Clinical Drug Trial Report: Drug Safety Analysis

Tiffany Nwanne

2025-05-16

Executive Summary

This report presents an analysis of a clinical drug trial aimed at evaluating the safety profile of a candidate drug compared to a placebo. Key endpoints include the incidence of adverse effects and their severity.

I. Introduction

Background

Understanding the safety profile of new pharmaceuticals is essential before clinical approval. This study examines adverse events reported by participants in a randomized trial.

Objective

To compare the incidence and severity of adverse effects between participants who received the trial drug and those who received a placebo.

II. Methods

Study Design

This is a two-arm randomized controlled trial comparing a candidate drug with a placebo.

Participants

Participant-level data includes demographic variables (e.g., sex), treatment group assignment, and outcomes such as number and presence of adverse effects.

Endpoints

- **Primary:** Incidence of adverse effects (binary)
- **Secondary:** Number of reported adverse effects

Statistical Methods

- Descriptive statistics and visualization
- Chi-squared test for comparing proportions
- Mann-Whitney U test for severity comparisons

III. Results

1. Data Import

```
df <- read_csv("drug_safety.csv")
```

```
## Rows: 16103 Columns: 8
## ── Column specification ──────────────────────────────────────
## Delimiter: ","
## chr (3): sex, trx, adverse_effects
## dbl (5): age, week, wbc, rbc, num_effects
##
## ℹ Use `spec()` to retrieve the full column specification for this data.
## ℹ Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

2. Baseline Data Exploration

```
str(df)
```

```
## spc_tbl_ [16,103 × 8] (S3: spec_tbl_df/tbl_df/tbl/data.frame)
##  $ age            : num [1:16103] 62 62 62 62 62 62 62 62 69 69 ...
##  $ sex            : chr [1:16103] "male" "male" "male" "male" ...
##  $ trx            : chr [1:16103] "Drug" "Drug" "Drug" "Drug" ...
##  $ week           : num [1:16103] 0 1 12 16 2 20 4 8 0 1 ...
##  $ wbc            : num [1:16103] 7.3 NA 5.6 NA 6.6 ...
##  $ rbc            : num [1:16103] 5.1 NA 5 NA 5.1 ...
##  $ adverse_effects: chr [1:16103] "No" "No" "No" "No" ...
##  $ num_effects    : num [1:16103] 0 0 0 0 0 1 1 1 0 0 ...
##  - attr(*, "spec")=
##   .. cols(
##   ..   age = col_double(),
##   ..   sex = col_character(),
##   ..   trx = col_character(),
##   ..   week = col_double(),
##   ..   wbc = col_double(),
##   ..   rbc = col_double(),
##   ..   adverse_effects = col_character(),
##   ..   num_effects = col_double()
##   .. )
##  - attr(*, "problems")=<externalptr>
```

```
summary(df)
```

```
##       age              sex                 trx                 week
##  Min.   :39.00   Length:16103       Length:16103        Min.   : 0.000
##  1st Qu.:58.00   Class :character   Class :character    1st Qu.: 1.000
##  Median :65.00   Mode  :character   Mode  :character    Median : 4.000
##  Mean   :64.12                                          Mean   : 7.741
##  3rd Qu.:71.00                                          3rd Qu.:12.000
##  Max.   :84.00                                          Max.   :20.000
##
##       wbc              rbc          adverse_effects      num_effects
##  Min.   : 1.800   Min.   :2.100   Length:16103        Min.   :0.0000
##  1st Qu.: 6.000   1st Qu.:4.400   Class :character    1st Qu.:0.0000
##  Median : 7.100   Median :4.700   Mode  :character    Median :0.0000
##  Mean   : 7.341   Mean   :4.673                       Mean   :0.1016
##  3rd Qu.: 8.400   3rd Qu.:5.000                       3rd Qu.:0.0000
##  Max.   :26.500   Max.   :7.600                       Max.   :3.0000
##  NA's   :6975     NA's   :6976
```

```
head(df)
```

```
## # A tibble: 6 × 8
##     age sex   trx    week   wbc   rbc adverse_effects num_effects
##   <dbl> <chr> <chr> <dbl> <dbl> <dbl> <chr>                 <dbl>
## 1    62 male  Drug      0  7.30  5.10 No                        0
## 2    62 male  Drug      1 NA    NA    No                        0
## 3    62 male  Drug     12  5.60  5    No                        0
## 4    62 male  Drug     16 NA    NA    No                        0
## 5    62 male  Drug      2  6.60  5.10 No                        0
## 6    62 male  Drug     20 NA    NA    Yes                       1
```

```
colSums(is.na(df))
```

```
##             age             sex             trx            week             wbc
##               0               0               0               0            6975
##             rbc adverse_effects     num_effects
##            6976               0               0
```

3. Data Cleaning

```
df <- df |> select(-wbc, -rbc)

# Basic frequency tables
table(df$sex)
```

```
##
## female   male
##   3775  12328
```

```
table(df$trx)
```

```
##
##    Drug Placebo
##   10727    5376
```

```
table(df$adverse_effects)
```

```
##
##    No    Yes
## 14567   1536
```

```
# Create binary outcome
df$adverse_binary <- ifelse(df$adverse_effects == "Yes", 1, 0)
```

4. Primary Endpoint Analysis: Adverse Effect Rate

```
# Overall adverse rate
mean(df$adverse_binary)
```
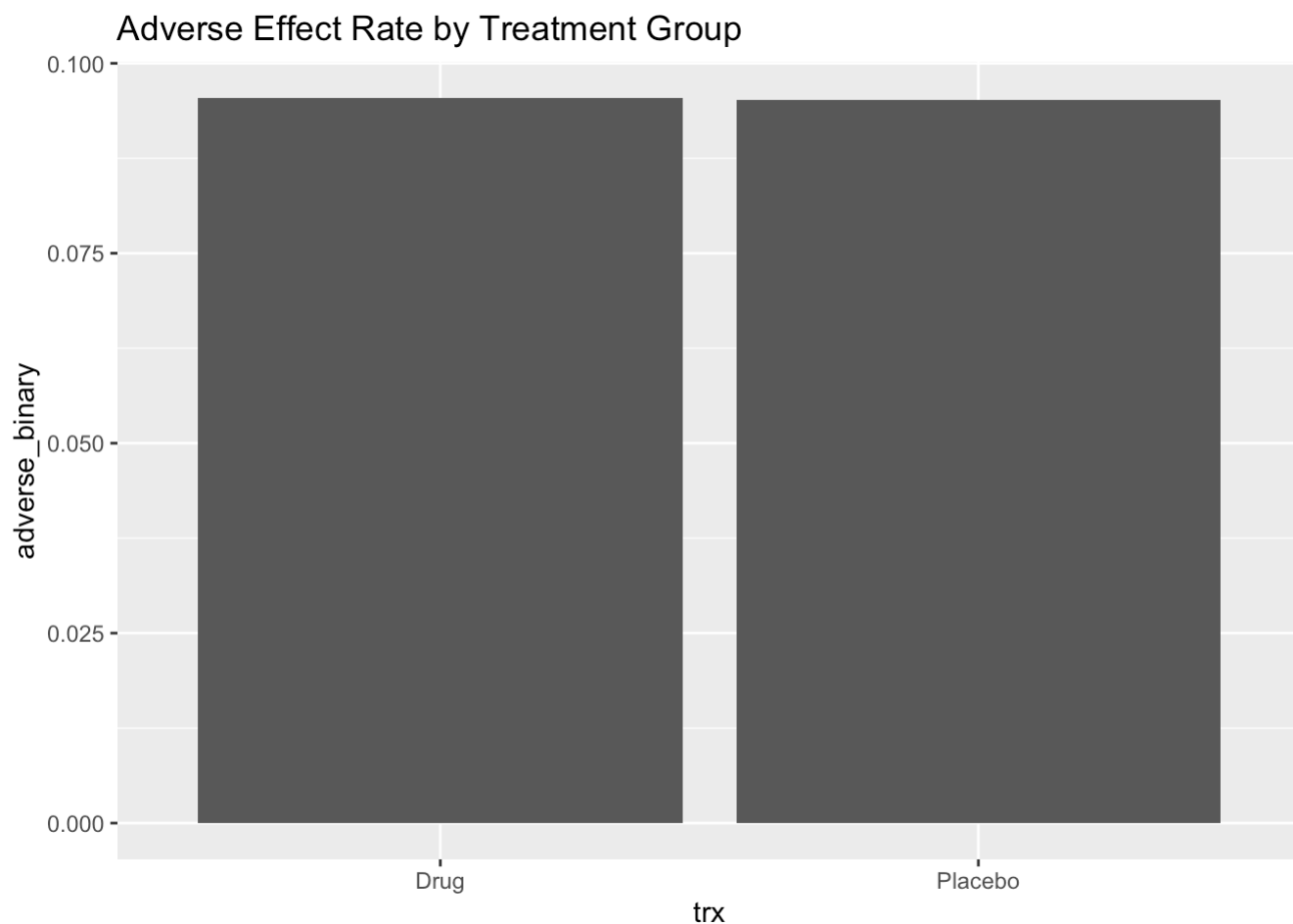
```
## [1] 0.09538595
```

```
# By treatment group
df |> group_by(trx) |> summarise(adverse_rate = mean(adverse_binary))
```

```
## # A tibble: 2 × 2
##   trx     adverse_rate
##   <chr>          <dbl>
## 1 Drug          0.0955
## 2 Placebo       0.0952
```

Figure 1: Adverse Effect Rate by Treatment Group

```
ggplot(df, aes(x = trx, y = adverse_binary)) +
  stat_summary(fun = mean, geom = "bar") +
  ggtitle("Adverse Effect Rate by Treatment Group")
```

Adverse Effect Rate by Treatment Group

5. Secondary Endpoint: Severity of Adverse Effects

Figure 2: Number of Adverse Effects by Group

```
ggplot(df, aes(x = trx, y = num_effects)) +
  geom_boxplot() +
  ggtitle("Number of Adverse Effects by Group")
```

Number of Adverse Effects by Group
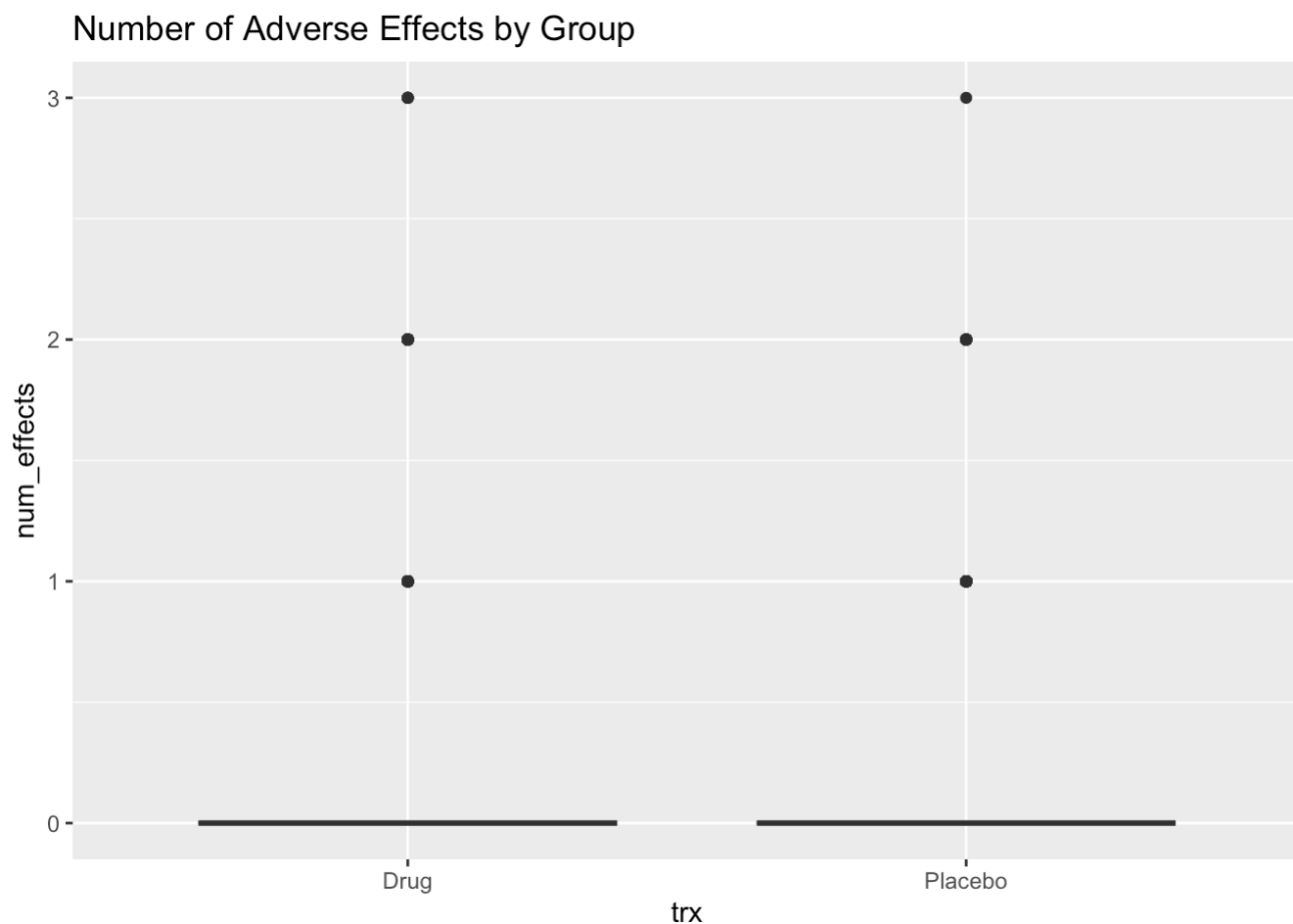
6. Statistical Testing

Chi-squared Test (Adverse Effect Incidence)

```
table1 <- table(df$trx, df$adverse_binary)
chisq.test(table1)
```

```
##
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  table1
## X-squared = 0.0002814, df = 1, p-value = 0.9866
```

Mann-Whitney U Test (Effect Severity)

```
drug <- df$num_effects[df$trx == "Drug"]
placebo <- df$num_effects[df$trx == "Placebo"]
wilcox.test(drug, placebo)
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  drug and placebo
## W = 28844703, p-value = 0.9407
## alternative hypothesis: true location shift is not equal to 0
```

IV. Discussion

The drug group showed a similar difference in adverse effect rate compared to the placebo group (9.55% vs. 9.52%). The Chi-square test ($\chi^2 \approx 0.00028$, $p \approx 0.987$) confirmed that there is no statistically significant association between treatment type and the occurrence of adverse effects.

Additionally, the Mann-Whitney U test ($U \approx 28{,}844{,}703$, $p \approx 0.941$) indicated no significant difference in the severity of adverse effects (measured by the number of effects reported) between the drug and placebo groups.

These findings suggest that the safety profile of the drug is comparable to that of the placebo with respect to both incidence and severity of adverse effects.

V. Conclusion

This analysis suggests that the drug is associated with a similar rate and severity of adverse effects compared to placebo. There is no evidence of increased safety concerns based on the current dataset. However, further trials involving larger sample sizes, diverse populations, and longer follow-up periods are warranted to confirm these findings and ensure a comprehensive assessment of the drug's safety.

VI. Appendices

Code: Saving Outputs

```
write.csv(df, "cleaned_drug_safety.csv", row.names = FALSE)
saveRDS(df, "cleaned_drug_safety.rds")

plot1 <- ggplot(df, aes(x = trx, y = adverse_binary)) +
  stat_summary(fun = mean, geom = "bar") +
  ggtitle("Adverse Effect Rate by Treatment Group")

ggsave("adverse_rate_by_group.png", plot = plot1, width = 6, height = 4, dpi = 300)
```

Session Information

```
sessionInfo()
```

```
## R version 4.5.0 (2025-04-11)
## Platform: x86_64-apple-darwin20
## Running under: macOS Monterey 12.7.4
##
## Matrix products: default
## BLAS:   /Library/Frameworks/R.framework/Versions/4.5-x86_64/Resources/lib/libRbla
s.0.dylib
## LAPACK: /Library/Frameworks/R.framework/Versions/4.5-x86_64/Resources/lib/libRlapa
ck.dylib;  LAPACK version 3.12.1
##
## locale:
## [1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8
##
## time zone: Africa/Kinshasa
## tzcode source: internal
##
## attached base packages:
## [1] stats     graphics  grDevices utils     datasets  methods   base
##
## other attached packages:
## [1] ggplot2_3.5.2 dplyr_1.1.4   readr_2.1.5
##
## loaded via a namespace (and not attached):
##  [1] bit_4.6.0          gtable_0.3.6       jsonlite_2.0.0     compiler_4.5.0
##  [5] crayon_1.5.3       tidyselect_1.2.1   parallel_4.5.0     jquerylib_0.1.4
##  [9] scales_1.4.0       yaml_2.3.10        fastmap_1.2.0      R6_2.6.1
## [13] labeling_0.4.3     generics_0.1.4     knitr_1.50         tibble_3.2.1
## [17] bslib_0.9.0        pillar_1.10.2      RColorBrewer_1.1-3 tzdb_0.5.0
## [21] rlang_1.1.6        utf8_1.2.5         cachem_1.1.0       xfun_0.52
## [25] sass_0.4.10        bit64_4.6.0-1      cli_3.6.5          withr_3.0.2
## [29] magrittr_2.0.3     digest_0.6.37      grid_4.5.0         vroom_1.6.5
## [33] hms_1.1.3          lifecycle_1.0.4    vctrs_0.6.5        evaluate_1.0.3
## [37] glue_1.8.0         farver_2.1.2       rmarkdown_2.29     tools_4.5.0
## [41] pkgconfig_2.0.3    htmltools_0.5.8.1
```