Phase III Clinical Trial Analysis

Bioinformatics Division

November 22, 2025

1 Statistical Analysis Protocol

We investigate the efficacy of Treatment A vs. Placebo. The analysis is performed using R version 4.5.1.

2 Data Ingestion and Summary

```r
# Simulate Clinical Data
set.seed(2024)
n <- 50
df <- data.frame(
  id = 1:(2*n),
  group = factor(rep(c("Placebo", "Treatment"), each=n)),
  recovery_days = c(rnorm(n, mean=14, sd=3), rnorm(n, mean=10, sd=2.5))
)

# Descriptive Statistics Table
summary_stats <- aggregate(recovery_days ~ group, data=df, FUN=function(x) c(Mean=mean(x), SD=sd(x))
summary_stats <- do.call(data.frame, summary_stats) # Flatten structure

# Print as LaTeX table using xtable
print(xtable(summary_stats, caption="Descriptive Statistics by Group"),
      booktabs=TRUE, include.rownames=FALSE)
```

group	recovery_days.Mean	recovery_days.SD
Placebo	13.77	3.19
Treatment	9.76	2.48

Table 1: Descriptive Statistics by Group

The table above is generated dynamically. If the underlying dataset `df` changes, the table updates automatically without manual re-formatting.
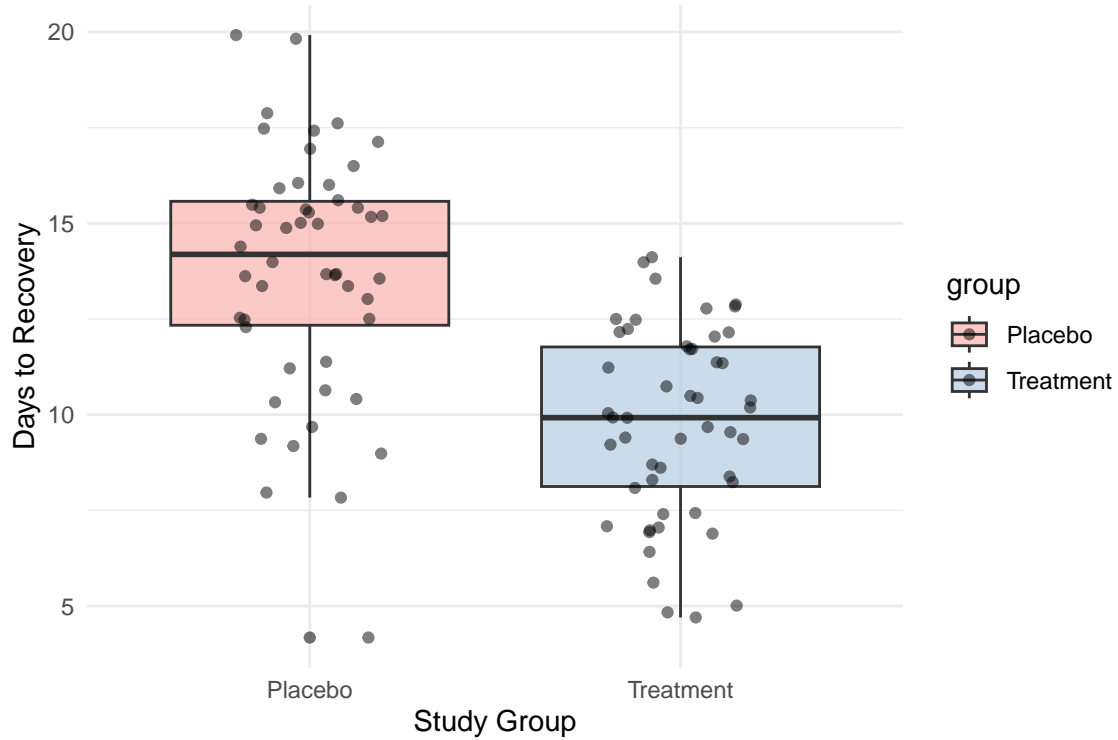
Figure 1: Distribution of Recovery Days by Treatment Group

3 Hypothesis Testing and Visualization

We perform an independent t-test to compare the groups.

The t-test yields a p-value of ¡ 0.001. This indicates a statistically significant difference.

```
ggplot(df, aes(x=group, y=recovery_days, fill=group)) +
  geom_boxplot(alpha=0.7) +
  geom_jitter(width=0.2, alpha=0.5) +
  theme_minimal() +
  labs(y="Days to Recovery", x="Study Group") +
  scale_fill_brewer(palette="Pastel1")
```