Pivotal Study of the Cerene Cryotherapy Device in Women with Heavy Menstrual Bleeding

Submitted to Channel Medsystems
By Berry Consultants
June 29, 2016

1.0 Introduction

This is a pivotal trial to evaluate the effectiveness of the Cerene device to treat heavy menstrual bleeding. This is a single arm trial that will enroll a maximum of 242 patients. The primary endpoint is the pictorial blood loss assessment chart (PBLAC) at 12 months after treatment. Patients who achieve a PBLAC score of <= 75 will be considered responders. The proportion of patients achieving a response will be compared against an objective performance criterion (OPC) of 66%. This trial will be monitored for early success.

2.0 Statistical Modeling

We use a longitudinal model that considers the PBLAC as continuous to allow the earlier observations of PBLAC to predict the final 12-month PBLAC value. The primary analysis is based on the dichotomized PBLAC score at 12 months where values less than or equal to a 75 indicate response.

2.1 Longitudinal Model

Patients will be assessed at 3, 6, and 12 months following treatment. We use a longitudinal model of the primary endpoint to allow these earlier observations to predict the final 12-month status. The primary endpoint for an individual patient, i, at time t is labeled as Y_{it}. At each interim analysis, some patients will have completed follow-up and we will have observed their 12-month endpoint Y_{i12}. These patients may also have earlier measures of the primary endpoint, such as at 3 and 6 months. There will be patients with earlier observations of the primary endpoint who have not yet reached 12 months of follow-up. We utilize the information from patients with incomplete follow-up to the extent that these earlier observations are correlated with the final 12-month value. A Bayesian model is built to learn how the early endpoints are correlated with the final 12-month endpoint and patients with complete follow-up information this model.

A linear regression model is created for the correlation between the baseline and 12-month, 3-month and 12-month, and 6-month and 12-month values. We suppress the subject index i and refer to the baseline, 3, 6, and 12 month values as Y_0, Y_3, Y_6,

and Y_{12}. There are three instances of the model. We model the log PBLAC score and each model instance is identical and has the following structure

$$\left[\log(Y_{12}) \mid \log(Y_0) \sim N\left(\alpha_0 + \beta_0 \log(Y_0), \lambda_0^2\right)\right]$$

$$[\alpha_0] \sim N\left(1, 2^2\right)$$
$$[\beta_0] \sim N\left(0.5, 1.5^2\right)$$
$$\left[\lambda_0^2\right] \sim IG\left(0.6, 1\right)$$

Identical models and prior distributions are assumed for the distribution of Y_{12} given Y_3 and for Y_{12} given Y_6

The joint posterior distribution of α_0, α_3, α_6, β_0, β_3, β_6, λ_0, λ_3, and λ_6, is updated based on all patients with observed values of Y_{12}. This model is then used (using Bayesian imputation within Markov chain Monte Carlo) to inform the primary analysis.

2.2 Primary Efficacy Analysis

Let I_{i12} be the response indicator for the i^{th} patient at 12 month. I_{i12} will be 1 where Y_{i12} is less than or equal to 75 and will be 0 where Y_{i12} is greater than 75. Let π = Pr(I_{i12} = 1) be the probability of response at 12 months. We model the log odds of response,

$$\theta = \log(\pi/1\text{-}\pi)$$

as normally distributed with prior a distributions for the log-odds of response

$$\theta \sim N(0, 2^2),$$

When converted from the log-odds scale back to the original probability scale, the resulting prior distribution for the log odds of response has a median of 50% and a 95% probability that the success rate is between 2% and 98%.

The longitudinal modeling will be used to inform the early success analyses at each planned interim analysis. However, consistent with the ITT population definition that the 66% OPC is based on, any patients with missing 12 month values for the primary endpoint will be considered non-responders in the primary efficacy analysis. Therefore, the longitudinal modeling will have no impact on the final primary efficacy analysis.

3.0 Allocation

A maximum of 242 patients will be enrolled and treated. This is a single arm trial and all enrolled patients will be treated with the Cerene device.

4.0 Evaluation of Trial Success

Interim analyses are planned when 100 and when 175 patients have been enrolled and treated and in the trial for 12 months, i.e. when the first 100 and 175 patients have had the opportunity to complete follow-up for the 12-month primary endpoint.

At an interim analysis, early success will be declared if there is a high probability that the response rate with the Cerene device is superior to the OPC of 66%. Formally if:

$$\Pr(\pi > 0.66) > 0.979.$$

A final determination of trial success will be applied when when all 242 patients have the opportunity to complete follow-up for the 12-month primary endpoint. Even if early success is declared, a final analysis will be conducted when all 242 patients have the opportunity to complete follow-up. If, at the completion of the trial, there is at least a 97.9% probability the Cerene device is superior to the OPC of 66%, this trial will be considered a success,

$$\Pr(\pi > 0.66) > 0.979.$$

5.0 Example Trials

In this section we present a single simulated trial to illustrate the statistical design of the trial.

When the first interim analysis is conducted, all 242 patients have been enrolled and treated and the first 100 patients have completed the 12-months of follow-up for the primary endpoint. We impute 12-month outcomes for the 142 patients still in follow-up based on their earlier visits. For all of these patients we have observed their 6-month visit and so we impute their 12-month outcome based on that most recent visit.

Among the 100 patients with complete follow-up, we have observed a 79% response rate at 6 months and an 80% response rate at 12 months. Among patients still in follow-up, we have observed an 84% response rate at 6 months. By the

longitudinal model, we estimate the 12-month response rate to be 77% (95% Interval = 63.2%, 87.4%). There is a 94% probability that the response rate is greater than the OPC of 66%. This is less than the 97.9% required for early success and the trial continues to the next interim analysis.

	Completers N = 100		Patients in Follow-up N = 142	
	Mean PBLAC Score (SD)	% Response	Mean PBLAC Score (SD)	% Response
Baseline	605 (1198)	25%	747 (1984)	25%
3 Months	80 (95)	66%	82 (100)	67%
6 Months	49 (53)	79%	44 (46)	84%
12 Months	54 (54)	80%	--	--

At the next interim analysis there is complete 12-month follow-up for 175 patients. We now observe a 12-month response rate of 78% among these patients. For the 67 patients still in follow-up, 88% are responders at 6-months. The estimated response rate at 12-months including all patients is 78% (95% Interval = 68.9%, 85.8%). There is a 99.7% probability that this response rate is greater than the OPC of 66%. This trial satisfies the early success criteria.

	Completers N = 175		Patients in Follow-up N = 67	
	Mean PBLAC Score (SD)	% Response	Mean PBLAC Score (SD)	% Response
Baseline	742 (1872)	24%	493 (817)	27%
3 Months	79 (94)	65%	89 (113)	71%
6 Months	49 (53)	80%	37 (30)	88%
12 Months	55 (57)	78%	--	--

At the final analysis of this trial, when all patients have complete follow-up through 12 months, the response rate is 78.4% (95% Interval = 73.1%, 83.5%) and there is a >99% probability that this response rate is greater than the OPC of 66%.

6.0 Simulation Scenarios

In order to characterize the performance of the trial design, we simulated the trial considering different scenarios for the proportion of patients achieving a response at 12 months and the longitudinal behavior of the primary endpoint.

6.1 Response Scenarios

We specify the scenarios for achieving a response at 12 months in terms of the mean and standard deviation of the PBLAC score at 12 months. Because we model the log of the PBLAC score in the longitudinal modeling, we also specify the scenarios for achieving a response at 12 months on the log scale. We consider 4 different standard deviations (SDs) for the log of the PBLAC score and then determine the mean that achieves a specific proportion of patients who would be considered responders. We consider SDs for PBLAC on the log scale of 1, 0.95, 0.75, and 0.5. Table 6.1 shows the means and SD on both the log PBLAC scale and PBLAC score scale that correspond to the indicted proportion of patients achieving a response.

Table 6.1 Response Scenarios: Mean and SD of the log PBLAC Score and Corresponding Proportion Responding

		Log PBLAC Scale				**PBLAC Scale**			
	Proportion Responding	SD = 1	SD = 0.95	SD = 0.75	SD = 0.5				
		Mean	Mean	Mean	Mean	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
1	60%	4.06	4.076	4.13	4.19	96 (126)	93 (112)	82 (72)	75 (40)
2	66%	3.905	3.925	4.009	4.111	82 (107)	80 (96)	73 (63)	69 (37)
3	70%	3.79	3.82	3.93	4.05	73 (94)	72 (87)	67 (59)	65 (35)
4	75%	3.64	3.676	3.81	3.98	63 (81)	62 (75)	60 (51)	61 (32)
5	80%	3.477	3.518	3.685	3.895	53 (70)	53 (64)	53 (46)	56 (30)
6	85%	3.281	3.333	3.539	3.799	44 (57)	44 (53)	46 (40)	51 (27)
7	90%	3.036	3.1	3.353	3.677	34 (45)	35 (42)	38 (33)	45 (24)

6.2 Longitudinal Profiles

We specify the longitudinal behavior of the PBLAC score by specifying the mean and SD of the log PBLAC score at each earlier visit as well as the correlation between the log PBLAC scores from visit to visit.

6.2.1 Response Fraction

We specify the mean and SD of the log PBLAC score at each earlier visit in terms of a multiplier on the final 12 month log PBLAC score. For example a multiplier of 1 would indicate the same mean PBLAC at an earlier visit as compared to the mean 12-month value and a multiplier of 2 would indicate a mean PBLAC score twice the mean 12-month value. In general we consider scenarios where baseline PBLAC scores are higher at baseline and where much of the reduction in the PBLAC score is achieved by the 3-month visit. For the variance of the log PBLAC at each visit, we consider scenarios where either there is a constant variance for the log PBLAC at each visit, or that the earlier values will have larger variability.

The table below shows the different simulation scenarios for the behavior of the mean and the variance of the log PBLAC at each visit.

Table 6.2.1: Multipliers for the Mean and Variance of the log PBLAC at Each Visit

	Baseline	3 Months	6 Months
Scenarios for Mean			
1. High Baseline and Early Response	1.48	1.07	0.96
2. Higher Baseline and Early Response	1.6	1.2	0.8
3. Higher Baseline and Late Response	1.6	1.2	1.05
Scenarios for Variance	1.5	1.12	1
1. Higher Variance Early	2	2	1.5
2. High Variance Early	1.5	1.12	1
3. Constant Variance	1	1	1

6.2.2 Visit-to-Visit Correlation

The values of the primary endpoint are correlated within each patient from visit to visit. The table below shows the correlations we assume for the log PBLAC score within a patient between each visit.

Table 6.2.2: Correlation Scenarios

	Weaker Correlation	**Stronger Correlation**
Baseline to 3 Months	0.1	0.35
3 Months to 6 Months	0.8	0.76
6 Months to 12 Months	0.3	0.6

6.3 Accrual and Dropout

We consider an accrual scenario in which accrual ramps up linearly for 12 weeks to a peak accrual rate of 11 patients per week. All patients will have been enrolled between 6 and 8 months of total trial time. We simulate no drop-outs, 5%, 10%, and 15% drop-out. In the primary analysis, any patient we does not have a 12-month value will be considered a failure.

7.0 Operating Characteristics

Table 7.1 shows the operating characteristics of the trial design. Results are based on at least 5000 simulations for each scenario. We show the probability of achieving early success and the probability the trial does not achieve an early success but claims success at trial completion (Probability of late success). The total probability of success is the sum of achieving a success, either early or late. All patients are expected to be treated at the time the first interim analysis occurs so the trial's mean sample size is 242 patients.

The operating characteristics presented here are based on the response scenarios in which the SD of the log PBLAC is 0.95. The longitudinal profile of the virtual patients is based on scenarios with a high baseline score and more patients achieving a response at 6 months than at 12 months (Table 6.2.1 mean scenario 1), a high variance for the PBLAC at earlier timepoints (Table 6.6.1 variance scenario 2), and the stronger visit-to-visit correlation. Based on patient-level data from a previous trial of the Cerene device, this is the expected profile of patients who will be enrolled and treated in this trial. We show results across the various assumptions for the proportion of patients dropping-out before 12 months. Operating characteristics based on other simulation scenarios are shown in the appendix.

Table 7.1: Operating Characteristics				
Response Among 12-Month Completers	**ITT Response**	**Pr (Early Success)**	**Pr (Late Success)**	**Total Pr (Success)**
No Drop-Outs				
60%	60.0%	0	0.000	0.000
****66%**	66.0%	0.000	0.025	0.025
70%	70.0%	0.015	0.230	0.245
75%	75.0%	0.295	0.571	0.866
80%	80.0%	0.900	0.098	0.998
85%	85.0%	0.999	0.001	1.000
90%	90.0%	1.000	0.000	1.000
5% Drop-Out				
60%	57.0%	0	0	0
66%	62.7%	0	0.001	0.001
70%	66.5%	0.001	0.038	0.039
75%	71.3%	0.044	0.357	0.401
80%	76.0%	0.483	0.441	0.924
85%	80.8%	0.966	0.034	0.999
90%	85.5%	0.999	0.000	0.999
10% Drop-Out				
60%	54.0%	0	0	0
66%	59.4%	0	0	0
70%	63.0%	0	0.001	0.001
75%	67.5%	0.003	0.067	0.070
80%	72.0%	0.105	0.391	0.497
85%	76.5%	0.643	0.303	0.947
90%	81.0%	0.988	0.012	1.000

**10,000 simulations

If the response rate is less than the OPC, 60%, there is a 0% probability of trial success. Under the null hypothesis, that response rate is the same as the OPC, 66%, this trial will not achieve early success, but will achieve success with complete follow-up on all patients with 2.5% probability. This is the simulated one-sided Type I error rate in this scenario. Given that the final evaluation of trial success is based on the dichotomous responses endpoint with complete data among all enroll patients, and that the longitudinal model has no role in the primary efficacy analysis, the overall Type I error rate does not depend on the mean or standard deviation of the PBLAC score, accrual rates, or longitudinal profile assumptions. The simulated Type I error rate here is shown to be controlled to < 2.5% with a critical value of

97.9%, but given there is a single final evaluation of trial success based on complete data, a critical value of 97.5% could also be considered appropriate.

Under an alternative hypothesis, that the response rate is 75%, this trial will stop early for success with approximately 30% probability. The total probability of trial success in this scenario is 87%. This is the power of the trial in this scenario. As the drop-out rate increases, the observed response rate at 12 month decreases and patients who are missing their 12 month observation are counted as non-responders. Correspondingly, the power of the trial decreases. If 5% of patients drop-out and there is an 80% response rate among completers at 12 months, the response rate in the ITT population will be 76% and there is a 92.4% probability of trial success. If 10% of patients drop-out and there is an 85% response rate among completers at 12 months, the response rate in the ITT population will be 76.5% and there is a 94.7% probability of trial success.