UROTRONIC

ROBUST III Clinical Study

Re-Establishing Flow Via Drug Coated Balloon For The Treatment Of Urethral Stricture Disease – A Randomized Control Trial

Protocol No.: PR1076-001

Protocol Version: J

Protocol Date: 13May2020

NCT03499964

SPONSOR
Urotronic Inc
2495 Xenium Lane N
Minneapolis, MN 55441

Title:	ROBUST III Re-Establishing Flow Via Drug Coated Balloon For The Treatment Of Urethral Stricture Disease- A Randomized Control Trial
Objective:	The study described below is designed to establish the safety and effectiveness for the Optilume™ Stricture Drug-Coated Balloon (DCB).
Study Design:	This is a prospective, multi-center, randomized controlled adaptive sample size clinical trial with one planned interim analysis for sample size re-estimation with a non-randomized pharmakonetics (PK) study arm.
Interventions:	Study Device: Urotronic Optilume Stricture DCB Control: Standard of care dilation method as determined by the treating physician
Enrollment:	A minimum of 126 Subjects, up to a maximum of 200 will be enrolled and randomized in the randomized arm of the study. An additional 15 subjects will be enrolled and treated in the PK arm of the study. A subject will be considered enrolled when they have signed informed consent and have been randomized. Individual site enrollment may not exceed 30% of the total study enrollment.
Randomized Arm Inclusion Criteria:	1. Male subjects ≥ 18 years' old 2. Visual confirmation of stricture via cystoscopy or urethrogram 3. Single, tandem or diffuse anterior urethral stricture(s), less than or equal to 3.0 cm total length measured by retrograde urethrogram. (Stricture length is defined as the distance between the most distal edge of the stricture to the most proximal edge of the stricture). 4. Two or more prior dilation treatments of the same stricture, including DVIU (Direct Vision Internal Urethrotomy), but no prior urethroplasty. *Note: Catheterization is not considered a dilation treatment*. 5. Significant symptoms of stricture such as frequency of urination, dysuria, urgency, hematuria, slow flow, feeling of incomplete emptying, recurrent urinary tract infections (UTI's). 6. International Prostrate Symptoms Score (IPSS) score of 11 or higher (assumed to be "35" if suprapubic catheter is present) 7. Lumen diameter ≤12F by urethrogram 8. Qmax <15 ml/sec (assumed to be "0" if suprapubic catheter is present) 9. Guidewire must be able to cross the lesion

Randomized Arms: Exclusion Criteria:	1. Subjects with diffuse stricture length, greater than 3.0 cm in total length. (Stricture length is defined as the distance between the most distal edge of the stricture to the most proximal edge of the stricture). 2. Subjects with a history of hypersensitivity reactions to TAXOL, on medication that may have negative interaction with paclitaxel, with solid tumors who have a baseline neutrophil counts of <1500 cells/mm^3 or subjects with AIDS-related Kaposi's sarcoma with baseline neutrophile counts of <1000 cells/mm^3. 3. Subjects who had an indwelling suprapubic catheter longer than three (3) months total prior to enrollment. 4. Previous urethroplasty within the anterior urethra. 5. Stricture dilated or incised within the last six (6) weeks (urethral catheterization is not considered dilation) 6. Presence of local adverse factors, including abnormal prostate making catheterization difficult, urethral false passage or fistula. 7. Presence of signs of obstructive voiding symptoms not directly attributable to the stricture at the discretion of the physician 8. Diagnosis of untreated and unresolved BPH or BNC 9. Untreated stress urinary incontinence (SUI). 10. History of diagnosed radiation cystitis 11. Diagnosis of carcinoma of the urethra, bladder or prostate within the last two (2) years 12. Active kidney, bladder, urethral or ureteral stone passage in the last six (6) weeks or concern of stone passage in the next 6 weeks at the discretion of the investigator. 13. Diagnosis of chronic renal failure and treatment with hemodialysis 14. New diagnosis of OAB (overactive bladder) within the last six (6) months 15. Use of alpha blockers, OAB (Overactive Bladder) medication, anticonvulsants (drugs that prevent or reduce the severity and frequency of seizures), and antispasmodics where the dose is not stable. (Stable dose is defined as having the same medication and dose in the last six months.) 16. Dependence on Botox (onabotulinumtoxinA) in urinary system 17. Presence of an artificial urinary sphincter, slings or stent(s) in the urethra or prostate 18. Known neurogenic bladder, sphincter abnormalities, or poor detrusor muscle function 19. Diagnosed with Lichen Sclerosus, or stricture due to balanitis xerotica obliterans (BXO) 20. Previous hypospadias repair 21. History of cancer in non-genitourinary system which is not considered in complete remission (except basal cell or squamous cell carcinoma of the skin). A potential participant is considered in complete remission if there has been no evidence of cancer within two (2) years of enrollment 22. Any cognitive or psychiatric condition that interferes with or precludes direct and accurate communication with the study investigator regarding the study or affect the ability to complete the study quality of life questionnaires 23. Unwilling to use protected sex for thirty (30) days' post treatment 24. Unwilling to abstain or use protected sex for ninety (90) days post treatment if sexual partner is of child bearing potential. 25. Inability to provide Informed Consent Form (ICF) and/or comply with all the required follow-up requirements

	26. Participation in other pre-market studies or treatment with an investigational drug or device. Long term follow up or post market study of an approved device is allowed. 27. Current active infection in the urinary system 28. Current uncontrolled diabetes (hemoglobin A1c > 8.0%) or evidence of poor wound healing due to diabetes 29. Diagnosed or suspected primary neurologic conditions such as multiple sclerosis or Parkinson's disease or other neurological diseases known to affect bladder function, sphincter function or poor detrusor muscle function. 30. Visible hematuria in subject's urine sample without known contributing factor 31. Invisible hematuria (or significant microscopic hematuria, i.e. hematuria of ≥ 3 RBC's/HPF) that may be caused by a clinically significant disease unless it is attributed to the urethral stricture disease or other causes which are benign and not requiring treatment.
PK Arm Inclusion Criteria:	In addition to meeting all the inclusion criteria for the Randomization arm of the study, subjects must also meet the following criterion in order to be eligible to participate in the PK arm of the study: • Urethral stricture measurements appropriate for treatment with a 10mm (30F) x 50mm, 12mm (36F) x 30mm, or 12mm (36F) x 50mm Optilume Drug Coated Balloon
PK Arm Exclusion Criteria:	In addition to not meeting any of the exclusion criteria for the Randomization arm of the study, subjects must also not meet the following criteria in order to be eligible to participate in the PK arm of the study: • Prior treatment with any device or medical therapy that contains paclitaxel, including drug coated balloons for vascular applications
Follow-up:	Follow-up visits required post-procedure: Foley Removal (2–5 days), 30 days, 3 months, 6 months, 12 months, and annually thereafter for 5 years post-procedure.
Clinical Sites:	Up to thirty (30) clinical sites in the US and Canada willing to participate in the study
Primary Efficacy Endpoint	Stricture Free Rate at six (6) months • Measured by passing a 16F flexible cystoscope at 6 months' post-treatment • If a 16F flexible cystoscope cannot pass, a 14F rubber catheter will be used A stricture is defined to be resolved when a 16F flexible cystoscope or a 14F rubber catheter can be passed through the stricture.
Primary Safety Endpoint	This safety endpoint is defined as a composite of specific device or procedure related serious complications at three (3) months. The proportion of subjects who experience a serious device or procedure related event of the following types will be reported: • Formation of fistula • Unresolved de novo stress urinary incontinence (requiring ≥1 pad/day) at ninety (90) days • Urethra rupture or burst
Secondary Endpoint 1 - Efficacy	Time to treatment failure Defined as any stricture treatment at the target treatment site.
Secondary Endpoint 2- Efficacy	Change in Qmax (Peak Flow Rate) at six (6) months

Secondary Endpoint -3- Efficacy	Percent responder at twelve (12) months (IPSS) • A responder is defined as a subject with a >50% improvement of IPSS score from baseline
Statistical Considerations	The study is designed as a prospective, multi-center, randomized, controlled adaptive sample size clincial trial. Subjects will be randomized in a 2:1 ratio to Test:Control and randomization will be stratifed by site. Sample size for the randomized portion of the study is based on power considerations for the primary efficacy endpoint. Sample size for the study is based on the primary effectiveness endpoint using the following assumptions: • 2:1 randomization allocation • Type 1 error of 0.025, one sided • Statistical power of approximately 90% • Assumed population success rate of 40% for the Control arm and 72% for Treatment arm, corresponding to a difference of 32%. Based on these assumptions, an initial sample size of 126 evaluable subjects (Test: 84; Control: 42) provides approximately 90% power. The primary safety endpoint will be analyzed with descriptive statistics and nominal 95% confidence intervals; there will be no formal statistical hypotheisis test. Due to uncertainty with respect to the design assumptions, an adaptive sample size methodology is planned. An interim analysis of the primary effectiveness endpoint is planned when primary effectiveness endpoint data is available on the first 60 randomized subjects (approximately 48% of the planned original total evaluable sample size). At the interim analysis, the primary effectiveness endpoint will be evaluated; if warranted, the sample size of the trial may be increased up to a maximum of two hundred (200) total subjects to maintain the study power. An additional 15 subjects will be enrolled in the non-randomized PK arm of the study, for a minimum of 141 randomized/treated subjects enrolled in the study. Data from non-randomized PK subjects will be reported separately.
Primary Efficacy Hypothesis	For the trial to be successful, the statistical evaluation for the resolution of the stricture at six (6) months of the Optilume arm will be statistically compared to the control arm on subjects in the randomized part of the study. Ho: $P_t \leq P_c$ Ha: $P_t > P_c$ Where P_t is the stricture free rate at 6 months in the Treatment arm and P_C is the stricture free rate at six (6) months in the Control arm.
Primary Safety Hypothesis	The primary safety endpoint will be analyzed with descriptive statistics and nominal 95% confidence intervals; there will be no formal statistical hypotheisis test.

Principal Investigators	Sean Elliott, MD University of Minnesota Department of Urology 420 Delaware St. SE, MMC 394 Minneapolis, MN 55455 Ramon Virasoro, MD Urology of Virginia 225 Clearfield Ave Virginia Beach, VA 23462
Independent Review Committees:	An independent Clinical Events Committee (CEC) will be utilized for this study. The CEC will be responsible for adjudicating the seriousness and relatedness of all potential device and/or procedure related adverse events occurring during the study period. A charter will be completed for the CEC outlining membership, duties and functions. An independent Data Monitoring Committee (DMC) will be utilized for this study. The DMC will be responsible for evaluating safety by reviewing overall study outcomes, adverse events and determining the implementation of the adaptive design protocol. The DMC will review the first interim analysis and provide feedback to the sponsor and study sites regarding the adaptation of the study design if required. A charter will be completed for the DMC outlining membership, duties and functions.

1.1 OVERALL STUDY DESIGN AND PLAN

This is a prospective, multi-national, multicenter, randomized, active control, single blinded, clinical trial. The minimum sample size is 126 evaluable subjects.

This study will include a single interim analysis with a sample size re-estimation (SSR). The interim analysis will be conducted when 6-month data are available on 60 randomized subjects. The required sample size will be re-estimated based on the results from interim analysis, but the trial size will not exceed 200 subjects regardless of interim results. Up to 30 investigational sites will be recruited for the study. Individual site enrollment may not exceed 30% of the total study enrollment. The Data Monitoring Committee (DMC) will review the interim analysis results, including the SSR, and make recommendations related to trial continuation to the sponsor.

Subjects will be stratified by prior pelvic radiation treatment and the number of prior treatments of their stricture prior to study enrollment (<5 or ≥ 5). After the completion of the 6-month follow up evaluation, the subjects will be unblinded.

At 6 months after randomization, the effectiveness of the Optilume will be demonstrated by comparison to the control group. Additionally, all subjects randomized to the Optilume group, or who later receive the Optilume, will be monitored for adverse events and continued questionnaire and Qmax assessment through 5 years post procedure.

The primary safety endpoint is an assessment of the proportion of subjects experiencing a composite of pre-defined procedure or device related serious complications in the treatment arm through 90 days (i.e., 3 months).The event types are listed in Section 1.2.2.

1.2 PRIMARY ENDPOINTS

1.2.1 Primary Effectiveness Endpoint – Stricture Free Rate at 6 Months

This endpoint will be evaluated by the ability to pass a 16Fr flexible cystoscope or a 14Fr rubber catheter at 6 months post-treatment. If a 16Fr cystoscope cannot be passed, a 14Fr red rubber catheter will be used.

This primary effectiveness endpoint of the study will be met if the rate of success in the Treatment Arm is statistically superior to the success rate in the control arm at 6 months post treatment, based on rejection of the null hypothesis for the following statistical test.

The statistical hypothesis test for the primary effectiveness endpoint will be based on a two-sample continuity corrected Chi-square test at the one-sided 0.025 alpha level (equivalent to a two-sided 0.05 alpha level) implemented with multiple imputation to account for missing data.

$$\text{Ho: } P_t \leq P_c$$

$$\text{Ha: } P_t > P_c$$

Where P_t is the stricture free rate at 6 months in the Treatment arm and P_C is the stricture free rate at 6 months in the Control arm. See also section 3.2 regarding a sample size re-estimation and associated statistical adjustments to the primary effectiveness endpoint analysis.

1.2.1.1 Primary Effectiveness Endpoint Treatment Failures Definition

Any subjects who have a second dilation procedure, pursue surgical intervention, or otherwise seek alternative treatment for the target stricture after the index procedure are considered

treatment failures for the primary analysis. Subjects who cross-over to receive treatment with the Optilume device will be considered a treatment failure for the primary therapy. At the 6 months follow up, if a 16F flexible cystoscope or a 14F rubber catheter cannot cross the treated stricture, the subject will be considered a treatment failure. This 6 month follow-up urethral lumen test will be analyzed as occurring at 180 days for time-to-event analyses.

Note: Treatment of obstructive symptoms due to other reasons (e.g. BPH, BNC) is not considered a treatment failure unless the stricture treated at the index procedure is also treated.

1.2.2 Primary Safety Endpoint

This safety endpoint will be analyzed with descriptive statistics for a composite of specific device or procedure related serious complications at 3 months. The proportion of subjects who experience a serious device or procedure related event of the following types will be reported:

- Formation of fistula
- Unresolved de novo stress urinary incontinence (requiring ≥1 pad/day)
- Urethra rupture or burst

The primary safety endpoint will be analyzed with descriptive statistics and nominal 95% confidence interval; there is no formal statistical hypothesis test planned for this endpoint.

1.3 SECONDARY ENDPOINTS

The following secondary endpoints employ formal statistical hypothesis tests for the purposes of supporting labeling claims with inferential quantities. Endpoints will be tested in the sequential order listed, following a gatekeeping strategy to control the type I error rate.

1.3.1 Secondary Endpoint 1: Time to treatment failure at 6 months

The following hypothesis will be tested in a one-sided, two sample log-rank test:

$$\text{Ho: } S_{Test} \leq S_{Control}$$

$$\text{Ha: } S_{Test} > S_{Control}$$

Where S_{Test} and $S_{Control}$ are the survival distributions for the time to first treatment failure (i.e. they represent the freedom from treatment failure). Treatment failure is defined in Section 1.2.1.1. The test will be performed at the one-sided 0.05 alpha level. Successful rejection of the null hypothesis will indicate the freedom from treatment failure for the treatment group is superior (i.e. longer time free from treatment failure) than for the control. Results will be described via Kaplan-Meier analysis.

1.3.2 Secondary Endpoint 2: Changes in Qmax at 6 months

The following hypothesis will be tested in a one-sided, two sample t-test for means:

$$\text{Ho: } \mu_{Test6} \leq \mu_{Control6}$$

$$\text{Ha: } \mu_{Test6} > \mu_{Control6}$$

Where μ_{Test3} is the change in Qmax at 6 months for subjects randomized to Optilume and $\mu_{Control3}$ is the change in Qmax at 6 months for subjects randomized to the Control. The test will be performed at the one-sided 0.05 alpha level. Successful rejection of the null hypothesis will indicate the change in Qmax for the device is statistically superior to the control. Non-parametric

statistics will be used as a supportive analysis to protect against violations of the normality assumption.

1.3.3 Secondary Endpoint 3: Percent Responder at 12 Months (IPSS)

A responder is defined as a 50% improvement in IPSS score. The following hypothesis will be tested in a one-sided, one-sample exact test for a binomial proportion.

$$H_o: P_{Test12} \leq 50\%$$

$$H_1: P_{Test12} > 50\%$$

Where P_{Test12} is the success rate at 12 months in all subjects treated with the Optilume, 50% is a performance goal. The test will be performed at the one-sided 0.05 alpha level. Successful rejection of the null hypothesis will indicate the success rate for the device is statistically greater than 50%.

1.3.4 Ancillary Endpoints and Other Analyses

The ancillary endpoints are to provide additional characterization of the safety and effectiveness of the Optilume in the treatment of stricture. No formal hypotheses tests for the purposes of supporting labeling claims are planned for ancillary endpoints. The ancillary endpoints and analyses include the following:

- Change from baseline in Qmax at 12, 24, 36, 48 and 60 months
- Change from baseline and percent responder in IPSS at 12, 24, 36, 48 and 60 months
- Rate of acute urinary retention requiring catheterization at 6 months
- Change Quality of Life (QoL)
- Procedure parameters including procedure time, treatment time and healing time
- Time to treatment failure through 12, 24, 36, 48, and 60 months of follow-up

In supportive analyses, data on the primary endpoints for control subjects who crossover to receive the active treatment will be summarized, based on treating their time of crossover as baseline. These results may be pooled with randomized treatment group subjects to better inform the performance of the device.

Generally, descriptive statistics will be used in reporting outcomes for other effectiveness secondary endpoints. Continuous variables will be summarized with means or medians, standard deviations. Adverse events, protocol deviations and device malfunction will be summarized with descriptive statistics.

All p-values in the ancillary endpoints and other analyses will be nominal and not adjusted for multiple testing; it is recognized this may limit the ability to use such p-values in product labeling.

2 RANDOMIZATION AND BLINDING

2.1 RANDOMIZATION PLAN

Subjects will be randomized in a 2:1 allocation of treatment vs control. Randomization will be stratified by investigational center and by prior radiation treatment (yes or no) and number of prior dilation treatments (i.e. less than 5 prior dilations versus ≥ 5prior dilations). Each treatment

group will have its own randomization schedule within each participating center. For each randomization schedule, randomization will be performed using randomly permuted blocks.

Those subjects who do not meet inclusion/exclusion criteria after baseline evaluation will be counted as screening failures and will be withdrawn from the study. Each subject will be randomized prior to initiation of the treatment/control procedure. Only randomized subjects will be considered enrolled and evaluable.

Subjects will be blinded to the treatment. The treating physicians will not be blinded. Blinding will be broken only to protect the subject's health. If a non-urgent clinical need requires that the subject be unblinded prior to the 6-month follow-up and if time allows, the physician will notify the Sponsor prior to unblinding the subject. Subjects may be unblinded after the 6-month follow-up examination.

2.2 CROSS OVER OF CONTROL SUBJECTS

Subjects randomized to the control group will be offered the option to receive the Optilume treatment if their stricture has been confirmed to have recurred (<12F as measured by urethrogram AND recurrent symptoms AND reduced flow rates). Crossover will be limited to those subjects with confirmed stricture recurrence by the end of the 12 month follow up window.

Alternatively, subjects in the control arm may choose to be treated with another commercially available treatment. Those subjects in the control arm that elect to have a different stricture procedure will be exited from the study and will only be followed until AE resolution. Those subjects who elect to cross over to the treatment arm will be followed up per the same schedule as the treatment group.

Follow-up of the subjects will be continued after the PMA submission and continue as part of the long-term follow up for up to 5 years post treatment. Subjects in the control arm who did not elect to cross over to the Test procedure or to have another procedure will be followed up only until all AEs are resolved or up to 12 months post randomization, whichever is later, at which point the subjects will exit the study.

3 STUDY SAMPLE JUSTIFICATION

3.1 PRIMARY EFFECTIVENESS ENDPOINT SAMPLE SIZE CONSIDERATIONS

Sample size for the primary effectiveness endpoint was based on the following assumptions:

- 2:1 randomization allocation
- Type 1 error of 0.025, one sided
- Statistical power of approximately 90%
- Assumed population success rate of 40% for the Control arm and 72% for Treatment arm, corresponding to a difference of 32%.

Based on these assumptions, an initial sample size of 126 subjects (Test: 84; Control: 42) provides approximately 90% power.

3.2 SAMPLE SIZE RE-ESTIMATION

Due to uncertainty with respect to the design assumptions, an adaptive sample size methodology is planned[1,2]. An interim analysis of the primary effectiveness endpoint is planned when primary effectiveness endpoint data is available on the first 60 randomized subjects (approximately 48% of the planned original total evaluable sample size). At the interim analysis, the primary effectiveness endpoint will be evaluated; if warranted, the sample size of the trial may be increased up to a maximum of 200 total subjects to maintain the study power.

The study will not stop because of effectiveness or futility at the interim analysis. The planned evaluable final sample size will not be adjusted downward from the originally planned 126 subjects, and the maximum sample size that may be randomized in this study following the sample size re-estimation is 200 subjects.

The interim analysis will be performed by an independent statistician unblinded to the treatment group success rate. The designated study statistician may perform the interim analysis if an independent, blinded statistician is consulted for any potential modifications to the study design after completion of the interim analysis but before the primary analysis. The statistician conducting the interim analysis will not be allowed to give input on study design modifications between the interim analysis and primary analysis to minimize the potential for introduction of bias.

For the interim analysis, conditional power calculations for sample sizes up to the maximum of 200 will be performed using the observed success rate among the first 60 randomized subjects with available primary effectiveness endpoint data.

The final recommendation regarding the final sample size will be based on a combination of factors, both statistical and logistic (i.e. the past and expected future enrollment rate, attrition rate, etc.). As a guiding base, the result from the following formula will be provided to the DMC:

$$M = \max\left(N, \min\left(200, N\left(\frac{\delta}{\hat{\Delta}_i}\,\frac{se(\hat{\Delta}_i)}{se(\delta)}\right)^2\right)\right)$$

where N is the original planned evaluable sample size (126), δ is the original assumed treatment effect defined (32%), $se(\delta)$ is the original assumed standard error of the treatment effect at the interim (0.143), $\hat{\Delta}_1$ is the estimated treatment effect at the interim using multiple imputation, $se(\hat{\Delta}_i)$ is the estimated standard error of the treatment effect at the interim using multiple imputation, and M is the planned evaluable final sample size. The standard error of the estimated treatment effect using multiple imputation is typically smaller than if only complete case data were analyzed but larger than if all data were observed. The sample size re-estimation accounts for both the assumed treatment effect and missing data. M is constrained to be no less than 126 subjects, and no more than 200 total subjects. If 126 subjects are randomized prior to having primary effectiveness endpoint data on the first 60 randomized subjects, enrollment will halt until the interim analysis takes place. If dictated by the results of the interim analysis, enrollment may then continue up to the maximum of 200 subjects.

---

[1] Proschan MA, Lan KKG, Wittes JT, "Statistical Monitoring of Clinical Trials: A Unified Approach", Springer 2006, Chapter 11.4.

[2] Cui L, Hung HMJ, Wang SJ, "Modification of Sample Size in Group Sequential Trials", Biometrics 55, 853-857, 1999.

Only the final sample size recommended by the DMC will be shared with Urotronic and the clinical team to avoid the potential for bias.

To account for the sample size re-estimation in the final analysis of the primary effectiveness endpoint, a weighted Z test will be employed. Let Z_1 and Z_2 be the corresponding stage-wise independent standard normal statistics calculated as

$$Z_i = \frac{\hat{\Delta}_i}{se(\hat{\Delta}_i)}, \ \ i = 1,2$$

where $\hat{\Delta}_i$ is the estimated difference between the treatment and control stricture rates at 6 months and $se(\hat{\Delta}_i)$ is the standard error of the estimated difference. Both $\hat{\Delta}_i$ and $se(\hat{\Delta}_i)$ are obtained from the multiple imputation analysis using Rubin's method, which estimates the standard error as the combination of within-imputation and between-imputation components (e.g. as implemented in PROC MIANALYZE in SAS). Specification of the multiple imputation model is provided in Section 6.1.1.

At the conclusion of the study, the final test statistic is calculated as

$$Z = \frac{\sqrt{n_1}Z_1 + \sqrt{n_2}Z_2}{\sqrt{n_1 + n_2}}$$

With $n_1 = \frac{60}{126}$ and $n_2 = \frac{126-60}{126}$, fixed weights that do not depend on the sample size re-estimation, the type I error rate is preserved[1]. Z_1 and Z_2 are the multiple imputation test statistics of stage 1 (before interim analysis) and stage 2 (between interim analysis and final analysis) respectively.

4 ANALYSIS SET

All subjects enrolled in the study (including those withdrawn from the investigation or lost to follow-up) will be accounted for and documented.

The primary endpoints analyses and the secondary endpoints will be performed on the intent-to-treat (ITT) set, under which all randomized subjects will be included for the analysis, regardless of whether or not the subjects received the treatment to which they were randomized during the index treatment.

In addition to the ITT analysis, the primary endpoint analyses will also be performed on the as treated set (i.e. subjects analyzed based on the treatment actually received at baseline) and per-protocol (PP) set, (*i.e.,* subjects treated and followed per the protocol) where appropriate.

The per protocol analysis set will exclude subjects with the following significant protocol violations:

- Subjects with significant violations of the inclusion and exclusion criteria including the following:
    - Subject's rights are violated or did not give consent and subject data was requested to be excluded by IRB
    - Subject's obstructive symptoms are primarily due to BPH or bladder neck contracture in addition to or rather than stricture.
    - Subject had UTI at the time of treatment

- Subject had a previous radical prostatectomy
- Subject had a previous urethroplasty

Other additional criteria may be added to the list to accommodate for unforeseen events that occurred during the conduct of the trial that resulted in noteworthy study protocol violations. Significant protocol violations will be summarized by category and by site.

The list of protocol violations including the major and minor classification will be reviewed and finalized before the database lock for final analysis.

All other ancillary endpoints will be analyzed for those subjects treated by the treatment or control device irrespective of the original randomization assignment (As Treated); subjects will be grouped based on the treatment received.

5 POOLABILITY OF DATA

5.1.1 Sites

This study is designed and conducted as a multicenter randomized-control clinical trial. Data from all the sites will be pooled. All subjects will be treated and evaluated following the same protocol to ensure generalizability of the study results.

Heterogeneity in the treatment effect for the primary effectiveness endpoint will be assessed via logistic regression models with covariates for site, treatment group, and the interaction of site and treatment group. A p-value for the interaction term <0.15 will suggest evidence of variation in the treatment effect and will trigger additional exploratory analyses to attempt to further quantify and explain the variation. Firth's penalized likelihood may be used if needed due to sparse data. Sites with fewer than 4 subjects will be combined into a single super-site for these analyses.

5.1.2 Subgroup Analysis

Subgroup analyses will be performed for the 6-month primary effectiveness endpoint to understand potential variation in the treatment effect. Statistical methods will follow the same as those outlined for poolability of site (Section 5.1.1). Subgroups will be defined based on the following baseline factors:

- Baseline number of prior treatments
- Baseline lesion length (≤ the median length vs. > the median length)
- Subject age at the time of randomization (≤ 50 years vs. > 50 year)
- Subject race (White vs. Non-White)
- Baseline IPSS Score (≤19 vs. ≥20, i.e. moderate vs. severe)
- Use of supra-pubic catheter at baseline (yes vs. no)
- Prior radiation (yes vs. no)
- Number of prior dilations (less than 5 prior dilations versus ≥ 5prior dilations)

6 MISSING DATA

Every effort will be made to reduce the incidence of missing data. All available data on subjects who drop out during the study will be included.

6.1 MISSING DATA HANDLING RULES

6.1.1 For the Primary Effectiveness Endpoint

1) Subjects will be considered a treatment failure for the primary effectiveness endpoint if the subject opts to break blinding and seek alternative treatment, including cross-over, or are otherwise exited due to treatment failure as described in Section 1.2.1.1 before the close of the 6-month follow-up visit window.
2) For subjects whose stricture assessment was done after the 6 months compliance window, assessment after the close of the 6 month window will be accepted as observed data for this endpoint through a maximum of 8 months post-index. [*Note: Strictures noted at late assessments represent the worst case for this endpoint, since the patient may have been stricture-free had the assessment been performed on time.*]
3) Subjects who either drop out, are lost to follow-up prior to the 6-month evaluation, or have a late stricture assessment >8 months post-index will be analyzed using multiple imputation methods as described below unless they are documented as a Treatment Failure as described in Section 1.2.1.1.

   Multiple imputation will be performed for the 6-month stricture primary endpoint under the missing at random (MAR) assumption using a logistic regression model (e.g. FCS logistic procedure in PROC MI in SAS) conditional on the observed predictor variables:

   - IPSS at 3, 6 and 12 month follow-up and at late 6 month cystoscopy (>8 months)
   - Qmax at 3, 6 and 12 month follow-up
   - Stricture per late 6 month cystoscopy (>8 months)
   - Stricture per 12 month cystoscopy
   - Baseline variables: Age, IPSS, dilation diameter, stricture length, prior radiation (yes/no) and prior dilations (count)

The above covariates for the multiple imputation model are based on clinical input. If multiple imputation predictor variables are missing, they will be imputed using multivariate normal Markov Chain Monte Carlo imputation (e.g. MCMC procedure in PROC MI in SAS) prior to imputing the primary endpoint. If necessary to facilitate model fitting, an augmented likelihood approach or removal of variables will be employed. Imputation is performed separately by treatment received to eliminate the possibility of cross-contamination of imputation models between subjects receiving the treatment and control interventions, thereby permitting potentially different relationships among the predictor variables in the imputation model by group. A total of 100 imputations per group is planned, resulting in a total of 100 imputed datasets. This number of imputations is selected to ensure accurate inference for confidence intervals and p-values. A common guideline is that the number of imputations should be greater than or equal to the percentage of data that is missing[3]. Using 100 imputations therefore covers all possible missing

[3] Graham, John W., Allison E. Olchowski and Tamika D. Gilreath. "How many imputations are really needed? Some practical clarifications of multiple imputation theory." *Prevention Science* 8, 206-213, 2007.

data scenarios. Imputed values for binary and count variables will not be rounded, based on current recommendations[4,5].

6.1.2 For the Primary Safety Endpoint

Unless there is evidence of occurrence of a primary safety endpoint, subjects with missing data for the primary safety endpoint are presumed to not have experienced a primary safety endpoint.

6.1.3 Secondary Endpoint 1: Time to Treatment Failure at 6 Months

1) Subjects will be considered a treatment failure for this endpoint if the subject opts to seek alternative treatment before the specified follow-up visit as described in Section 1.2.1.1.
2) Subjects who either drop out or are lost to follow-up or have missing endpoint data will be analyzed using the same multiple imputation methods as the primary endpoint as unless they are documented as a Treatment Failure described in Section 1.2.1.1. For subjects with an imputed failure, failure time will be analyzed as occurring at 180 days.

6.1.4 For the Secondary and Ancillary Endpoints: Uroflow, QoL, Responder Analyses

Subjects who dropped out, are lost to follow-up, or otherwise have missing data for secondary endpoints with hypothesis tests will be evaluated using the same multiple imputation variables and methods as the primary endpoint. For subjects considered a "treatment failure" for the primary effectiveness endpoint based on clinical status as described in, 6.1.1, a failure will be imputed for the secondary and ancillary endpoints as appropriate (i.e. failure for binary endpoints or the worst observed value in the study for continuous endpoints). For other subjects, under the missing at random (MAR) assumption, a linear regression model (e.g. FCS reg procedure in PROC MI in SAS) will be used to impute 6 month Qmax and 12 month IPSS. The IPSS responder analysis will use the imputed 12 month IPSS score to derive an imputed responder status.

6.2 MULTIPLE RESPONSES

If a subject has multiple baseline evaluations, the last baseline value prior to the procedure will be used in the analysis.

If a subject has multiple post-baseline evaluations within the same time visit window period, the latest results will be used for the subject in all analyses unless there is scientific valid reason(s) to exclude one or more of the evaluations.

6.3 SENSITIVITY ANALYSES FOR MISSING DATA

For the primary effectiveness endpoint, sensitivity analysis, *e.g.*, tipping point analysis, will be performed to evaluate the impact of missing data on study conclusion. Complete-case (i.e. observed data) and last observation carried forward analyses will be used as sensitivity analyses for the primary effectiveness endpoint and secondary endpoints. For continuous outcomes, linear interpolation between the nearest visits with observed data will serve as an additional sensitivity analysis.

---

[4] Allison, Paul D. 2005. "Imputation of Categorical Variables with PROC MI," Presented at the 30th Meeting of SAS Users Group International, April 10–13, Philadephia, PA, 2005.

[5] Nicholas J Horton, Stuart R Lipsitz & Michael Parzen. "A Potential for Bias When Rounding in Multiple Imputation", The American Statistician 57:4, 229-232. 2003.

Primary effectiveness endpoint will also be analyzed by adjusting for the randomization factors (i.e., prior radiation and number of prior dilations). This analysis will be performed using a multivariable logistic regression model with the randomization factors as covariates. The odds ratio for the randomized arm and its standard error will be pooled across the multiple imputations using Rubin's method.