Galil Medical

Multicenter Study of Cryoablation for Palliation of Painful Bone Metastases (MOTION)
(CGC15-BNE098)

Statistical Analysis Plan (SAP)

Table of Contents

Protocol Summary

Objective

The objective of this study is to assess the efficacy of cryoablation for palliation of painful metastases in patients with metastatic lesions involving bone who have failed, are not candidates for, or are not experiencing adequate pain relief from current pain therapies (e.g. radiation, analgesics).

Study Design

This is a pivotal, multicenter, prospective, single arm study with subjects serving as their own control. This study is to enroll and treat 60 subjects at approximately 10 centers in the US and internationally.

Patients with painful metastatic lesions involving bone who meet the eligibility criteria and who have been determined to be an appropriate candidate for cryoablation therapy will be offered enrollment into the study. Patients agreeing to participate will read and sign an informed consent form and thus become subjects in the study. Treatment will be performed using a Galil Medical cryoablation system and Galil Medical cryoablation needles. Subjects will have one cryoablation procedure and will be followed for up to 6 months for palliation of pain, quality of life and analgesic usage.

Study visits include a screening visit, a cryoablation procedure visit and follow-up visits at 1, 4, 8, 12, 16, 20, and 24 weeks post cryoablation procedure.

Sample Size Justification

A 2 point reduction from baseline in worst pain score is considered clinically meaningful. The expected device performance is at least a 3.2 point reduction from baseline in mean worst pain score and the expected standard deviation is 2.7. With these assumptions, PASS 13 software was used to compute a sample size for a one-sample T-test with superiority by a margin:

> A sample size of 42 achieves 80% power to detect superiority using a one-sided t-test when the margin of superiority is -2.0 and the true difference between the mean and the reference value is -3.2. The data are drawn from a single population with a standard deviation of 2.7. The significance level (alpha) of the test is 0.025.

Galil Medical completed a pilot study for the same indication. In this pilot study, 28 subjects were enrolled and approximately 82% (23/28) of subjects completed the 7 week follow up interval. Nearly all of the attrition (10/11 early withdrawals) observed in the pilot study was attributed to death due to disease progression. This is similar to the population accrued for the Callstrom et al study with the same duration of follow up 57% (35/61) completed 7 weeks. In the Goetz et al (RFA) multicenter study, 77% (33/43) completed the study at 24 weeks. The comparative analysis of study participation summarized the attrition for each of these studies in which ablation was used to treat metastatic bone lesions. The significant attrition between 7 and 12 weeks follow up should be noted across all studies and is not unanticipated due to this population with metastatic disease. Furthermore, the 30% attrition rate at 8 weeks for this

terminally ill population was acknowledged by FDA reviewers in preparation for the launch of this study.

The final study sample size accounts for 30% attrition to ensure the appropriate number of evaluable subjects at 8 weeks. Additionally, sensitivity analyses will be conducted using methods attempt to account for drop-out subjects. Missing at random (MAR) approach is assumed as the pilot data shows nearly all attrition will be due to subject death from disease progression unrelated to cryoablation treatment. We do not anticipate any systematic differences between those who do and do not drop out.

With an attrition rate at 8-weeks estimated to be 30%, the total treated study sample size will be 60 subjects. This sample size will be considered clinically sufficient to give indications of device effectiveness.

General Analysis Methods

Data will be analyzed using descriptive statistics. Continuous variables (e.g., age) will be summarized by the number of patients, mean, standard deviation, median, interquartile range, minimum and maximum. Categorical variables (e.g., race) will be summarized by frequencies and percentages of patients in each category.

Analysis Populations

Intent-to-Treat (ITT) analyses will be performed on all subjects with a Galil Medical cryoablation system as the attempted treatment. Endpoints will also be evaluated using per protocol population (PP).

Efficacy Endpoint Analysis

Primary Efficacy Endpoint

The primary endpoint for this study will be measured as follows: improvement in self-reported pain scores defined by ≥ 2 point reduction in worst pain in the last 24 hours using the Brief Pain Inventory (BPI) from baseline to 8 weeks. A mean difference of 2 points reduction is considered clinically significant.

Analysis Methods

Pain assessments using the BPI will be made by examining the average difference of pre- and post-treatment worst pain in 24 hours at week 8 follow-up as measured on the Brief Pain Inventory (BPI) scale. A mean pain reduction of 2 points from baseline in worst pain score is considered clinically meaningful.

Calculations:

- 8-Week follow-up minus baseline worst pain will be calculated for each subject.
- Mean change over all subjects will be calculated.

- The mean reduction in pain at 8-weeks will be evaluated for clinical meaning.

To be considered clinically meaningful the 95% confidence interval's upper bound on the mean reduction in worst pain score must be less than -2.

Subjects without missing endpoint data will be considered completers. Intent-to-Treat Completers (ITTC) analysis as well as Per-Protocol Completers (PPC) analysis will be supplemented by missing data sensitivity analyses. Sensitivity analyses will be performed on the primary efficacy endpoint using two methods of missing data imputation: multiple imputation and last observation carry forward. These imputation methods will be employed on both the ITT and PP populations resulting two additional outcome analyses: Intent-to-Treat Imputed (ITTI) and Per-Protocol Imputed (PPI). Of the four analyses, the primary hypothesis test will be considered the ITTI analysis with multiple imputation.

For any complete case analysis, the subjects missing BPI at week 8 would be excluded from the analysis. For any imputed analysis, the subjects missing BPI at week 8 would use the imputed week 8 value to determine if the responder threshold was met for that subject.

The multiple imputation model will consider the BPI score at each time point (BPI at 1, 4, 12, 16, 20, and 24 weeks will be part of analyses) in addition to age and gender.

Safety Endpoint Analysis

Safety Endpoint

The safety endpoint for this study is to assess the incidence and severity of intraoperative events, post-operative adverse events, serious adverse events and unanticipated adverse device effects.

Analysis Methods

Point estimates and two-sided Exact 95% confidence intervals will be generated for procedural related events in the following categories:

- intra-operative events
- post-operative adverse events
- serious adverse events
- unanticipated adverse device effects

The safety analysis will be performed on the per-protocol population. No formal statistical hypothesis testing will be performed for safety endpoints.

Other Endpoint Analyses

Other Endpoints

Other endpoints for this study will include, but are not limited to:

- A responder analysis will be conducted. A responder is defined as a subject having ≥ 2 point reduction in the worst pain score in the last 24 hours using the BPI and no more than 25% increase in medication use (morphine equivalent) from baseline to 8 weeks
- Quality of Life (QoL) as measured by the overall BPI pain interference average score from baseline to 1, 4, 8, 12, 16, 20, and 24 weeks after cryoablation
- Change in physical function as measured by Karnofsky Performance Status (KPS) from baseline to 1, 4, 8, 12, 16, 20, and 24 weeks after cryoablation
- Additional therapies for persistent/recurrent pain associated with the index tumor under study or new metastases through 24 weeks
- Change in analgesic medications (morphine equivalent daily dosing and NSAID use) from baseline to 1, 4, 8, 12, 16, 20, and 24 weeks after cryoablation
- Difference in worst pain scores from baseline to 1, 4, 12, 16, 20, and 24 weeks after cryoablation as measured on the numeric 0 to 10 BPI scale
- Difference in average pain scores from baseline to 1, 4, 8, 12, 16, 20, and 24 weeks after cryoablation as measured on the BPI numeric rating scale
- Self-assessed Overall Treatment Effect (OTE) at 1, 4, 8, 12, 16, 20, and 24 weeks after cryoablation

Tumor Characteristics

Tumor characteristics will include, but are not limited to:

- Histology of primary cancer
- Proof of metastases
- Location of the index tumor
- Size of the index tumor

Physical Function

Physical function is considered routine clinical practice in medicine. Specifically, in oncology practice, physical function assessments such as the Karnofsky scale have been incorporated. This study will use the Karnofsky scale to assess physical function.

Additional Data Summaries

Demographics, disease status, labs, procedural information, and protocol deviations will be assessed by mean, standard deviation, median, interquartile range, minimum and maximum for continuous measures, and frequencies and percentages of patients in each category for categorical measures. These measures include, but may not be limited to:

Demographics

- Age
- Gender
- Race and ethnicity
- Height, weight, and BMI
- Comorbidities
- Smoking history
- Karnofsky Performance Status

Disease Status

- Biopsy information
- Primary cancer diagnosis
- Treatment/therapy information
- Tumor location

Procedural Information

- Tumor location and description
- Tumor size measured by two methods:
    - 3-dimensional measurements (cm) and
    - Greatest trans-axial diameter (cm)
- Number and type of cryoablation needles used
- Number of freeze and thaw cycles
- Freeze and thaw times, per cycle
- Estimation of % tumor ablation during procedure
- Concomitant procedures
- Number and type of intraoperative complication(s)
    - Severity of complication(s)
- Duration of hospital stay
- Safety assessments

All protocol deviations

Subject Disposition

Subject disposition will be presented as:

- Number of subjects enrolled
- Number of subjects completing each follow-up visit
- Number of subjects withdrawn or lost to follow-up
- Number of subject deaths
- Number of subjects evaluable for ITT and PP analyses

Randomization and Blinding

As this is a single arm study, no randomization or blinding will be performed. To minimize selection bias within the enrolled patient population, participation in the study is to be offered to all eligible patients on a consecutive basis as they present to the physician at the site.

Handling of Missing Data

Missing data will be prospectively minimized through training of the participating investigator and site staff, and through appropriate clinical trial management. Every effort will be made to collect all data points in the study. All patient data that is available on patients who drop out during the course of the study will be included where possible.

Subjects without missing endpoint data will be considered completers. Intent-to-Treat Completers (ITTC) analysis as well as Per-Protocol Completers (PPC) analysis will be supplemented by missing data sensitivity analyses. Sensitivity analyses will be performed on the primary efficacy endpoint using two methods of missing data imputation: multiple imputation and last observation carry forward. These imputation methods will be employed on both the ITT and PP populations resulting two additional outcome analyses: Intent-to-Treat Imputed (ITTI) and Per-Protocol Imputed (PPI). Of the four analyses, the primary hypothesis test will be considered the ITTI analysis with multiple imputation.

Subset Analyses

The primary efficacy endpoint will be evaluated by gender.

Poolability Analyses

Though individual treatments may vary (needle used, number of needles used, and freeze/thaw times, etc.), it is expected that data will be poolable across study sites since all sites and investigators will follow a common protocol with identical inclusion/exclusion criteria (same population) and outcome assessment. The primary efficacy endpoint will be evaluated by study site.

Statistical Software

All statistical analyses will be produced using SAS software Version 9.3 or above (SAS Institute, Cary, NC.).

Version, Revision History, and Approvals

Current Version

Version 3.0 – Approval Date: 18MAY2016

Revision History

Version	Approval Date	Key Changes
1.0	28JUL2015	Initial Release
2.0	29MAR2016	Updated sample size justification including pilot study data to provide support of estimated attrition rate and sensitivity analyses to account for drop-out subjects; added clarification for missing data analyses.
3.0	18MAY2016	Revision to expand participating site number and clarification of sample size

Approvals

Name	Study Role	Signature	Date
[redacted]	Galil – [redacted]	DocuSigned by: [redacted]	18-May-2016
[redacted]	Galil – [redacted]	DocuSigned by: [redacted]	18-May-2016
[redacted]	NAMSA – [redacted]	DocuSigned by: [redacted]	18-May-2016