

```

---
title: "speffanalysis"
author: "The R Bootcamp"
date: "10/8/2017"
output: html_document
---

```{r setup, include=FALSE, echo = FALSE}
knitr::opts_chunk$set(fig.width = 6,      # Figure width (in)
                      fig.height = 6,     # Figure height (in)
                      echo = FALSE,       # Repeat code
                      eval = TRUE,        # Evaluate chunks
                      message = FALSE,    # Don't print messages
                      warning = FALSE,    # Don't print warnings
                      fig.align = 'center') # Center figures

options(digits = 2) # Round all output to 2 digits
```

```{r}
# Loading Packages -----
library(tidyverse)
library(knitr)
library(speff2trial)
```

```{r}
# Load the ACT trial data
ACTG175 <- read_csv("1_Data/ACTG175.csv")
```

# Study
ACTG175 was a randomized clinical trial to compare monotherapy with *zidovudine* or
*didanosine* with combination therapy with zidovudine and didanosine or zidovudine and
*zalcitabine* in adults infected with the human immunodeficiency virus type I whose CD4 T
cell counts were between 200 and 500 per cubic millimeter.

# Data

In R, the data is stored as ACTG175. The data originally come from the `speff2trial`
package. However, for demonstration purposes, the analyses in this document are conducted
from a copy of the dataset saved as a text file ACTG175.csv on
[http://therbootcamp.github.io](http://therbootcamp.github.io). The analyses in this
document are based on this text file.

# Analyses

The full dataset contains data from `r nrow(ACTG175)` patients. For each patient, there
are `r ncol(ACTG175) - 1` observations (i.e.; columns in the data). Table 1 shows the
first 5 rows and first 4 columns of the full dataset:

```{r}
knitr::kable(ACTG175[1:5, 1:5], caption = "Table 1: First 5 rows and 5 columns of the full
ACTG175 dataset")
```

One of the primary measures was the number of days until a major nevasive event. Across
all patients, the median number of days was `r median(ACTG175$days)`. However, the results
did differ between treatment arms. Summary statistics of the number of days, separated by
each treatment arm, are presented in Table 2:

```{r, fig.cap="Summary statistics of the number of days until a major negative event for
different treatment arms."}
# Create Table 2

```

```
trial_summary <- ACTG175 %>%
  group_by(arms) %>%
  summarise(
    N = n(),
    Mean = mean(days),
    Median = median(days),
    SD = sd(days),
    Max = max(days)
  )

```

A plot showing the relationship between treatment arm and number of days until a major negative event are presented in Figure 2:

```
```{r}
# Boxplots
ggplot(data = ACTG175,
        mapping = aes(x = factor(arms), y = days)) +
  geom_boxplot() +
  labs(x = "Treatment Arm",
        y = "Number of days until a major negative event",
        title = "ACTG175",
        subtitle = "Created within an RMarkdown Document!",
        caption = "Source: speff2trial R package") +
  theme_bw()

```

```
```{r}
# Correlation test between cd40 and days

cd4_cor <- cor.test(formula = ~ cd40 + days,
                    data = ACTG175)

cd4_cor_r <- cd4_cor$estimate # Get the correlation
cd4_cor_p <- cd4_cor$p.value  # Get the p-value

```

The correlation between CD4 T cell count at baseline and number of days until a major negative event was  $r = \text{round}(cd4\_cor\_r, 2)$ ,  $p < .01$ .

```
```{r}
# Scatterplot
ggplot(data = ACTG175,
        mapping = aes(x = cd40, y = days)) +
  geom_point(alpha = .5) +
  labs(x = "CD4 T cell count at baseline",
        y = "Number of days until a major negative event",
        title = "ACTG175",
        subtitle = "Created within an RMarkdown Document!",
        caption = "Source: speff2trial R package") +
  theme_bw()

```

# Conclusions

Add your conclusions here! Include mini-chunks like ``min(ACTG175$days)`` which will return the value of ``r min(ACTG175$days)``