KU DEPARTMENT OF BIOSTATISTICS & DATA SCIENCE
The University of Kansas Medical Center

DRIVE: Daily Route Investigation Via an Efficient Bayesian Adaptive Trial Design

Statistical Analysis Plan

Table of Contents

1. Introduction

1.1. Background

This document describes the statistical methods that will be used for the analysis and reporting of data collected for the DRIVE trial. This includes the statistical models, decision rules, and simulation scenarios as input into the FACTS™ (Fixed and Adaptive Clinical Trial Simulator) software. A small set of operating characteristics for the simulations is also summarized.

1.2. Primary Endpoint

The primary endpoint is driving time in minutes continuous endpoint, measured at every study day. Lower driving times correspond to improvement.

1.3. Treatment Arms

The trial will involve up to a maximum of 33 drives, randomized among 3 arms, including a control arm and 2 comparator arms. We label these arms generically by their arm index as: $d = 0$ (Control), 1 (Rainbow), 2 (State Line).

2. Statistical Modeling

This section describes the statistical modeling used in the design. The modeling is Bayesian in nature.

2.1. Final Endpoint Model

Let Y_i be the primary outcome measured at each study day for the i^{th} drive. We model the outcomes as

$$Y_i \sim N\left(\theta_{d_i}, \sigma^2\right)$$

where θ_d is the mean response for arm d .

The mean response is modeled independently for each route as:

$$\theta_1 \sim N(15, 5^2),$$

$$\theta_2 \sim N(15, 5^2).$$

Thus, θ_d for each route is estimated separately using only data from that route.

The mean response for the control arm is modeled separately as

$$\theta_0 \sim N(15, 5^2).$$

The error variance is modeled as:

$$\sigma^2 \sim IG(0.05,1.25),$$

where IG(a, b) is the inverse gamma distribution defined by:

$$f(x|a, b) = \frac{b^a e^{-b/x}}{x^{a+1}\Gamma(a)}.$$

2.2. Evaluation of Posterior Estimates

The Bayesian final endpoint model is fitted to the data at each update. The posterior is calculated as:

$$p(\omega|Y) \propto \prod_{i=1}^{n} p\,(y_i|\varphi)p(\varphi)$$

where φ is the set of parameters for the final endpoint model, $p(\varphi)$ is the prior for those parameters, y_i is the final response for each drive, and n is the number of drives. The posterior is evaluated using MCMC with individual parameters updated by Metropolis Hastings (or Gibbs sampling where possible), using only the y_i data available at the time of the update.

2.3. Quantities of Interest

We define a number of quantities that will be tracked and may be used to make decisions during the trial.

2.3.1. Posterior Probabilities

For each route, we calculate the following quantities from the posterior:

- the probability that the mean response on route d is less than on Plaza:

$$Pr(\theta_d < \theta_0)$$

- the probability that the mean response on route d is less than on Plaza by at least -1:

$$Pr(\theta_d - \theta_0 < -1)$$

2.3.2. Target Routes

We consider the following target routes:

- The maximum effective route d_{max} is the route with the greatest treatment effect (difference from control). For each route, we calculate the probability of being the d_{max}:

$$Pr(Max).$$

2.3.3. Decision Quantities

The above quantities are computed for each arm (thus making them vector quantities). To facilitate decisions in the trial, we attach a particular arm to the quantity that will be used for the decision (thus reducing the vector quantity to a scalar quantity). Throughout the trial, decisions may be based on the following quantities:

- Maximum $Pr(\theta_d - \theta_0 < -1)$ across all routes
- $Pr(\theta_d < \theta_0)$ for $d =$ greatest $Pr(Max)$

2.4. Conventions for Missing Data

The primary endpoint will be evaluated using only completed drives. If a drive is randomized, but not completed, the assigned route will be completed at a later date prior to the next interim analysis.

2.5. Analysis Sets

2.5.1. Primary Set

All drives that were randomized and completed will be included in the primary set. This set will be used for the primary analysis.

2.5.2. Safety Set

All drives that were randomized, whether completed or not, will be included in the safety set. This set will be used for all adverse event summaries and for assessing the effect of route on the number of adverse events and serious adverse events.

3. Study Design

3.1. Timing of Interim Analyses

Interims will occur based on the number of drives with complete data.The schedule for the 3 interims is specified in Table-2.

Table-2: Interim Analysis Schedule

Interim	Drives (complete)
1	8
2	16
3	24

3.2. Response Adaptive Randomization

Up to the first interim analysis, drives will be randomized in blocks of size 4 with ratio 2:1:1 . After this initial burn-in period, adaptive randomization will begin, with the goal of preferentially allocating drives to the routes that appear more promising. Once adaptive allocation begins, drives will be randomized in blocks of size 4 .
Within each block, 1 drive will be allocated to route d_0. The remaining drives in each block will be allocated adaptively, with allocation probabilities weighted according to $Pr(Max)$ for each route. First, we calculate:
The response adaptive allocation uses the following weights:

$$V_d = [Pr(Max)]^1$$

The randomization probabilities for the adaptively allocated arms will be updated at each interim. They will be weighted according to the V_d and the weights will be renormalized to sum to 1.

To avoid assigning drives to a route with a minimal chance of being the best route, any probability less than 0.05 is set to zero at that interim and the resulting probability is reallocated among the remaining routes. In this manner, a route may be temporarily dropped but may be re-introduced if the adaptive randomization probability increases at subsequent interims.

3.3. Criteria for Stopping Accrual

3.3.1. Stopping for Futility

For interims 1-3, the trial may stop accrual for futility if any of the following criteria is satisfied:

$$\text{Maximum } Pr(\theta_d - \theta_0 < -1) \text{ across all routes} < 0.1$$

If a futility stopping rule is met at an interim analysis, then the final evaluation criteria will be applied to the currently available data.

3.3.2. Stopping for Expected Success

For interims 1-3, the trial may stop accrual for expected success if the following criteria is satisfied:

$$Pr(\theta_d < \theta_0) > 0.998 \text{ for } d = \text{ greatest } Pr(Max)$$

If a success stopping rule is met at an interim analysis, then a final analysis will be conducted using the currently available data.

3.4. Final Evaluation Criteria

No final futility criteria have been defined for this trial.

At the final analysis, the trial will be considered successful if the following criteria is satisfied:

$$Pr(\theta_d < \theta_0) > 0.9836 \text{ for } d = \text{ greatest } Pr(Max)$$

4. Simulation Scenarios

We evaluate the proposed design through trial simulation. We hypothesize several possible underlying truths for the mean response, as well as for trial execution variables such as accrual and dropout. For each of these scenarios, we generate data according to those truths and run through the design as specified above. We repeat this process to create multiple ``virtual trials" and we track the behavior of each trial. In this section, we describe the parameters used to generate the virtual drive-level data.

4.1. Virtual Drive Response Profiles

We consider 3 profiles for which drive outcomes for the final endpoint are simulated to have means as shown in Table-3 and standard deviations shown in Table-4.

Table-3: Virtual drive response means

VSR	d_0	d_1	d_2
Expected	15	13	13
Null	15	15	15
OneBest	15	13	14

Table-4: Virtual drive standard deviations

VSR	d_0	d_1	d_2
Expected	1.2	1.2	1.2
Null	1.2	1.2	1.2
OneBest	1.2	1.2	1.2

4.2. Accrual Profiles

We assume deterministic accrual of one drive per business day.

4.3. Dropout Profiles

We assume no dropouts for the purpose of this simulation.

5. Operating Characteristics

For the scenarios described above, we simulate multiple virtual trials and track the behavior of each trial, including the final outcome of the trial, the estimated mean response, etc. The results in this section are summarized across all simulated trials for each scenario.

5.1. Overall

This section gives a high-level description of the operating characteristics. Table-6 shows the following information per scenario:

- N sim: the number of simulated trials
- E[N]: the expected sample size
- Pr(success): the proportion of trials that met the final success criteria
- E[duration]: the expected duration of the trial in weeks.

Table-6: Overall Operating Characteristics

		Expected Allocation				
Effect	E[N]	Plaza (Control)	Rainbow	State Line	Pr(Success)	E[duration]
Expected	23.8	8.0	7.3	8.2	0.9796	4.6
Null (No Effect)	23.8	7.2	6.4	7.5	0.0238	4.6
One Best	25.3	8.6	11.1	6.4	0.9453	4.9

5.2. Trial Outcomes

This section summarizes the outcomes of the simulated trials. For each scenario in Table-7, the columns represent the proportion of simulated trials meeting each of the following definitions:

- Late Success (LS): enrolled to the maximum sample size and successful at the final analysis
- Late Futility (LF): enrolled to the maximum sample size and met the futility criteria at the final analysis
- Inconclusive (Inconc.): met neither the success nor the futility criteria at the final analysis

Table-7: Trial Outcomes

Effect	Early Success (ES)	Late Success (LS)	Early Futility (EF)	Inconc.
Expected	0.6801	0.2995	0.0010	0.0194

Effect	Early Success (ES)	Late Success (LS)	Early Futility (EF)	Inconc.
Null (No Effect)	0.0032	0.0206	0.5713	0.4049
OneBest	0.5864	0.3589	0.0060	0.0487

6. Computational Details

This report reflects the design parameters contained within the `ICD_KY.facts` file. The simulations were run using FACTS (Berry Consultants, LLC, Austin, TX) version 7.0.0.

7. Analysis Tables

1. Table 1

		Route			
	Category/ Statistic	Plaza (N=XX)	Rainbow (N=XX)	State Line (N=XX)	Overall (N=XX)
Start Time, n (%)	7:00-7:29AM	X (XX.X)	X (XX.X)	X (XX.X)	X (XX.X)
	7:30-7:59AM	X (XX.X)	X (XX.X)	X (XX.X)	X (XX.X)
	8:00-8:29AM	X (XX.X)	X (XX.X)	X (XX.X)	X (XX.X)
Day of Week, n (%)	Monday	X (XX.X)	X (XX.X)	X (XX.X)	X (XX.X)
	Tuesday	X (XX.X)	X (XX.X)	X (XX.X)	X (XX.X)
	Wednesday	X (XX.X)	X (XX.X)	X (XX.X)	X (XX.X)
	Thursday	X (XX.X)	X (XX.X)	X (XX.X)	X (XX.X)
	Friday	X (XX.X)	X (XX.X)	X (XX.X)	X (XX.X)
Time of Sunrise, n (%)	7:00-7:29AM	X (XX.X)	X (XX.X)	X (XX.X)	X (XX.X)
	7:30-7:59AM	X (XX.X)	X (XX.X)	X (XX.X)	X (XX.X)
	8:00-8:29AM	X (XX.X)	X (XX.X)	X (XX.X)	X (XX.X)
	8:30AM or later	X (XX.X)	X (XX.X)	X (XX.X)	X (XX.X)
Start time of First Meeting, n (%)	8:00 AM or earlier	X (XX.X)	X (XX.X)	X (XX.X)	X (XX.X)
	8:00-8:29AM	X (XX.X)	X (XX.X)	X (XX.X)	X (XX.X)
	8:30-8:59AM	X (XX.X)	X (XX.X)	X (XX.X)	X (XX.X)
	9:00 AM or later	X (XX.X)	X (XX.X)	X (XX.X)	X (XX.X)
Driving Time	Mean	XX.X	XX.X	XX.X	XX.X
	SD	X.XX	X.XX	X.XX	X.XX
	Median	XX.X	XX.X	XX.X	XX.X
	Range	XX.X, XX.X	XX.X, XX.X	XX.X, XX.X	XX.X, XX.X
Car, n (%)	GMC	X (XX.X)	X (XX.X)	X (XX.X)	X (XX.X)
	Toyota	X (XX.X)	X (XX.X)	X (XX.X)	X (XX.X)
	Honda	X (XX.X)	X (XX.X)	X (XX.X)	X (XX.X)
	Other	X (XX.X)	X (XX.X)	X (XX.X)	X (XX.X)

Weather Condition, n (%)	Clear	X (XX.X)	X (XX.X)	X (XX.X)	X (XX.X)
	Icy	X (XX.X)	X (XX.X)	X (XX.X)	X (XX.X)
	Snowy	X (XX.X)	X (XX.X)	X (XX.X)	X (XX.X)
	Wet	X (XX.X)	X (XX.X)	X (XX.X)	X (XX.X)
	Windy	X (XX.X)	X (XX.X)	X (XX.X)	X (XX.X)

2. Evaluation of Stopping Criteria at Interim Analysis 1

Parameter	Description	Estimate (95% Credible Interval)
θ_0	Mean time for Plaza (Control)	XX.X (XX.X, XX.X)
θ_1	Mean time for Rainbow	XX.X (XX.X, XX.X)
θ_2	Mean time for State Line	XX.X (XX.X, XX.X)
σ^2	Standard deviation for drive time	XX.X (XX.X, XX.X)
$\theta_1 - \theta_0$	Mean difference between Rainbow and Control	XX.X (XX.X, XX.X)
$\theta_2 - \theta_0$	Mean difference between State Line and Control	XX.X (XX.X, XX.X)
$Pr(\theta_d < \theta_0)$ for $d =$ greatest $Pr(Max)$	Probability of being better than control for the best route	XX.X
Max $Pr(\theta_d - \theta_0 < -1)$	Maximum probability of being better than control by one minute	XX.X
$Pr(\theta_d < \theta_0) > 0.998$ for $d =$ greatest $Pr(Max)$	Success Criteria	Yes/No
Max $Pr(\theta_d - \theta_0 < -1) < 0.1$	Futility Criteria	Yes/No

3. Evaluation of Stopping Criteria at Interim Analysis 2 (if applicable)

   Same as Table 1

4. Evaluation of Stopping Criteria at Interim Analysis 3 (if applicable)

   Same as Table 1

5. Final Analysis

Parameter	Description	Estimate (95% Credible Interval)
θ_0	Mean time for Plaza (Control)	XX.X (XX.X, XX.X)
θ_1	Mean time for Rainbow	XX.X (XX.X, XX.X)
θ_2	Mean time for State Line	XX.X (XX.X, XX.X)
σ	Standard deviation for drive time	XX.X (XX.X, XX.X)

$\theta_1 - \theta_0$	Mean difference between Rainbow and Control	XX.X (XX.X, XX.X)
$\theta_2 - \theta_0$	Mean difference between State Line and Control	XX.X (XX.X, XX.X)
$Pr(\theta_d < \theta_0)$ for $d =$ greatest $Pr(Max)$	Probability of being better than control for the best route	XX.X
$Pr(\theta_d < \theta_0) > 0.9836$ for $d =$ greatest $Pr(Max)$	Success Criteria	Yes/No

6. Adverse Events Summary

	Route			
	Plaza (N=XX)	Rainbow (N=XX)	State Line (N=XX)	Overall (N=XX)
Category	n (%)	n (%)	n (%)	n (%)
Total Number of Events	X (XX.X)	X (XX.X)	X (XX.X)	X (XX.X)
Category 1	X (XX.X)	X (XX.X)	X (XX.X)	X (XX.X)
Category 2	X (XX.X)	X (XX.X)	X (XX.X)	X (XX.X)
Etc.	…	…	…	…

7. Adverse Events by Category and Severity

		Route			
		Plaza (N=XX)	Rainbow (N=XX)	State Line (N=XX)	Overall (N=XX)
Category	Severity	n (%)	n (%)	n (%)	n (%)
Total Number of Events					
	Serious	X (XX.X)	X (XX.X)	X (XX.X)	X (XX.X)
	Non-Serious	X (XX.X)	X (XX.X)	X (XX.X)	X (XX.X)
Category 1					
	Serious	X (XX.X)	X (XX.X)	X (XX.X)	X (XX.X)
	Non-Serious	X (XX.X)	X (XX.X)	X (XX.X)	X (XX.X)
Category 2					
	Serious	X (XX.X)	X (XX.X)	X (XX.X)	X (XX.X)
	Non-Serious	X (XX.X)	X (XX.X)	X (XX.X)	X (XX.X)
Etc.					
	…	…	…	…	…

Secondary Analysis:

- Model the adverse events
- Take the final model and adjust for time of departure (and other covariates)