Create an Efficacy Table with gt Package in R

Hengwei Liu, Hengrui USA, Princeton, New Jersey

ABSTRACT

In oncology studies, we often need to create an efficacy table for confirmed best overall response and present objective response rate (ORR) and clinical benefit rate (CBR).

In this paper we will use the gt package in R to create such a table.

INTRODUCTION

We will use two made-up datasets ADSL and ADRS. In the datasets there are three treatments, Group A, Group B and Group C. For each treatment we will present the count and percentage of patients who had confirmed best overall response CR, PR, SD, PD or NE.

We will also calculate the ORR and CBR with confidence intervals.

We will use the gt package in R to produce this table. See the reference [1] and [2] for more details about the gt package.

The datasets, R program and output in this paper are available in reference [3].

THE DETAILS

We will need five packages in the R program:

- use the haven package to read the SAS® datasets
- use the summarise function from the package dplyr to count patients
- use the pivot_wider function from the package tidyr to transpose a dataset
- use the binconf function from Hmisc package to calculate the proportion and confidence interval
- use the gt package to create the table in HTML format.

This is the R program efficacy.R.

```
library(haven)
library(dplyr)
library(tidyr)
library(gt)
library(Hmisc)

the_date <- as.character(Sys.Date())

# read the data
adsl <- read_sas("C:\\efficacy\\adsl.sas7bdat")
```

```
adrs <- read_sas("C:\\efficacy\\adrs.sas7bdat")


adsl$TRT01P <- gsub(" ", "", adsl$TRT01P)
adrs <- adrs[(adrs$RSEVAL=="Independent Central Review" &
adrs$PARAMCD=="CBRSP"), ]
adrs$TRT01P <- gsub(" ", "", adrs$TRT01P)


# get the big N in column headers from adsl
bign <- table(group=adsl$TRT01P)


# do the counting by TRT01P, AVALC


count0 <-
  adrs %>%
  group_by(TRT01P, AVALC) %>%
  summarise(unique_subj = n_distinct(USUBJID))


# generate a frame data called comb
# it has all the treatment group and all response values CR, PR, SD, PD, NE


mat <- matrix(NA, nrow = 5, ncol = 2)
xy <- data.frame(mat)
xy[[1,1]] <- data.frame(X1='CR')
xy[[2,1]] <- data.frame(X1='PR')
xy[[3,1]] <- data.frame(X1='SD')
xy[[4,1]] <- data.frame(X1='PD')
xy[[5,1]] <- data.frame(X1='NE')


for (i in seq(1,5)){
  xy[[i,2]] <- data.frame(X2=i)
}


for (i in seq(1:length(unique(count0$TRT01P)))){
  xy$X3 <- unique(count0$TRT01P)[i]
  assign(paste0('comb', i, sep=''),  xy)
}


comb <- do.call("rbind", mget(sprintf("comb%d",
1:length(unique(count0$TRT01P)))))
```

```
comb$AVALC <- unlist(comb$X1)
comb$TRT01P <- unlist(comb$X3)
comb$X2 <- unlist(comb$X2)

# merge comb with count0 and calculate percentage
m_count0 <- merge(count0, comb, by=c("TRT01P", "AVALC"), all=TRUE)
m_count0$denom <- ifelse(m_count0$TRT01P=='GroupA', bign[1],
ifelse(m_count0$TRT01P=='GroupB', bign[2], bign[3]))
m_count0$value <- ifelse(is.na(m_count0$unique_subj),"0",
paste(m_count0$unique_subj, "(",
format(round(100*m_count0$unique_subj/m_count0$denom, 1), nsmall = 1), ")"))

# do the transpose

a1 <- m_count0 %>%
 pivot_wider(id_cols=c(X2, AVALC), names_from = TRT01P, values_from = value,
                  names_prefix = "")

a1$catlabel <- a1$AVALC
a1$catlabel=ifelse(a1$AVALC=='CR','Complete Response (CR)', a1$catlabel)
a1$catlabel=ifelse(a1$AVALC=='PR','Partial Response (PR)', a1$catlabel)
a1$catlabel=ifelse(a1$AVALC=='SD','Stable Disease (SD)', a1$catlabel)
a1$catlabel=ifelse(a1$AVALC=='PD','Progressive Disease (PD)', a1$catlabel)
a1$catlabel=ifelse(a1$AVALC=='NE','Not Evaluable (NE)', a1$catlabel)
a1$block <- "Confirmed Best Overall Response"

a1 <- a1[order(a1$X2), ]


# do the ORR

mat <- matrix(NA, nrow = 1, ncol = 3)
z <- data.frame(mat)
z$GroupA <- z$X1
z$GroupB <- z$X2
z$GroupC <- z$X3

x <- table(group=adsl$TRT01P)
```

```
n <- table(group=adrs[(adrs$AVALC=='CR' | adrs$AVALC=='PR'),]$TRT01P)

# create a function to combine the proportion and confidence interval
getci <- function (grp){
  grp <- {{grp}}
  n[grp] <- ifelse(is.na(n[grp]), 0, n[grp])
  t <- binconf(n[grp], x[grp], method="exact")
  ci <- paste0(round(100*t[1], digits=1), ' (', round(100*t[2], digits=1), ',
', round(100*t[3], digits=1),')')
  return(ci)
}

z$GroupA <- getci("GroupA")
z$GroupB <- getci("GroupB")
z$GroupC <- getci("GroupC")
z$catlabel <- "ORR (95% CI) [a]"
z$block <- "Objective Response Rate"

# do the CBR

mat <- matrix(NA, nrow = 1, ncol = 3)
y <- data.frame(mat)
y$GroupA <- y$X1
y$GroupB <- y$X2
y$GroupC <- y$X3

x <- table(group=adsl$TRT01P)
n <- table(group=adrs[(adrs$AVALC=='CR' | adrs$AVALC=='PR' |
adrs$AVALC=="SD"),]$TRT01P)

# create a function to combine the proportion and confidence interval
getci <- function (grp){
  grp <- {{grp}}
  n[grp] <- ifelse(is.na(n[grp]), 0, n[grp])
  t <- binconf(n[grp], x[grp], method="exact")
  ci <- paste0(round(100*t[1], digits=1), ' (', round(100*t[2], digits=1), ',
', round(100*t[3], digits=1),')')
  return(ci)
}
```

```
y$GroupA <- getci("GroupA")
y$GroupB <- getci("GroupB")
y$GroupC <- getci("GroupC")
y$catlabel <- "CBR (95% CI) [a]"
y$block <- "Clinical Benefit Rate"




a1 <- a1[c("block","catlabel","GroupA", "GroupB", "GroupC")]
z <- z[c("block","catlabel","GroupA", "GroupB", "GroupC")]
y <- y[c("block","catlabel","GroupA", "GroupB", "GroupC")]
df <- rbind(a1, z, y)

df %>%
  gt(groupname_col="block")



# use gt to do the reporting
tab_html <- df %>%
  gt(groupname_col="block") %>%


  tab_header(
    title = "Table 14.2.1 Confirmed Best Overall Response based on BICR
assessment",
    subtitle = "ITT Population"
  ) %>%


  tab_source_note(
    source_note = "[a]: the confidence interval is based on Clopper-Pearson
method."
  ) %>%



  tab_source_note(
    source_note = paste('Program Source: efficacy.R            Executed:
(Draft)',  the_date)
  ) %>%
```

```
  cols_label(
    catlabel= " ",
    GroupA = paste0("Group A (N=", bign[1], ")"),
    GroupB = paste0("Group B (N=", bign[2], ")"),
    GroupC = paste0("Group C (N=", bign[3], ")")

  ) %>%

  tab_options(
    table.border.top.color = "white",
    heading.border.bottom.color = "black",
    table.border.bottom.color = "white",
    table_body.border.bottom.color = "black",
    table_body.hlines.color = "white",
    row_group.border.bottom.color = "white",
    row_group.border.top.color = "white",
    column_labels.border.top.color = "black",
    column_labels.border.bottom.color = "black",
  ) %>%

  cols_align(
    align = "left",
    columns = c(catlabel)
  )

# output the HTML table

tab_html %>%
  gtsave("efficacy.html", path = "C:\\efficacy" )
```

The table created with this R program is shown in display 1.

Table 14.2.1 Confirmed Best Overall Response based on BICR assessment

ITT Population

	Group A (N=17)	Group B (N=16)	Group C (N=17)
Confirmed Best Overall Response			
Complete Response (CR)	0	0	0
Partial Response (PR)	0	6 (37.5)	7 (41.2)
Stable Disease (SD)	11 (64.7)	3 (18.8)	3 (17.6)
Progressive Disease (PD)	3 (17.6)	4 (25.0)	3 (17.6)
Not Evaluable (NE)	3 (17.6)	3 (18.8)	4 (23.5)
Objective Response Rate			
ORR (95% CI) [a]	0 (0, 19.5)	37.5 (15.2, 64.6)	41.2 (18.4, 67.1)
Clinical Benefit Rate			
CBR (95% CI) [a]	64.7 (38.3, 85.8)	56.2 (29.9, 80.2)	58.8 (32.9, 81.6)

[a]: the confidence interval is based on Clopper-Pearson method.

Program Source: efficacy.R Executed: (Draft) 2022-03-16

Display 1. Confirmed Best Overall Response Table in HTML Format, Created with gt Package in R

CONCLUSION

The package gt is a great tool to create tables.

REFERENCES

[1] Some detailed discussion about gt package, available at:

https://aosmith16.github.io/spring-r-topics/slides/week04_gt_tables.html#1

[2] Presentation by Rich Iannone, available at:

https://www.youtube.com/watch?v=h1KAjSfSbmk&t=872s

[3] the datasets, R program and output in this paper are available at:

https://github.com/hengweiliu2020/efficacy-table-with-gt-package

CONTACT INFORMATION

Hengwei Liu
Hengrui USA
400 Alexander Park
Princeton, NJ 08540
Hengwei_liu@yahoo.com