EXPAND URO

A prospective, multi-center, single-arm study of the Medtronic Hugo™ Robotic Assisted Surgery (RAS) System in Urologic Surgery

Statistical Analysis Plan v3.0
04/NOV/2024

NCT05696444

Medtronic **Statistical Analysis Plan**	
Clinical Investigation Plan Title	A prospective, multi-center, single-arm study of the Medtronic Hugo™ Robotic Assisted Surgery (RAS) System in Urologic Surgery (Expand URO)
Clinical Investigation Plan Identifier	MDT19051EINURO
Clinical Investigation Plan Version	Version 9.0 / 21-AUG-2024 Version 8.0 / 15-NOV-2023 Version 4.0 / 16-SEP-2021 (first 15 patients)
Statistical Analysis Plan Version Date	04-NOV-2024
Sponsor/Local Sponsor	Covidien LP, Medtronic Medical Surgical Surgical Robotics 60 Middletown Ave North Haven, CT 06473 USA

Table of Contents

1. Version History

Version	Summary of Changes	Author(s)/Title
1.0	• Initial Release	Zengri Wang, PhD, Director Biostatistics Evelyn Wu-Allen, MS, Principal Biostatistician
2.0	• Updates and clarifications were added throughout the document to be consistent with CIP 8.0 • Moved extraneous details from section 4 to be included with the analysis descriptions in section 7.9 • Added R version 4.0.0 as an allowable statistical software in section 7.2 • Added sub-group analysis by surgery type to primary effectiveness analysis • Added detail to section 7.9 to clarify the hypothesis, endpoint definition, performance requirements, performance requirements rationale, analysis methods, determination of subjects, and supporting analyses for all primary (Sections 7.9.1 and 7.9.2) and secondary objectives (Section 7.9.3). Analysis of ancillary objectives are also clarified in section 7.9.4	Charles Cain, PhD, Senior Statistician Casey Blaser, MS, Senior Principal Statistician
3.0	• Made minor updates to be consistent with CIP V9.0: Added wording “up to” to attrition-adjusted sample sizes statements throughout the document to account for lower-than-expected attrition rates. • Clarified the derivation of the time window for events in scope of the primary safety endpoint to be within 30 calendar-days post procedure (Section 7.9.2.2) Clarified the definition of the events in scope of the primary safety endpoint to be any CEC-adjudicated related (device, procedure, underlying condition/disease) adverse event (Section 7.9.2.2) • Clarified secondary endpoint definitions involving the 30-day follow-up to refer to the 30 calendar-day post procedure to align with the primary safety endpoint (Section 7.9.3.1) • Added an overview of the Holm-Bonferroni method (Section 7.5) • Added confidence intervals as a general methodology (Section 7.2)	Casey Blaser, MS, Senior Principal Statistician

2. List of Abbreviations and Definitions of Terms

Abbreviation	Definition
AE	Adverse Event
ADE	Adverse Device Effect
CEC	Clinical Event Committee
CIP	Clinical Investigation Plan
CSR	Clinical Study Report
DD	Device Deficiency
eCRF	Electronic Case Report Form
FAS	Full Analysis Set
FDA	United States Food and Drug Administration
MedDRA	Medical Dictionary for Regulatory Activities
PG	Performance Goal
PPAS	Per Protocol Analysis Set
RAS	Robotic Assisted Surgery
SADE	Serious Adverse Device Effect
SAP	Statistical Analysis Plan
SAS	Statistical Analysis System
SOP	Standard Operating Procedure
UAE	Unavoidable Adverse Event
UADE	Unanticipated Adverse Device Effect

3. Introduction

Over the course of history, surgery has generally evolved towards less invasive methods for performing the same procedures. Presently, surgical technique is undergoing another revolution – the growth of robotic-assisted surgery (RAS). Much as the laparoscope changed how surgery was practiced in the 20th century, RAS will similarly propel forward surgical specialties in the 21st century. RAS is based on the accurate translation of user input to a robotically assisted output. Similar to laparoscopic surgery, RAS involves the use of endoscopic instrumentation for manipulation of tissues and vessels in the insufflated body cavity.

The Medtronic Hugo™ RAS System is a modular robotic platform for performing robotically assisted minimally invasive surgery. This clinical study will evaluate the safety and effectiveness of the Medtronic Hugo™ RAS System when used for urologic RAS.

The statistical analysis plan (SAP) specifies the statistical methods to be implemented for the analysis of data used in this clinical study. It elaborates on the statistical analyses specified in the Clinical Investigation Plan (CIP) version 8, Dated 15-NOV-2023 (the first 15-patient summary was based on CIP 4.0, dated 16-SEP-2021).

This document was created for internal use as a guideline for the study Biostatistician, Statistical Programmer(s), and other relevant stakeholders. Analysis results obtained from the statistical analyses outlined in this document will be the basis of the Clinical Study Report (CSR) for this study.

As with any statistical analysis plan, the proposed methods and approaches to the data analysis should be considered as flexible to accommodate necessary changes. Changes to the plan may arise if emerging data suggest that deviations from the original plan would provide a more reliable and valid analysis of the data. Sound statistical reasoning will substantiate any and all deviations from this plan. The purpose of this plan is to provide general, and in some instances, specific guidelines from which the analysis will proceed.

The planned analyses identified in this SAP may be included in regulatory submissions and/or future manuscripts. Additional exploratory analyses, not identified in this SAP, may be performed to support the clinical development program. Any post-hoc, or unplanned, analyses that are performed for the CSR, but not identified in this SAP, will be clearly delineated in the CSR.

4. Study Objectives

4.1 Primary Objective

The primary objective of this study is to confirm that the Medtronic Hugo™ RAS System is safe and effective when used for urologic robotic-assisted surgery.

4.1.1 Primary Effectiveness Objective

The primary efficacy objective of this study is to confirm that the Medtronic Hugo™ RAS System is effective when used for urologic robotic assisted surgery. The primary effectiveness endpoint is the surgical success rate, defined as the procedure not going into conversion. Conversion is defined as the switch from a robotic-assisted approach using the Hugo™ system to a robotic-assisted approach utilizing an FDA cleared robotic-assisted device, laparoscopic or open surgery.

4.1.2 Primary Safety Objective

The primary safety objective of this study is to confirm that the Medtronic Hugo™ RAS System is safe when used for urologic robotic assisted surgery. The primary safety endpoint is the rate of subjects with major complications from the first incision through 30 days post-procedure. A major complication is defined as Grade III or higher complication per the Clavien-Dindo Classification system.

4.2 Secondary Objective

The secondary objective is to demonstrate that the Medtronic Hugo™ RAS System performs as intended when used in urologic robotic assisted surgery. The secondary endpoints will assess the overall performance of the Medtronic Hugo™ RAS System when used for urologic robotic surgery.

The following short-term secondary endpoints to assess the overall safety and performance of the Medtronic Hugo™ RAS System when used for urologic robotic surgery will be assessed.

- Complication rate: Overall rate of subjects with one or more complications (Clavien-Dindo Grade I or higher), from the first incision through 30 days post-procedure
- Operative time
- Intraoperative estimated blood loss (mL)
- Transfusion rate
- Rate of device-related conversion
- Hospital length of stay
- Readmission rate (through 30 days)
- Reoperation rate (through 30 days)
- Mortality rate (through 30 days)
- Rate of device deficiencies

The following long-term secondary endpoints will be assessed through 5 years in oncologic subjects:

- Overall survival
- Progression-free survival
- Disease-free survival

4.3 Ancillary Objectives

Descriptive analyses of other pre-specified outcome measures beyond the primary and secondary objectives will be exploratory in nature and are not intended as a focus of the study for the evaluation of the study device.

The following data will be collected as applicable:

- Rate of negative surgical margins
- Lymph node yield
- Warm ischemia time
- Surgeon experience

5. Investigation Plan

This is a prospective, multicenter, single-arm pivotal study to be performed in up to 141 subjects undergoing a urologic RAS procedure using the Medtronic Hugo™ RAS System. Subjects without an oncologic indication will be followed for 30 days (+7 days) post procedure. Oncologic subjects will be followed through 5 years. This study will be conducted using up to six investigative sites in the United States of America (USA).

This study will have two phases, a roll-in phase and open enrollment phase. Two (2) sites will consecutively enroll a total of 15 subjects who will be treated with the Hugo™ RAS System as “roll-in cases”. These subjects will undergo the same preoperative and postoperative assessments (with the same schedule) as patients who are enrolled in the open enrollment phase. Outcomes through 30 days for the first 15 roll-

in cases will be provided in a summarizing fashion and provided to the FDA for initial review without statistical inferences. The roll-in subjects will be included in the primary analysis with the open enrollment subjects and for all required reports.

Following completion of the roll-in phase and sponsor approval, all sites will be allowed to enroll patients in the study. During the open enrollment phase, all patients enrolled will follow the Clinical Investigation Plan.

The expected study duration is approximately 5 years. The duration of individual subject participation will be approximately 67 days (-30 to 30+7) for non-oncologic subjects and 5 years for oncological subjects.

6. Determination of Sample Size

The sample size for the study was estimated based on the primary safety and effectiveness hypotheses.

Power calculations were performed using statistical software NCSS PASS 2023.

The effectiveness hypothesis is to test if the surgical success rate is above the performance goal. The surgical success rate is expected to be 90% or higher. The effectiveness performance goal is set to be 85%. A sample size of 126 subjects will provide more than 80% power at one-sided alpha of 0.025 (with an expected rate of 95.2% or higher in this study).

Safety was also considered for the sample size determination. The primary safety endpoint is major complication rate. A statistical hypothesis with a performance goal is specified for each of the three urologic procedures (nephrectomy, prostatectomy, cystectomy). The safety performance goals were determined based on analyses of published literature by taking into account clinically meaningful margins and statistical precisions. Each individual hypothesis was adequately powered with at least 80% power at an alpha level of 0.008 with multiplicity adjustment taken into consideration. The sample sizes for the three surgery types are up to: 55, 55, and 31 for nephrectomy, prostatectomy, and cystectomy, respectively. After taking 10% attrition into account the total number of subjects planned for this study based on the sample size calculations was up to 141 subjects. Full calculations for the safety sample sizes and the associated power for each urologic procedure are below.

Nephrectomy: The mean effect size is 0.029 with a 95% confidence interval of 0.022 to 0.038. The 95% prediction interval is 0.017 to 0.047. The 99% prediction interval is 0.014 to 0.058. A performance goal of 0.20, while greater than the predicted interval, broadly aligns with the upper limits for each nephrectomy study with a comparable study size (95% upper limits ranging from 0.100 to 0.539) and is clinically appropriate. A performance goal of 0.20, assuming a sample size of up to 55 (with 10% attrition n=49), one-sided alpha of 0.008 (Bonferroni Correction), and mean effect size of 0.029, results in a power of 95%.

Prostatectomy: The mean effect size is 0.029 with a 95% confidence interval of 0.016 to 0.051. The 95% prediction interval is 0.004 to 0.173. The 99% prediction interval is 0.002 to 0.316. A performance goal of 0.20 (20%), is clinically appropriate. A performance goal of 0.20, assuming a sample size of up to 55 (with 10% attrition n=49), one-sided alpha of 0.008 (Bonferroni Correction), and mean effect size of 0.029, results in a power of 95%.

Cystectomy: The mean effect size is 0.141 with a 95% confidence interval of 0.095 to 0.205. The 95% prediction interval is 0.047 to 0.354. The 99% prediction interval is 0.024 to 0.520. A performance goal of

0.45 (45%), is clinically appropriate. A performance goal of 0.45, assuming a sample size of up to 31 (with 10% attrition n=27), one-sided alpha of 0.008 (Bonferroni Correction), and mean effect size of 0.141, results in a power of 83%.

With the consideration of both effectiveness and safety, a total of up to 141 subjects are planned to be treated in this study. Up to six study sites will be used. To keep enrollment balanced, each study site will be allowed to enroll no more than 30% of the total population.

7. Statistical Methods

7.1 Study Subjects

7.1.1 Disposition of Subjects

Subject disposition will be illustrated in a CONSORT diagram. The number of subjects screened, treated, discontinued during the study, as well as the reasons for discontinuations will be summarized for all centers combined and each center separately. Disposition and reason for study discontinuation will also be provided as a by-subject listing. All demographics and baseline characteristics will be summarized using descriptive statistics.

7.1.2 Clinical Investigation Plan (CIP) Deviations

Protocol deviations will be summarized, listed and discussed in the study report.

7.1.3 Analysis Sets

The following populations will be considered for the analysis of data for this study:

Full Analysis Set

The full analysis set (FAS) is defined using the modified intent-to-treat (mITT) principle as all enrolled subjects for whom a Hugo™ RAS procedure was begun, with the beginning of a procedure defined as the first skin incision being made. In the event that a subject is consented, but the first incision does not occur (e.g., if the subject becomes ineligible during the timeframe between consent and the procedure day), that subject will not be considered part of the FAS. The FAS will be the primary analysis set for the evaluation of the primary and secondary endpoints.

Adverse Events (AEs) will be collected from the time of consent. AEs occurring in patients excluded from the FAS will be followed for 30 days post-consent through study exit and will be reported in a listing in the CSR. These AEs will not be included in the primary FAS analysis for either AE reporting or the analysis of the primary and secondary endpoints. The number and proportion of subjects experiencing each type of AE will be summarized by site, surgery type and overall for the FAS.

Per Protocol Analysis Set

The per protocol analysis set (PPAS) is a subset of the FAS including only those subjects without any major protocol deviations. Major deviation reasons are: “Enrolled subject did not meet enrollment criteria” and

"Unauthorized use of investigational device." Reasons for exclusion of subjects from PPAS will be documented in tabular format and reported.

All enrolled subjects will be included in a subject disposition table indicating reasons for exclusion from the FAS and PPAS analysis sets.

7.2 General Methodology

All data analyses will be performed by Medtronic or its designee. All statistical analyses will be performed using the Statistical Analysis System (SAS) for Windows (version 9.4 or higher, SAS Institute Inc. Cary, NC), R (version 4.0.0 or higher), or other widely accepted statistical or graphical software.

Descriptive statistics will be used to summarize study outcomes. Continuous variables will be summarized using number of subjects (n), mean, standard deviation, median, interquartile range (IQR), and ranges. Categorical variables will be summarized using frequencies and percentages. Confidence intervals may be produced for either continuous or categorical variables as warranted.

All tests of treatment effects will be conducted at a two-sided alpha level of 0.05 unless otherwise stated. A p-value less than 0.05 is considered statistically significant. Confidence intervals will be presented at the 95% level unless otherwise stated.

The study safety and performance analyses will occur after all subjects complete the 30-day follow-up. The analyses will include both primary and secondary objectives which are related to 30-day follow-up. A CSR will be prepared once all data collection has ended and all subjects have completed the 30-day follow-up or have exited. Annual progress reports will be prepared for oncologic subjects through 5 years. A final report will be completed at the end of the 5-year follow-up.

7.3 Center Pooling

Study centers will be pooled together for all planned analyses as they are assumed to be homogeneous in terms of clinical practice, efficacy and safety. This assumption will be tested for the primary efficacy endpoint as later described in **Section 7.9.1.7.1**.

7.4 Handling of Missing, Unused, and Spurious Data and Dropouts

The primary analysis will be based on the full analysis set with no imputation of missing data. For secondary objectives, analyses will be conducted using subjects with available data from the FAS. Sensitivity analyses will be performed for the primary endpoints using multiple imputation for missing data when the missing percentage is greater than 5% as well as tipping-point analysis on the primary effectiveness endpoint to fully understand the missing data impact on the study results.

7.5 Adjustments for Multiple Comparisons

For the primary effectiveness endpoint only one hypothesis will be tested at an alpha of 0.025 one-sided. If the one-sided primary objective p-value is <0.025, then the null hypothesis will be rejected, and the test will be declared statistically significant.

After the primary effectiveness objective is met, the primary safety objective for each surgery type will be tested as described in **Section 7.9.2.1**. For the primary safety endpoint (major complication rate), the closed test procedure (Holm's method, i.e. Holm-Bonferroni method) will be used to protect the overall study-wise error rate for surgery types (nephrectomy, prostatectomy and cystectomy).[1]

The Holm-Bonferroni method is applied in context of this study as follows: start by listing all the p-values obtained from the multiple statistical tests and arranging them in ascending order, from the smallest to the largest. For each ordered p-value, calculate a sequentially adjusted significance level using the formula (alpha(i) = 0.025 / (3+1-i)), where '0.025' is the desired overall significance level, '3' is the total number of tests, 'i' is the rank of the p-value in the ordered list, and 'alpha(i)' is the adjusted significance level for test 'i'. Next, compare each p-value with its adjusted significance level, starting with the smallest p-value. If the smallest p-value is less than or equal to alpha(1), reject the null hypothesis for the first test and move to the next smallest p-value, comparing it to alpha(2). Continue this process sequentially for each p-value. As soon as a p-value is greater than its corresponding adjusted significance level alpha(k), the procedure is stopped and the remaining p-values for the null hypotheses are not rejected. The null hypotheses corresponding to the p-values that were rejected are considered statistically significant, while those corresponding to the p-values that were not rejected are not considered statistically significant. The Holm-Bonferroni method is more powerful than the simple Bonferroni correction because it adjusts the significance levels in a stepwise manner, making it less conservative.

There will be no multiplicity controls applied to the secondary endpoints as strict hypothesis testing is not being utilized.

7.6 Demographic and Other Baseline Characteristics

Subject demographics, medical, and surgical history will be summarized using descriptive statistics for continuous variables (number of observations, mean, standard deviation, median, IQR, and ranges) and frequency tables for discrete variables.

Age will be reported in years. Subjects with available data will be summarized; gender will be summarized using counts and percentages. In addition to the reported values, unknown or unreported values will also be in the summaries as a count of subjects (if any). The supportive data for the demographics table will be presented in a listing entitled "Subject Demographics".

Summary of procedural characteristics including procedure time, primary diagnosis, and robotic performance will be presented.

7.7 Treatment Characteristics

A descriptive data summary for surgical procedure characteristics and exposure to the study device will be provided. Summaries will be discussed in the study report.

7.8 Interim Analyses

No interim analysis is planned for this study. Outcomes through 30 days for the first 15 roll-in cases will be summarized descriptively and provided to the FDA for initial review without statistical inferences. The roll-in subjects will be included in the final data analysis for the study.

7.9 Evaluation of Objectives

7.9.1 Primary Effectiveness Objective

The primary effectiveness endpoint is the surgical success rate, defined as the procedure not going into conversion. Conversion is defined as the switch from a robotic-assisted approach using the Hugo™ system to a robotic-assisted approach utilizing an FDA cleared robotic-assisted device, laparoscopic or open surgery.

7.9.1.1 Hypothesis

The primary effectiveness hypothesis is to test if the surgical success rate is above the performance goal. Based on the Clinical Data Analysis of Robotic-Assisted Surgical Procedures in Urology Rev B (Medtronic on file), henceforth referred to as the Urology Clinical Data Analysis and clinical practice, a surgical success rate of 90% or higher is expected. To account for natural variability between sites, surgeons, and surgery types an overall performance goal of 85% (based on the literature) is pre-defined to evaluate the surgical success rate. Let *P* be the surgical success rate in this study. The statistical hypothesis is as follows:

$$H_0: P \leqslant 85\% \quad vs. \quad H_a: P > 85\%$$

7.9.1.2 Endpoint Definition and Derivation

Sites are asked if the procedure was completed according to the surgical plan (Yes/No). If the answer was "No", they had to specify what was not done per the surgical plan. If the site specifies "Converted" as what was not done per the surgical plan, then the surgery is identified as a conversion. Surgical success is defined as a surgery that has not been specified as a conversion. The proportion of subjects with a successful surgery needs to meet a performance goal of 85%.

7.9.1.3 Performance Requirements

The null hypothesis will be rejected if the p-value for the one-sided hypothesis test is less than 0.025 or, equivalently, the one-side 97.5% lower confidence bound is greater than 85%.

7.9.1.4 Rational for Performance Criteria

Analyses were performed based on available literature. Point estimates along with 95% and 99% predicted intervals were calculated to provide plausible ranges of potential effect sizes. Clinically meaningful margins and statistical precisions have been taken into account to determine the appropriate performance goals.

The mean effect size for conversions is 0.024 with a 95% confidence interval of 0.019 to 0.032. The 95% prediction interval is 0.004 to 0.125. The 99% prediction interval is 0.003 to 0.20. A performance goal of 85% for surgical success rate was set, which is equivalent to 0.15 as the conversion rate (1-success rate), aligned with the 95% to 99% of predictive intervals upper limit of 0.125 to 0.2. A single-group design will be used to test whether the proportion is greater than 0.85 (H_0: $P \leqslant 0.85$ vs. H_1: $P > 0.85$).

7.9.1.5 Analysis Methods

The surgical success rate will be calculated as the proportion of subjects with a successful surgery, with a one-sided 97.5% confidence lower bound. This proportion will be tested against 85% using a Clopper-Pearson binomial exact test. The confidence lower bound needs to be greater than 85%, the p-value must be less than 0.025, to reject the null hypothesis.

7.9.1.6 Determination of Subjects/Data for Analysis

The analysis of the primary effectiveness will be based on the FAS but following the mITT principle for the conversion endpoint by excluding any subject with that endpoint missing.

7.9.1.7 Supporting Analyses

7.9.1.7.1 Poolability Analysis

An assessment of data poolability of the sites will be performed using the logistic regression with the outcome being surgery success (1=successful surgery, 0=converted surgery). A p-value of 0.15 or less will be considered significant (per FDA recommendation). Sites with fewer than five subjects will be combined into larger sites to ensure statistical robustness. If the sites are found to be significantly heterogeneous with respect to the primary endpoint, additional analyses will be conducted to further assess variations across sites in baseline and procedural variables that might contribute to the variations.

7.9.1.7.2 Sensitivity Analysis

If any primary effectiveness endpoint data are missing, up to two sensitivity analyses will be performed.

The first sensitivity analysis will use multiple imputation (MI) to impute missing outcome data if the percentage of missingness is greater than 5%. The model variables used for the MI analysis may include, but are not limited to: study site, age, gender, and indication for surgery. The fully conditional specification method with 10 burn-in iterations and 10 repetitions (M=10) will be used for imputation if MI is utilized.

Another sensitivity analysis will be a tipping point analysis with the missing outcome data. In this analysis, any missing primary outcome data will be imputed over a range of possible scenarios to determine the

proportion of subjects with missing data that would need to be converted to another form of surgery (e.g., open or laparoscopic) to change the primary effectiveness endpoint.

A listing of subjects with missing data, along with reasons for missing data, will be provided.

7.9.1.7.3 Sub-group Analysis by Surgery Type

The surgical success rate for each surgery type will be calculated as the proportion of subjects with a successful surgery, with a one-sided 97.5% confidence lower bound. Additionally, a summary of type of conversion (open, laparoscopic, or other RAS) will be provided.

7.9.2 Primary Safety Objectives

The primary safety objective of this study is to confirm that the Hugo™ RAS System is safe when used for urologic robotic assisted surgery. The primary safety endpoint is the rate of subjects with major complications (meeting Grade III criteria or higher per the Clavien-Dindo Classification system) from the first incision through 30 days post-procedure.

7.9.2.1 Hypothesis

The primary safety hypothesis is to test the major complication rate (Clavien-Dindo Grade III or higher) through 30-days post-procedure against a performance goal. Since the safety profiles (i.e., expected major complication rates) are very different among the three urologic procedures (nephrectomy, prostatectomy, cystectomy), a separate performance goal is pre-specified for each of the three procedures (as per FDA recommendation). The performance goals are determined based on published literature data as summarized in the Urology Clinical Data Analysis.

Let R be the 30-day major complication rate, and PG be performance goal for a specific surgery group. The statistical hypothesis is formulated as follows:

$$H_0: R \geqslant PG \quad vs. \quad H_a: R < PG$$

7.9.2.2 Endpoint Definition and Derivation

Rate of major complications within 30 days is calculated as the proportion of subjects who experience an AE with Grade III Clavien-Dindo classification or higher that is related to any of device, procedure, or underlying condition/disease prior to or during the 30 calendar-days post-procedure. The procedure date, AE start date, CEC-adjudicated relatedness (device, procedure, underlying condition/disease) and Clavien-Dindo Classification collected with the adverse event will be used to determine whether a related major complication was experienced within 30 days. The proportion of subjects experiencing a major complication must meet a performance goal based on the indication for surgery.

7.9.2.3 Performance Requirements

The null hypothesis will be rejected if the p-value for the one-sided hypothesis test is less than the multiplicity-adjusted alpha as described in **Section 7.5**. The performance goal for each indication is as follows:

Nephrectomy: 20%

Prostatectomy: 20%

Cystectomy: 45%

7.9.2.4 Rational for Performance Criteria

The literature search was performed based on available literature for each of the three surgery types (nephrectomy, prostatectomy, cystectomy). Point estimates along with 95% and 99% predicted intervals were calculated to provide plausible range of potential effect sizes. Clinically meaningful margins and statistical precisions have been considered to determine the appropriate performance goals. In some situations, considerations were also given to align with the studies from comparable sample sizes.

The full text detailing the performance goal rationales and justification for each is shown in the IDE CIP, Section 14.4.2.

7.9.2.5 Analysis Methods

The proportion of subjects experiencing a major complication will be calculated, with a one-sided 97.5% confidence upper bound. This proportion will be tested against the performance goals in **Section 7.9.2.3** using a Clopper-Pearson binomial exact test. The null hypothesis will be rejected if one-sided p-value is significant comparing to an alpha determined after multiplicity adjustments described in **Section 7.5** (Holm's method, i.e. Holm-Bonferroni method)[1].

7.9.2.6 Determination of Subjects/Data for Analysis

Subjects from the FAS will be included in this analysis.

7.9.2.7 Supporting Analysis

Two sensitivity analyses may be performed.

If primary safety endpoint data are missing for more than 5% of cases in the FAS, multiple imputation (MI) will be used to impute the missing outcome data. The model variables used for the MI analysis may include but are not limited to study site, age, and gender. A summary of the variables included, and the related model selection process will be provided. The fully conditional specification method with 10 burn-in iterations and 10 repetitions (M=10) will be used for imputation if MI is utilized.

Another sensitivity analysis will be to evaluate the primary safety endpoint using the PPAS.

Three exploratory analyses are planned for this endpoint as well, wherein the rate of major complications within 30 days will be limited to just the related to device, related to procedure, and related to underlying condition. Relatedness in this exploratory analysis will be any rating other than "Not Related" or equivalent.

7.9.3 Secondary Objectives

Descriptive statistics will be used to summarize secondary endpoints. In addition, performance goals are pre-specified based on the Urology Clinical Data Analysis, as well as considerations of clinically meaningful margins and statistical confidences and 95% confidence intervals might be provided as appropriate.

The role of the secondary endpoints is to explore additional effects and outcomes associated with either the procedure or the disease state. The pre-specified performance goals serve as benchmarks for interpretation rather than thresholds for statistical significance of clinical trial success in this study.

Statistical evaluations for secondary objectives with performances goals are not powered in this study. Therefore, the secondary endpoints are planned to be analyzed descriptively with no plans for statistical testing.

7.9.3.1 Objectives, Endpoint Definitions and Derivations

Endpoint #	Objective	Endpoint Definition and Derivation
1	Overall Complication Rate	Rate of overall complications within 30 days is calculated as the proportion of subjects who experience any related (device, procedure, or underlying condition/disease) adverse event prior to or during the 30 calendar-days post-procedure. The AE start date collected with the adverse event will be used to determine whether a complication was experienced.
2	Operative Time	Operative time is defined as the difference in time in minutes between the procedure end time and the procedure start time. The procedure start time is the time of first incision and the end time is the time of skin closure.
3	Intraoperative Estimated Blood Loss	The estimated amount of blood loss is reported during the procedure in mL.
4	Transfusion Rate	Occurrence of a blood transfusion is captured as an adverse event. A blood transfusion is defined as an AE that resulted in any treatment and then a bleeding complication led to a blood transfusion and started prior to or during the 30 calendar-days post-procedure.
5	Device-related Conversion	A conversion is defined the same way as in **Section 7.9.1.2**. A device related conversion is defined as a conversion where either the primary or any other reason for conversion is denoted as Adverse Event that is related to the device or procedure, Device Deficiency, Non-Subject Adverse Event, or Poor Visualization.

6	Hospital Length of Stay	The length of hospital stay is the number of days that the subject was in the hospital from the start of the surgical procedure to hospital discharge. The number of days in the hospital can be found as the difference in minutes between the discharge date and time and the procedure date and time divided by 1440 min/hour, i.e., (Discharge Date and time – procedure date and time)/1440.
7	Readmission Rate	A readmission in a subject is identified if a related (device, procedure, or underlying condition/disease) adverse event with a start date of or after the index procedure date and prior to or during the 30-calendar days post-procedure that resulted in the subject being hospitalized. Additionally, the hospital start date reported within the AE must be prior to or during the 30 calendar-days post-procedure and on or after the reported index procedure discharge date.
8	Reoperation Rate	Reoperation through 30 days is identified in a subject if they experience a related (device, procedure, or underlying condition/disease) adverse event that resulted in treatment with a surgical procedure and the AE start date was prior to or during the 30-calendar days post-procedure.
9	Mortality Rate	Death through 30 days is identified as an AE that led to death which occurs during the procedure or up to the 30 calendar-days post-procedure.
10	Device Deficiencies	A subject is identified as experiencing a device deficiency if a device deficiency occurred before, during, or after the subject's procedure.
11	Overall Survival	Subject death will be identified as an adverse event that leads to death where the adverse event start date is on or after the procedure date or if the subject is noted by the site to have died during follow-up. Time to death will be calculated as the difference in days between death and the procedure date.
12	Progression-free Survival	Disease progression will be identified in a subject if their disease state is noted by the site, during follow-up, as "disease reoccurrence, disease progression", or if subject death had been noted. This is assessed as either at the time of the follow-up visit or at the time of last contact, whichever is earlier will be denoted as the time of disease

		progression. Time to disease progression is calculated as the difference in days between disease regression and the procedure date.
13	Disease-free Survival	Disease reoccurrence will be identified in a subject if the subject disease state is noted by the site, during follow-up, as "disease reoccurrence, disease progression", "disease reoccurrence, disease stable", or if subject death had been noted. This is assessed either at the time of a follow-up visit or at the time of last contact, whichever is earlier will be denoted as the time of disease reoccurrence. Time to disease reoccurrence will be calculated as the difference in days between disease reoccurrence date and the procedure date.

7.9.3.2 Hypotheses

Secondary objectives are not powered for statistical significance, no hypothesis testing will take place.

7.9.3.3 Performance Requirements

Performance goals for each of the secondary objectives can be found in **Section 10**.

7.9.3.4 Rational for Performance Criteria

Rationale for each of the performance goals for secondary objectives can be found in **Section 10**.

7.9.3.5 Analysis Methods

Secondary objectives 1, 4, 5, 7, 8, 9 and 10 will be summarized by surgery as the count and proportion of subjects experiencing the given endpoint, with a 95% confidence interval.

Secondary objectives 2, 3 and 6 will be summarized by surgery performed using mean, standard deviations (SD), median, IQR, and range. A 95% confidence interval will also be calculated.

Secondary objectives 11, 12 and 13 will be evaluated through separate Kaplan-Meier analyses by surgery.[2] The event for each Kaplan-Meier will be their respective objective endpoint. Subjects will be censored upon study completion (5-year follow-up visit) or they are lost to follow-up without event. For subjects that are lost to follow-up, their last visit date or date of last subject contact will be used as their censor date, whichever occurred later. The 5-year survival will be estimated from the Kaplan-Meier model and a 95% confidence interval will be calculated.

Secondary objective 1 will have three additional exploratory analyses wherein the overall complication rate within 30 days will be limited to each of: related to device, related to procedure, and related to underlying condition. Relatedness in this exploratory analysis will be any rating other than "Not Related" or equivalent.

7.9.3.6 Determination of Subjects/Data for Analysis

Subjects from the FAS with data available will be included for each secondary analysis. Secondary endpoint 2 will add the additional requirement that it will be for subjects who did not undergo conversion.

7.9.4 Ancillary Objectives

Descriptive analyses of ancillary objectives will be exploratory in nature and are not intended as a focus of the study for the evaluation of the study device.

7.9.4.1 Rate of Negative Surgical Margins

The surgical margin at the 30-day follow-up will be summarized. The positive margin rate and the margin, recorded for each oncological subject, will be summarized using mean (SD), median, minimum, and maximum in subjects from the FAS with complete data at the 30-day follow-up. A lower value for these outcomes will indicate a more positive outcome.

7.9.4.2 Lymph Node Yield

The proportion of lymph nodes suspected to contain cancer cells at the 30-days follow-up will be calculated as the total positive lymph nodes divided by the total number of lymph nodes retrieved. This proportion will be summarized using mean (SD), median, minimum, and maximum in oncological subjects from the FAS with complete data at the 30-day follow-up. A lower value for these outcomes will indicate a more positive outcome.

7.9.4.3 Warm Ischemia Time

Warm ischemia time captured during the procedure will be summarized in nephrectomy subjects from the FAS. The time in minutes will be summarized using mean (SD), median, minimum, and maximum. A lower value for these outcomes will indicate a more positive outcome.

7.9.4.4 Surgeon Experience

Responses to the surgeon experience survey will be summarized for all surgeons who complete the survey. Number and proportion of surgeons who have “No”, “Moderate”, or “Constant” problems with the communication, coordination, equipment, and training will be summarized. Additionally, the ergonomic assessment will be summarized using the number and proportion of surgeons who responded “Yes” or “No” to each question in the survey.

7.10 Safety Evaluation

In addition to the primary and secondary safety endpoints, other safety measures will be assessed for the FAS population as well. These include data of adverse events and device deficiencies (DD).

Adverse Event definitions used in this study are based on ISO 14155:2020 (Clinical Investigation of Medical Devices for Human Subjects -- Good Clinical Practice). DD is defined as any inadequacy of a medical device with respect to its identity, quality, durability, reliability, usability, safety, or performance. This includes malfunctions, use errors, and inadequacy in the information supplied by the manufacturer such as labeling.

For a summary of adverse events (AEs), the frequencies of adverse device effects (ADEs), serious adverse events (SAEs), serious adverse device effects (SADEs) and unanticipated adverse device effects (UADEs) will be presented in the CSR in tabular format as well as listings. AEs relatedness will be summarized and presented. For AEs related to procedure and the investigational device Hugo™ system, a detailed summary will be presented by MedDRA preferred term and system organ class (SOC).

DDs will be presented in summary tables displaying the number of deficiencies, and the number and percentage of subjects with deficiencies.

7.11 Health Outcomes Analyses

This section is not applicable to this study.

7.12 Changes to Planned Analysis

A clarification on the sensitivity analyses for the primary endpoints has been provided. The tipping point analysis is for the primary effectiveness endpoint only and will not be performed for the primary safety endpoint. For the primary safety endpoints, since the endpoint would require hospitalization, the likelihood of missing data is low. Lastly, multiple imputation will only be used if 5% or more of the data is missing in each case.

8. Validation Requirements

Statistical outputs will be validated by Level I or Level II validation, the activities required for each are defined in internal standard operating procedures (SOPs). The primary objectives analysis will be validated by Level I. The analyses of the secondary, ancillary objectives, as well as baseline demographics, AE and PD summaries will be validated by at least Level II.

Level I: The peer reviewer independently programs output and then compares the output with that generated by the original Statistical Programmer.

Level II: The peer reviewer reviews the code; where appropriate, performs manual calculations or simple programming checks to verify the output.

9. References

1: Holm, S. (1979). A Simple Sequentially Rejective Multiple Test Procedure. Scandinavian Journal of Statistics, 6, 65-70.

2: Kaplan, E. L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. J. Amer. Statist. Assoc. 53 (282): 457–481.

10. Statistical Appendices

10.1 Performance Goals for Secondary Endpoints

Please refer to Section 14.5 of the clinical investigation protocol for the secondary endpoint performance goals and associated justification statements.