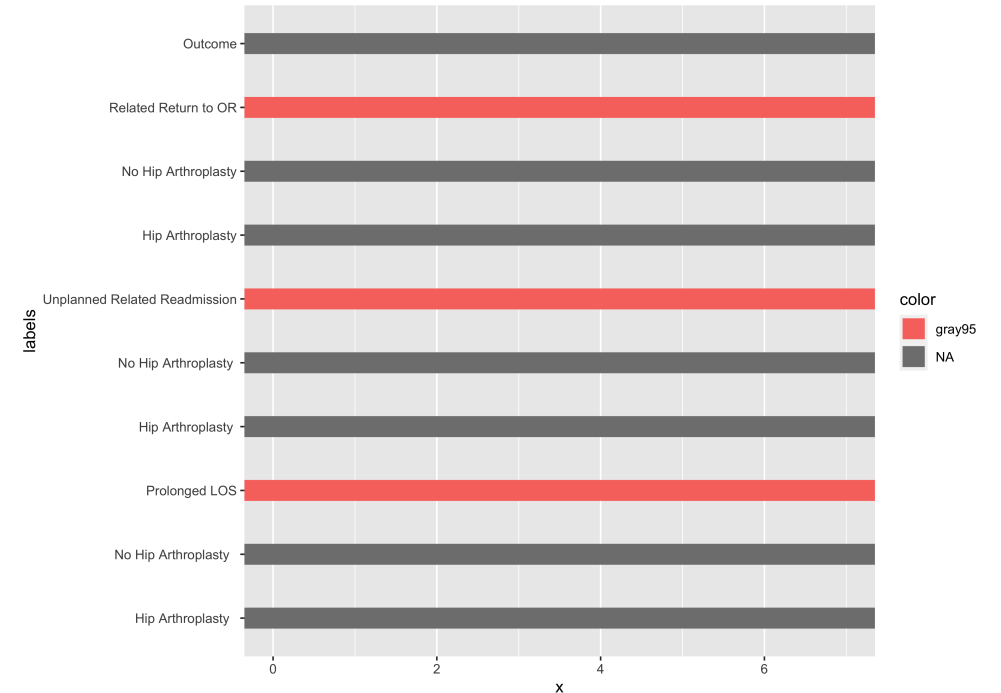


Table - ggplot

```
#ggplot for table
ggplot(table_plot_df, aes(y = labels)) +
  scale_y_discrete(limits = rev(table_plot_df$labels)) +
  geom_text(aes(x = 0, label = labels), hjust = 0,
            fontface = ifelse(table_plot_df$labels == "Outcome", "b",
                              "normal"),
            size = ifelse(table_plot_df$labels == "Outcome", 3.5, 3)) +
  geom_text(aes(x = 4, label = eventnum),
            fontface = ifelse(table_plot_df$eventnum == "events/N", "b",
                              "normal"),
            size = ifelse(table_plot_df$eventnum == "events/N", 3.5, 3)) +
  geom_text(aes(x = 7, label = or_ci), hjust = 1,
            fontface = ifelse(table_plot_df$or_ci == "aOR (95% CI)", "b",
                              "normal"),
            size = ifelse(table_plot_df$or_ci == "aOR (95% CI)", 3.5, 3)) +
  geom_hline(aes(yintercept = labels, colour = color), size = 6.5) +
  scale_colour_identity()
```

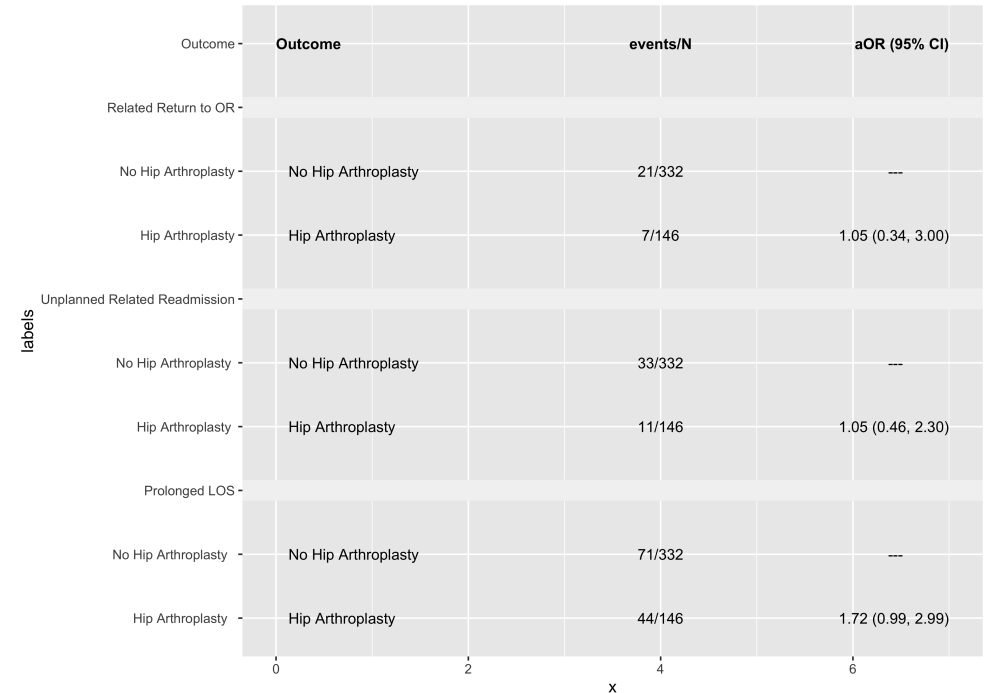


Table - ggplot

```
#ggplot for table
ggplot(table_plot_df, aes(y = labels)) +
  scale_y_discrete(limits = rev(table_plot_df$labels)) +
  geom_text(aes(x = 0, label = labels), hjust = 0,
            fontface = ifelse(table_plot_df$labels == "Outcome", "b",
                              "n"),
            size = ifelse(table_plot_df$labels == "Outcome", 3.5, 3),
  geom_text(aes(x = 4, label = eventnum),
            fontface = ifelse(table_plot_df$eventnum == "events/N",
                              "b", "n"),
            size = ifelse(table_plot_df$eventnum == "events/N", 3.5, 3),
  geom_text(aes(x = 7, label = or_ci), hjust = 1,
            fontface = ifelse(table_plot_df$or_ci == "aOR (95% CI)",
                              "b", "n"),
            size = ifelse(table_plot_df$or_ci == "aOR (95% CI)", 3.5, 3),
  geom_hline(aes(yintercept = labels, colour = color), size = 6.5) +
  scale_colour_identity() +
  theme_void()
```

Outcome	events/N	aOR (95% CI)
No Hip Arthroplasty	21/332	---
Hip Arthroplasty	7/146	1.05 (0.34, 3.00)
No Hip Arthroplasty	33/332	---
Hip Arthroplasty	11/146	1.05 (0.46, 2.30)
No Hip Arthroplasty	71/332	---
Hip Arthroplasty	44/146	1.72 (0.99, 2.99)

Table - ggplot

```
#ggplot for table
ggplot(table_plot_df, aes(y = labels)) +
  scale_y_discrete(limits = rev(table_plot_df$labels)) +
  geom_text(aes(x = 0, label = labels), hjust = 0,
            fontface = ifelse(table_plot_df$labels == "Outcome", "b",
                              "normal"),
            size = ifelse(table_plot_df$labels == "Outcome", 3.5, 3)) +
  geom_text(aes(x = 4, label = eventnum),
            fontface = ifelse(table_plot_df$eventnum == "events/N", "b",
                              "normal"),
            size = ifelse(table_plot_df$eventnum == "events/N", 3.5, 3)) +
  geom_text(aes(x = 7, label = or_ci), hjust = 1,
            fontface = ifelse(table_plot_df$or_ci == "aOR (95% CI)", "b",
                              "normal"),
            size = ifelse(table_plot_df$or_ci == "aOR (95% CI)", 3.5, 3)) +
  geom_hline(aes(yintercept = labels, colour = color), size = 6.5) +
  scale_colour_identity() +
  theme_void() +
  theme(plot.margin = margin(55, -10, 57, 0))
```

Outcome	events/N	aOR (95% CI)
No Hip Arthroplasty	21/332	---
Hip Arthroplasty	7/146	1.05 (0.34, 3.00)
No Hip Arthroplasty	33/332	---
Hip Arthroplasty	11/146	1.05 (0.46, 2.30)
No Hip Arthroplasty	71/332	---
Hip Arthroplasty	44/146	1.72 (0.99, 2.99)

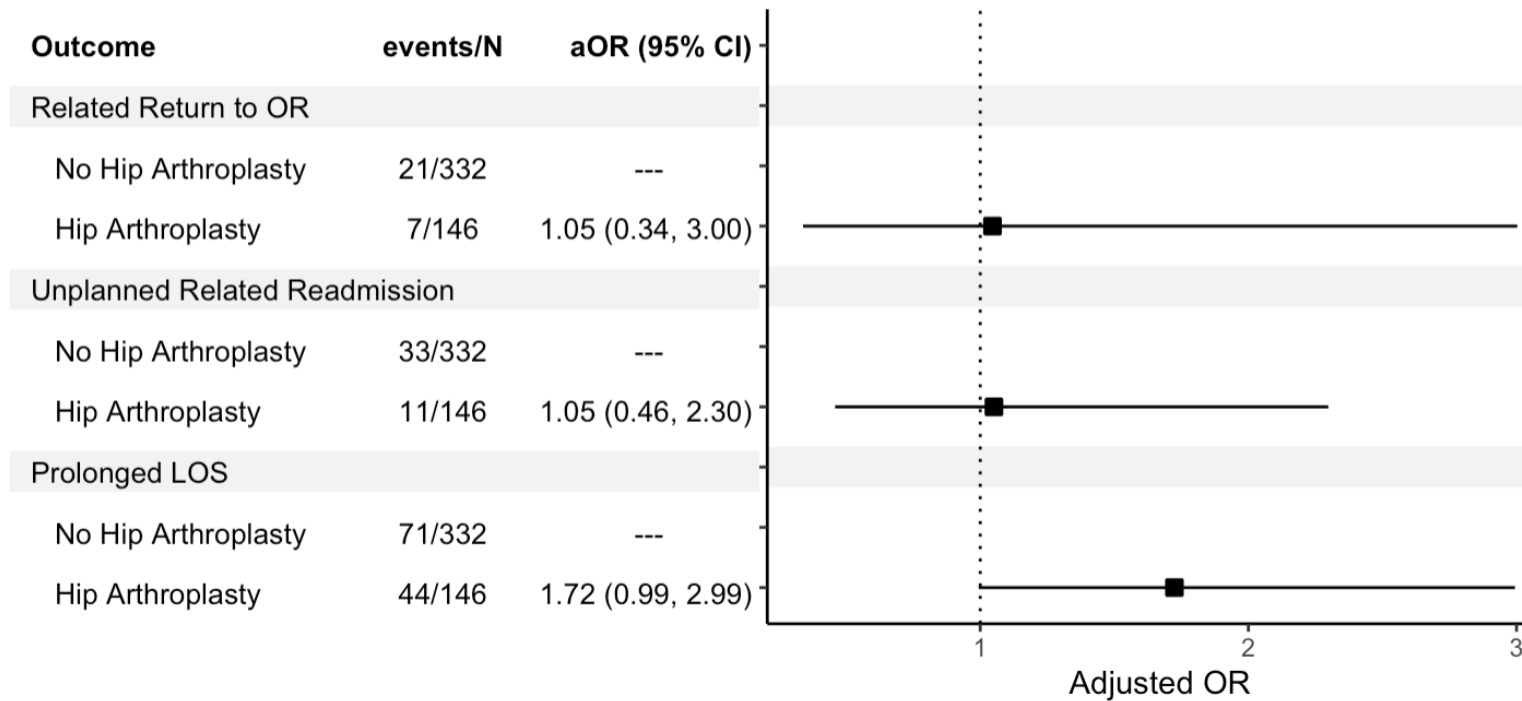
Table - ggplot

```
#ggplot for table
ggplot(table_plot_df, aes(y = labels)) +
  scale_y_discrete(limits = rev(table_plot_df$labels)) +
  geom_text(aes(x = 0, label = labels), hjust = 0,
            fontface = ifelse(table_plot_df$labels == "Outcome", "b",
                              "normal"),
            size = ifelse(table_plot_df$labels == "Outcome", 3.5, 3)) +
  geom_text(aes(x = 4, label = eventnum),
            fontface = ifelse(table_plot_df$eventnum == "events/N", "b",
                              "normal"),
            size = ifelse(table_plot_df$eventnum == "events/N", 3.5, 3)) +
  geom_text(aes(x = 7, label = or_ci), hjust = 1,
            fontface = ifelse(table_plot_df$or_ci == "aOR (95% CI)", "b",
                              "normal"),
            size = ifelse(table_plot_df$or_ci == "aOR (95% CI)", 3.5, 3)) +
  geom_hline(aes(yintercept = labels, colour = color), size = 6.5) +
  scale_colour_identity() +
  theme_void() +
  theme(plot.margin = margin(55, -10, 57, 0))
```

Outcome	events/N	aOR (95% CI)
No Hip Arthroplasty	21/332	---
Hip Arthroplasty	7/146	1.05 (0.34, 3.00)
No Hip Arthroplasty	33/332	---
Hip Arthroplasty	11/146	1.05 (0.46, 2.30)
No Hip Arthroplasty	71/332	---
Hip Arthroplasty	44/146	1.72 (0.99, 2.99)

Final Figure - ggplot

Use `grid.arrange` to combine the two ggplot objects



Notes

- You can customize the visual even further (e.g., if you wanted the estimates on the right side of the panel figure)
- It will probably take some adjustments to get the margins between the two figures to line up

Thank You!!! Questions?