STATISTICAL ANALYSIS PLAN FOR LUMICELL PROTOCOL NUMBER CL0007r08

Version Date: 17NOV2021

Study Title: PIVOTAL STUDY OF THE LUM IMAGING SYSTEM FOR ASSISTING INTRAOPERATIVE DETECTION OF RESIDUAL CANCER IN THE TUMOR BED OF FEMALE PATIENTS WITH BREAST CANCER

System: LUM Imaging System (LUM015 imaging agent and LUM Imaging Device)
IDE # G140195

Jorge Ferrer, Ph.D.
jmferrer@lumicell.com
Phone: 617-571-0592

Manna Chang
mchang@lumicell.com

Gheorghe Doros
Statistical Consultant
gheorghe.doros@gmail.com

TABLE OF CONTENTS

1 Introduction

This document describes the statistical analysis plan to be performed as part of the pivotal study of the Lumicell Imaging System under protocol number CL0007r08. The objective of this prospective, multi-center, single-arm study is to demonstrate the safety and efficacy of the LUM Imaging System (LUM015 imaging agent in conjunction with the LUM Imaging Device and decision software), in identifying residual cancer in the lumpectomy bed of female breast cancer patients. For safety assessment, we propose two analyses: one analysis that pools all the patients that have received administration of LUM015 across all the clinical trial protocols including non-breast cancer indication, and another analysis for just breast cancer patients.

1.1 Considerations for integrated safety analysis and integrated efficacy analysis

For the pivotal trial clinical study report, safety data will be presented from the 406 patients enrolled in the study. However, for the Integrated Summary of Safety (ISS for NDA submission) and the Summary of Safety and Efficacy (SSE for PMA submission) we plan to include safety data from patients that have been injected with LUM015 across multiple indications. This integrated safety analysis will be detailed in the statistical analysis plan for the ISS.

2 Analysis Populations

- Safety (SAF) Population: the safety population includes all subjects who received study drug (LUM015).
- Modified Intent To Treat (mITT) population: includes the subjects in SAF, but excluding subjects who are not able to be imaged with the LUM Imaging System. That is, the mITT population includes all patients imaged with the LUM Imaging Systeme and will be used as the primary analysis population for the three co-primary efficacy endpoints and efficacy analysis of the secondary endpoints.
- Per-Protocol Population: Includes mITT subjects who complete all study evaluations and are without any major protocol deviations that may impact data collection or efficacy analysis. A sensitivity analysis of the primary and secondary efficacy endpoints will be conducted using the per-protocol analysis set.

3 Study endpoints

The pivotal trial will have three co-primary endpoints and other secondary endpoints, which are summarized in the next sections.

3.1 Definition of positive margins

For this study, positive margins are defined using the latest consensus from the Society of Surgical Oncology as follows:

- For invasive cancer with or without associated ductal carcinoma in situ (DCIS): cancer cells or DCIS present on ink [1]
- For pure DCIS lesions: DCIS present less than 2 mm from the ink [2]. Note: when microinvasive cancer is reported, the positive margin criterion will be the same as for pure DCIS instead of the criterion for invasive cancer per NCCN Clinical Practice Guidelines in Oncology.

3.2 Three co-primary endpoints

3.2.1 Removal of residual cancer

Ratio of patients who have **residual cancer** found in at least one LUM-guided shave (also known as therapeutic shave or "T-shaves") among all patients randomized to the device arm. **Residual cancer** is defined as tumor found by pathology in a therapeutic shave after the SOC surgical procedure is completed; that is, tumor that current SOC surgery failed to remove (Figure 1). Mathematically, it is defined as:

$$\frac{\#\ patients\ with\ residual\ cancer\ found\ in\ at\ least\ one\ therapeutic\ shave}{Total\ number\ of\ patients}$$

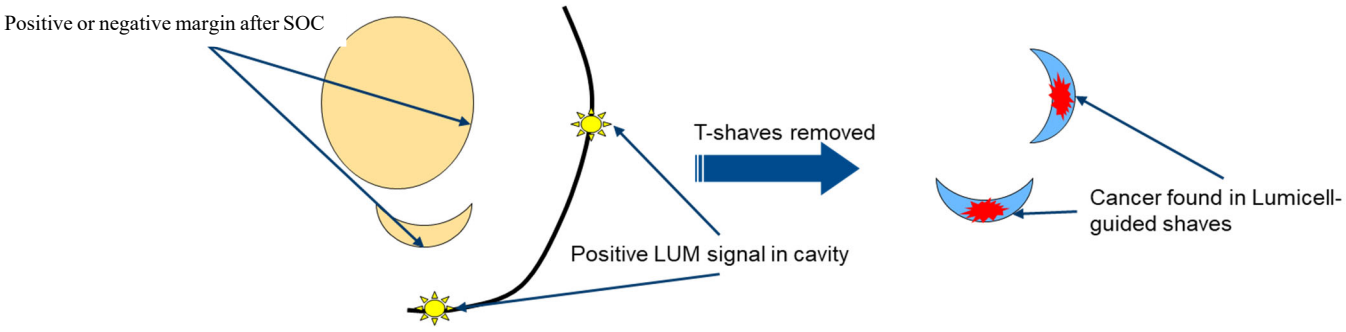

Figure 1: Depiction of primary endpoint

This proposed primary endpoint for determination of whether tumor is found in at least one therapeutic shave, the **ground truth** of histopathology assessment of the therapeutic shave is used.

3.2.2 Sensitivity and specificity

The instrument diagnostic accuracy is measured by sensitivity and specificity on a per-tissue basis. Because a LUM-guided shave does not exist when a negative LUM image is indicated we propose a hierarchical approach to determine the truth standard as depicted in Figure 2. The highest truth standard is whether pathology finds cancer in a tissue shave removed from the area that was imaged in the cavity. When a shave does not exist, then whether tumor is found in a re-excision surgery is considered as the truth standard. When neither a shave nor a second surgery exists for a given orientation that was imaged, the margin assessment of that orientation is used as truth standard. We believe that the hierarchical approach outlined to determine the truth standard is the most meaningful and practical assessment of device accuracy to predict the presence of cancer in the cavity.

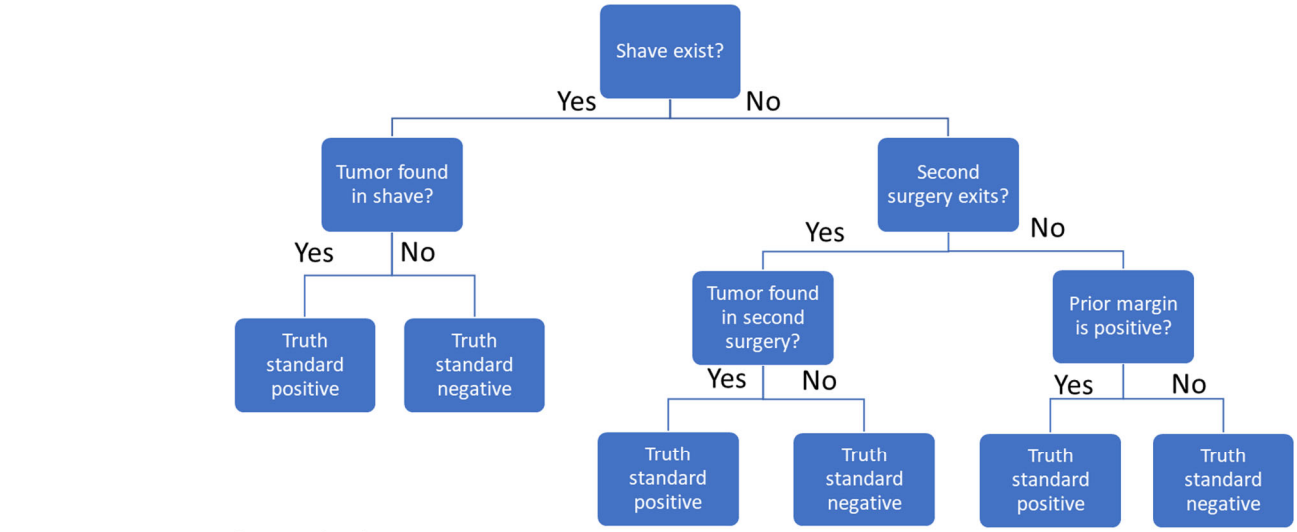

Figure 2: Definition of truth standard to evaluate the LUM Imaging System diagnostic performance.

When a therapeutic shave and re-excision are not available for a given orientation, the margin from the prior resected tissue of the corresponding orientation will be used as ground inferred truth for evaluation (Figure 3). The number of prior positive margins contributing to the ground inferred truth for a patient will not exceed the number of orientations with positive margins.

The orientation of the tissue removed during a re-excision is typically noted in the pathology report and matches the orientation in which a positive margin was identified from the initial surgery. However, based on our prior experience there will be instances in which the orientation of the tissue removed during a re-excision is not specified. In cases of unclear orientation of the tissue removed during re-excisions, it will be assumed that the re-excision was intended to address only the orientations with positive margins and pathology findings from the re-excsions (tumor found/not found) will be applied to those orientations. Also, if there are two consecutive shaves during the re-excision from overlapping orienations, they will be considered as one.

If a mastectomy is performed as a follow-up surgery and no cancer is found in the mastectomy specimen, all the orientation will be considered as cancer negative for truth standard. If a mastectomy is performed and cancer is found but there is no indication on the orientation in which the cancer was found, the orientation in which a positive margin was identified in the initial lumpectomy will be considered as cancer positive for truth standard unless otherwise noted.

The diagnostic performance will be computed by using the hierarchical approach to populate the 2x2 contingency table (Table 1). The Youden Index will be reported but is not hypothesis-tested.

Table 1: Definitions for 2x2 contingency table.

	Hierarchical Truth standard (+)	Hierarchical Truth standard (–)
LUM positive cavity signal (+)	True Positives	False Positives
LUM negative cavity signal (-)	False Negatives	True Negatives

$$sensitivity = \frac{True\ Positives}{True\ Positives + False\ Negatives}$$

$$specificity = \frac{True\ Negatives}{True\ Negatives + False\ Positives}$$

$$Youden\ Index = sensitivity + specificity - 1$$

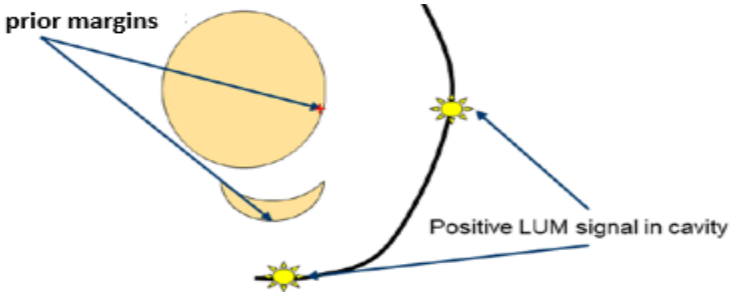

Figure 3: Depiction of the prior margin assessment for the LUM signal in the cavity when no additional tissue from the cavity is available.

3.2.3 Additional sensitivity and specificity analysis

In the pivotal study protocol, images of the cavity are being recorded with the LUM Imaging System before the surgeon removes any SOC shaves. These images are saved with the tumor detection inactive to prevent potentially influencing surgeon's decision to remove the SOC shave. This procedure was implemented to acquire additional data for a potential future use of the LUM Imaging System before removing SOC shaves. However, this procedure does not represent the initial intended use of the LUM Imaging System, more so with the tumor detection algorithm turned off. The initial intended use of the LUM Imaging System is after the SOC procedure including removal of SOC shaves, is completed. Thus, the primary analysis for sensitivity and specificity to support the co-primary endpoints will be done with imaging data collected after the SOC procedure is completed, that is, based on therapeutic shaves, second surgery and prior margin using a hierarchical approach as described in Table 1. As part of the pivotal trial analysis, a secondary evaluation of sensitivity and specificity will include the results from the LUM imaging of the cavity prior to removing SOC shaves.

3.3 Secondary endpoints

a. Proportion of patients with positive margins after standard of care breast-conserving surgery who have a LUM Imaging System signal in the cavity above the threshold as defined by the tumor detection algorithm (i.e., a positive LUM signal). The endpoint is mathematically defined as:

$$\frac{\#\ positive\ margin\ patients\ after\ SOC\ with\ LUM\ signal\ above\ the\ threshold}{\#positive\ margin\ patients\ after\ SOC}$$

If a subject has more than one orientation with an SOC positive margin, LUM imaging must correctly identify all orientations with an SOC positive margin for success of this endpoint.

b. Proportion of subjects with pathology-positive margins after standard of care breast-conserving surgery for whom additional LUM Imaging System-guided shaves resulted in pathology-negative margins. Note: This endpoint also estimates the potential reduction in re-excision surgeries following standard of care breast-conserving surgery that resulted from a positive SOC margin. This is estimated as:

$$\frac{\#\ positive\ margin\ patients\ after\ SOC\ with\ all\ final\ negative\ margins\ after\ the\ LUM\ procedure}{\#positive\ margin\ patients\ after\ SOC}$$

c. Proportion of subjects with pathology-positive margins after standard of care breast-conserving surgery for whom additional LUM Imaging System-guided shaves resulted in pathology-negative margins among all patients. This is estimated as:

$$\frac{\#\ positive\ margin\ patients\ after\ SOC\ with\ all\ final\ negative\ margins\ after\ the\ LUM\ procedure}{Total\ number\ patients}$$

d. Ratio of patients with negative margins after the SOC procedure who have residual cancer found in at least one Lumicell-guided (or therapeutic) shave among patients with negative margins. This is estimated as:

$$\frac{\#\ negative\ margin\ patients\ after\ SOC\ with\ residual\ cancer\ found\ in\ at\ least\ one\ therapeutic\ shave}{\#negative\ margin\ patients\ after\ SOC}$$

e. Ratio of patients with negative margins after the SOC procedure who have residual cancer found in at least one Lumicell-guided (or therapeutic) shave among all patients

$$\frac{\#\ negative\ margin\ patients\ after\ SOC\ with\ residual\ cancer\ found\ in\ at\ least\ one\ therapeutic\ shave}{Total\ number\ of\ patients}$$

f. Mean and median incremental volume (in cubic centimeters) of tissue removed from therapeutic shaves: For each patient, absolute shave volume will be calculated as length (c) × width (c) × depth (c) for each therapeutic shave with the sum by patient of all therapeutic shaves. Subjects with no therapeutic shaves taken will have a therapeutic shave volume of 0cc.

g. Mean and median contribution of therapeutic shave volume to total tissue removed.

h. Average number of SOC and therapeutic shaves taken per subject overall and by type of SOC lumpectomy procedure (i.e., lumpectomy with comprehensive shaves vs. lumpectomy with or without selective shaves).
i. Number of second surgeries recommended as a result of final positive margins.
j. Rate of second surgeries as a result of a positive margin status following SOC (control arm) and LUM Imaging procedures (device arm).
k. Rate of cancer found in second surgeries
l. Number of second surgeries for each patient
m. Number of images per subject from the cavity after SOC and first round of therapeutic shaves
n. Secondary analysis of sensitivity and specificity including imaging before removal of SOC shaves.
o. Number of device issues and malfunctions and their impact to data capture.
p. Patient-level sensitivity and specificity
q. Collect exploratory data on tissue types found in therapeutic shaves.
r. Collect exploratory data on patient reported outcomes.
s. Analysis of tissue-level sensitivity and specificity for the SOC procedure based on the outermost SOC resected surface.
t. Adverse events stratified by severity and relatedness to drug/device.
u. Serious adverse events stratified by severity and relatedness to drug/device
v. Adverse events by preferred term (sorted by descending occurrences for each preferred term).
w. Summary of adverse events, overall and split by expectedness to drug/device.
x. Comparison of adverse events in subjects that have at least one therapeutic shave versus subjects that have no therapeutic shaves.

4 Statistical methods

4.1 Demographics and baseline characteristics

Demographic and baseline characteristic including: age, race and ethnicity, menopausal status, mammographic breast density, palpability, cancer histologic type, tumor size, node positive patients, and tumor receptor status (ER/PR/HER2), and baseline body mass index (BMI) calculated from weight and height will be summarized for the study sample. Statistics for continuous variables will include mean, median, standard deviation (SD), minimum, maximum and sample size for the overall sample and grouped by control and device arms, and two-sided 95% CI of the mean difference between (control and study arms). Binary variables will be described with frequencies and percentages for the overall sample. Demographic and baseline characteristics will be summarized by overall using descriptive statistics for analysis populations.

4.2 Residual cancer removal (co-primary endpoint)

The null hypothesis is that the true percentage of patients in whom at least one LUM-guided shave removed after the SOC procedure contains cancer as confirmed by pathology examination is smaller than or equal to 3%:

$$H_0: p \leq 3\%$$

The alternative hypothesis is that the true percentage of patients in whom at least one LUM-guided shave removed after the SOC procedure contains cancer as confirmed by pathology examination is larger than 3%:

$$H_A: p > 3\%$$

where p is the true percentage of patients in whom at least one LUM-guided shave removed after the SOC procedure contains cancer as confirmed by pathology examination.

The hypothesis will be tested using an exact Binomial test at a one-sided significance level of 2.5%. The true proportion p will be estimated along with a two-sided 95% exact confidence interval. The null hypothesis H_0 is rejected if the lower bound of the confidence interval for removal of residual cancer sensitivity is larger than the performance goal of 3%.

Performance goal selection for removal of residual cancer endpoint

To choose the success criteria for the proportion of patients who have residual cancer found in at least one LUM-guided shave, we used published results for estimates of local recurrence before and after adjuvant radiation, assuming that most local recurrences are a consequence of unresected cancer during the initial surgery. The meta-analysis conducted on approximately 28,000 patients to support the consensus guideline for margin in 2014 from SSO-ASTRO, reported an overall recurrence rate of 5.3% for patients undergoing lumpectomy and receiving whole breast radiation therapy (including patients with positive and negative margins) [1]. Other investigators have published local recurrence rates for patients with negative margins at 10.2% without a radiation boost, while 6.2% for those with radiation boosts [3]. Based on this data, we estimate that radiation addresses approximately 50% of residual cancer, but not all. Our Phase C results shows a proportion of 26/230 or 11.3% [95% CI: 7.5%-16.1%], supporting the performance goal selection of this endpoint.

Sample size for testing residual cancer removal

Our preliminary analysis of the Phase C exploratory data indicated that LUM detected residual cavity cancer in 11% of patients (26/230). To power the study, we are using a slightly more conservative assumption of a ~9% true percentage of patients in whom at least one therapeutic shave has cancer confirmed by pathology. With this assumption, enrolling 220 evaluable subjects, a binomial exact test for proportions at a one-sided significance level of 2.5%, will have 90% power to declare the percentage of patients in whom at least one therapeutic shave has cancer confirmed by pathology to be larger than 3%.

4.3 **Per tissue sensitivity and specificity (co-primary endpoints)**

4.3.1 Analysis methods and intrapatient correlation

Three separate analyses will be conducted for sensitivity and specificity assessment. The first analysis uses the Binomial estimator and the other two approaches address potential intra patient correlations.

- Binomial estimator: the probability and the confidence interval are estimated according to binomial distribution

- Generalized Estimating Equation (GEE): This method is used to analyze correlated data which is modeled with generalized linear model having binomial link function, and including correlation cluster and matrix information (here within each subject). Procedure Genmod in SAS is used to apply this method (https://support.sas.com/rnd/app/stat/topics/gee/gee.pdf). This method with a compound symmetry working correlation structure, will be used to estimate sensitivity and specificity along with a 97.5% lower bounded one-sided confidence interval (CI) (* Tessa S. S. Genders, Sandra Spronk, Theo Stijnen, Ewout W. Steyerberg, Emmanuel Lesaffre, M. G. Myriam Hunink. Methods for Calculating Sensitivity and Specificity of Clustered Data: A TutorialRadiology: Volume 265 (3):910 (2012)).

- Bootstrapping sampling method: This method is applied to sample the subjects with replacement at the same sample size of mITT for 10,000 rounds. The sampled subjects with corresponding imaging, and standard truth positive and negatives will be used to estimate the sensitivity and specificity. The median sensitivity and specificity will be used as the estimate from bootstrapping method. The 2.5% and 97.5% percentiles will be used as lower and upper bounds, respectively, for 95% confidence intervals.

4.3.2 Performance goals

The verification of the sensitivity and specificity co-primary endpoints will be a refutation of the null hypotheses

H_{01}^{s}: $sen \leq 40\%$ against the alternative H_{11}^{s}: $sen > 40\%$

H_{02}^{s}: $spec \leq 60\%$ against the alternative H_{12}^{s}: $spec > 60\%$

where sen and $spec$ represent the true sensitivity and specificity, respectively. The hypotheses involving sensitivity and specificity will be assessed based on the data collected per-tissue level.

These hypotheses will be tested using a one-sided significance level of 2.5%. To account for within-subject correlation, Generalized Estimating Equations (GEE), with a compound symmetry working correlation structure, will be used to estimate sensitivity and specificity along with an equi-tail 95% two-sided confidence interval (CI). The null hypothesis H_{01}^{s} is rejected if the lower bound of the confidence interval for sensitivity is larger than the performance goal of 40%. Similarly, the null hypothesis H_{02}^{s} is rejected if the lower bound of the confidence interval for specificity is larger than the performance goal of 60%.

Performance Goals selection for sensitivity and specificity endpoints

Data from the Phase C Feasibility Study showed pathology margin assessment sensitivity to predict cancer in the cavity of ~40%. However, this pathology assessment is done several days after the surgery and can lead to second surgeries. Because the LUM Imaging System provides the additional benefit of assessing the cavity in real-time, a performance goal for instrument sensitivity > 40% is proposed as the success criterion for this hypothesis-tested endpoint. We expect the sensitivity to be similar to that observed in the Phase C study of 69% [95% CI: 56%-80%].

The Phase C data shows that at the current specificity performance of 70% [95% CI: 68%-73%], an average of 1.1 additional shaves were removed guided by the LUM Imaging system, accounting for ~10% of all the total resected tissue. We selected a performance goal for specificity of > 60% to ensure that similar performance in the pivotal trial is obtained regarding tissue removal. Also, with this lower bound, a Youden Index > 0 will be demonstrated to indicate non-random diagnostic performance.

Sample size for testing sensitivity:

As sensitivity and specificity are estimated at the per tissue level, within subject correlation is considered for the sample size estimation. Intracluster Correlation Coefficient (ICC) was estimated with data from the Feasibility Phase C study. The ICC value was obtained at 0.145 for the full phase data set and at 0.22 for the validation data set which the tumor detection algorithm is same as that used in the pivotal study. We selected ICC at 0.22 for sample size estimation.

The parameters for the simulation is as below:

- True sensitivity is 60%
- Intracluster Correlation Coefficient (ICC) value = 0.22
- Truth standard positive rate 4%
- Performance Goal (PG) = 40%
- 2.5% one-sided significance level
- 70 truth standard positive shaves

Assuming 70 truth standard positive shaves, under the alternative hypothesis, using Generalized Estimating Equation (GEE) model with a compound symmetry working correlation structure, will have >90% power to show that the true sensitivity is > 40%.

Sample size for testing specificity:

The parameters for the simulation is as below:

- True specificity is 65%
- Intracluster Correlation Coefficient (ICC) value = 0.22
- Truth standard negative rate 96%
- Performance Goal (PG) = 60%
- 2.5% one-sided significance level

Under the alternative hypothesis, using Generalized Estimating Equation (GEE) model with a compound symmetry working correlation structure, 1023 truth standard negatives are needed to reach >90% power to show that the true specificity is >60%.

4.4 Per patients sensitivity and specificity (secondary endpoint)

Based on feedback from FDA, Lumicell is including an analysis of sensitivity and specificity at the patient level. Because a single patient can have up to 6 lumpectomy cavity orientations, each orientation with its own image, there will be patients that can have truth standard negatives and truth standard positives within the cavity. Thus, we propose to evaluate the per-patient truth standards as described below.

Definitions for per-patient truth standards:

- Truth standard positive patient: a patient having at least one orientation as a truth standard positive using the per-tissue hierarchical approach
- Truth standard negative patient: a patient having all orientations as truth standard negatives using the per-tissue hierarchical approach.

Definitions for per-patient Lumicell imaging results:

- Positive detection patient by Lumicell imaging device: Lumicell device detects at least one positive signal in a patient
- Negative detection patient by Lumicell imaging device: Lumicell device does not detect any positive signal in a patient; that is, all Lumicell signals for that patient are negative.

Using these definitions, the per-patient 2x2 contingency table will be populated as:

- True positives: truth standard positive patients with positive detection by Lumicell imaging device
- True negatives: truth standard negative patients with negative detection by Lumicell imaging device
- False negatives: truth standard positive patients with negative detection by Lumicell imaging device
- False positives: truth standard negative patients with positive detection by Lumicell imaging device

4.5 Study sample size, event-driven design, and overall study power

The pivotal study is powered based upon targeting success for the three co-primary endpoints as outlined above. The efficacy endpoints will be assessed in the patients who undergo LUM-guided imaging (not the $1/10^{th}$ of subjects randomized to the control arm).

Based on Phase C data, it is expected that 268 patients will provide the necessary 70 truth standard positive events to power for sensitivity. Also, based on Phase C data, it is expected that 152 patients will provide the necessary 1021 truth standard negative events to power for specificity. To power for the removal of residual cancer endpoint, approximately 220 patients are required. Thus, the size of this study is driven by the necessary number of truth standard positive events.

Adding 15% to the sample size of 268 patients to include subjects in the control arm and some expected data loss, it is estimated that 310 total patients will be needed for the pivotal study. However, given the uncertainty of translating the number of truth standard positive events to actual number of patients, the pivotal trial is planned as event-driven clinical study. Patients will be enrolled until 70 truth standard positive events are reported, up to a maximum of 450 patients. The number of truth standard positive events will be counted based on the truth standard hierarchical approach from Figure 2. Counting of these events will be conducted via software programming to read from the electronic data capture system; no endpoints will be calculated at that time. Results from the event counting process will be accessible to the DSMB which will then inform Lumicell

personnel on whether the targeted number of events has been reached or if enrollment needs to continue.

With the proposed sample size from sensitivity, under the alternative hypothesis for specificity using the GEE model with a compound symmetry correlation structure, the power to show true specificity above the performance goal of > 60% will be >99%. For removal of residual cancer, the power is estimated at 96%. Thus, the overall study power is 90% x 99% x 96% = 86%.

4.6 Overall primary hypotheses

The overall primary hypotheses for the primary endpoint is

$$H_0^p \text{ (null)}: H_0 \text{ or } H_{01}^s \text{ or } H_{02}^s \quad \text{versus} \quad H_1^p \text{ (alternative)}: H_1 \text{and } H_{11}^s \text{ and } H_{12}^s$$

The removal of residual cancer, sensitivity and specificity will be tested at a one-sided significance level of 2.5%.

4.7 Methods for additional analyses

Secondary outcomes involving proportions will be analyzed using methods for Binomial distributions akin to the ones described above. Continuous measurement will be summarized using means and standard deviation and 95% confidence intervals. The percent of people with a specific AE will be reported in prespecified groups. The percent of subjects with adverse events in subjects that have at least one therapeutic shave will be compared with the percent of subjects with adverse events that have no therapeutic shaves using Fisher's exact test or goodness of fit test.

4.8 Safety analysis

Safety data will be reported as:

- Adverse events (AEs) stratified by severity and relatedness to drug/device.
- List of serious adverse events.
- Summary of AEs overall, by expectedness and by drug/device
- Adverse events coded per MedDRA guidelines
- Adverse events by preferred term per MedDRA (sorted by descending occurrences for each preferred term).
- Comparison of AEs in subjects that have at least one therapeutic shave versus subjects that have no therapeutic shaves
- Summary of AEs in system organ class and preferred terms (per MedDRA) and stratified by severity and relatedness to drug/device

All adverse events will be evaluated and assigned the "expectedness" by a Lumicell representative as described in the study plans or protocols and will be reported accordingly.

Two safety analyses sets will be completed: one for all the patients injected with LUM015 across all the clinical studies summarized in **Error! Reference source not found.**, and another one for just the breast cancer patients.

The safety variables of interest are listed below.

- AE and adverse device effects

- Complete blood count
  - Platelet Count
  - Red blood cell Count
  - Hemoglobin
  - Hematocrit
  - Mean corpuscular volume
  - Mean corpuscular hemoglobin
  - Neutrophils
  - Lymphocytes
  - Monocytes
  - Eosinophils
  - Basophils
  - White blood cell count
- Serum chemistry
  - Albumin
  - Alkaline Phosphatase
  - Total Bilirubin
  - Blood urea nitrogen
  - Calcium
  - Chloride
  - Creatinine
  - Estimated gromerular filtration rate (eGFR)*
  - Glucose
  - Potassium
  - Total Protein
  - SGOT (AST)
  - SGPT (ALT)
  - Sodium

*Lumicell will calculate eGFR based on CKD-EPI Creatinine Equation per Levey et al [4]:

$$eGFR = 141 \times \min\left(\frac{S_{cr}}{K}, 1\right)^{\alpha} \times max\left(\frac{S_{cr}}{K}, 1\right)^{-1.209} \times 0.993^{age} \times 1.018(if\ female) \times 1.159(if\ black)$$

Abbreviations / Units:

- eGFR = estimated glomerular filtration rate in mL/min/1.73 m^2
- S_{cr} = serum creatinine in mg/dL per data entered in the CRF
- K = 0.7 (females) or 0.9 (males)
- α = -0.329 (females) or -0.411 (males)

The resulting table with the calculated eGFR will be recorded as an external dataset.

4.9 **Reported device issues**

Device issues as reported or interpreted by the clinical site are captured in the eCRF. An external dataset will be created to evaluate and adjudicate by Lumicell personnel for impact to subject safety, impact to data capture, and categorization. Issues that resulted in an impact to data capture may be removed from the efficacy analysis. The 'Device Issues' external dataset will be used for analysis and not generated by the device issues reported in the eCRF.

4.10 Protocol Deviations

Protocol Deviations are identified by either Lumicell personnel or clinical sites and reported in the eCRF. An external dataset will be created to evaluate and adjudicate these events by Lumicell personnel as described in the study plans and procedures. The adjudication will include whether there was impact to data capture and integrity, and may be indicate data to be removed from the efficacy analysis. The Protocol Deviations external dataset will be used for analysis and not generated by the protocol deviations reported in the eCRF.

4.11 General considerations

All analyses will be performed under Good Clinical Practice (GCP) standards using a prespecified statistical analysis plan (SAP). Analyses will be conducted and tables and listings generated using Statistical Analysis System (SAS®) Version 9.4 or higher. Graphics may be prepared with SAS Version 9.4 or higher; R version 3.6.1 or higher, SigmaPlot 12.3 or higher. Subject level data will be provided as a listing to support tables and figures. Summaries will be provided overall, by SOC surgery type (lumpectomy/selective vs comprehensive), and by site.

4.11.1 Numerical Precision

For reporting of descriptive statistics, the mean, SD, median, min, max and quantile will be presented to 1-digit precision. The minimum, median, and maximum will be presented to the same precision as the source data. The maximum number of decimals will be two, no matter how precise the source data is. The percentages will be reported to 1 decimal (format of xx.x%). Sensitivity and specificity will be reported as percentages with 1-digit precision (format of xx.x%). P-values will be reported to three (3) decimal places or as < 0.001.

4.11.2 Multiple images to cover entire shave or multiple shaves for one image

To accurately evaluate the study endpoints, it is critical that images are collected from the entire lumpectomy cavity surface. To achieve this, multiple images may be required to record the LUM assessment for an entire orientation (e.g. 2 images to cover the entire superior orientation) but a single shave may exist. When multiple images exist for a given orientation or shave, the combination of the images will be considered as one single output as described below:

- All images for a given orientation or shaved region are LUM negative, then the LUM result for the entire orientation or shaved region is considered as LUM negative.
- If one or more of the images for a given orientation or shaved region is LUM positive, then the LUM result for the entire orientation or shaved region is considered as LUM positive.

When there are more than one shave for a single image, the pathology finding for the combination of the shaves will be considered as one as follows:

- If no tumor is found on the shaves, then the combination of the shaves is considered as truth standard negative.
- If any of the shaves has tumor, then the combination of the shaves is considered as truth standard positive.

4.11.3 Missing and spurious data

During LUM-guided imaging, when there is not enough tissue to be removed (e.g. too close to skin or too close to the chest wall) no LUM image shall be recorded because it does not generate an actionable result.

In the case where the LUM Imaging System indicates that a therapeutic shave needs to be removed but the surgeon decides not to follow this guidance, the reason for not doing so will be recorded in the patient's CRF or during imaging. These instances are considered as a protocol deviation. There are other cases when the LUM-guide shave reach the protocol limit of 2 guided shaves per orientation, so a third guided shave will not be performed. In these cases, the data will be included in the performance evaluation as is; that is, because these will not generate a LUM-guided shave they will not be evaluated as removing cancer or converting positive margins to negative marginsn. However, they will be evaluated for the per-tissue sensitivity and specificity calculation according the hierarchical approached described in section 3.2.2.

In occasions, a cavity image may be recorded from an unintended region or the image may be deemed as low quality (blurry, out of focus, etc.). For example, the surgeon realizes that he/she placed the device in the wrong location within the cavity after recording an image. Since the LUM software does not allow to overwrite saved images, the error images will be labeled as "image error delete", and the image will be excluded from the analysis.

Other specific cases will be handled as described below.

a. When a patient was prescribed Letrozole adjuvant therapy prior to surgery, the patient will be removed from the device performance evaluation (major protocol deviation) in the per-protocol analysis but will be included in the mITT population for primary analysis.
b. When the time from injection of LUM015 to initial imaging time is greater than 6 hours due to unplanned surgical delays, the patients will be removed from the device performance evaluation (major protocol deviation) in the per-protocol analysis but will be included in the mITT population for primary analysis.
c. For the re-excision rate endpoints, all patients undergoing a re-excision will be included in the analysis.
d. In the case of a therapeutic shave taken but no image saved before the shave was removed, the shave will be excluded from analysis but will be included in the estimation of final margin.
e. In the case of a SOC shave taken but no image saved before the shave was removed, the shave will be excluded from analysis but will be included in the estimation of SOC margin.
f. In the case of a protocol deviation where an SOC shave is taken after use of the LUM Imaging System, the analysis will be as follows:

i. For per tissue sensitivity and specificity: if there is a Lumicell image for this SOC shave, then the SOC shave will be used as truth standard for that orientation. If there is no Lumicell image, then the SOC shave will not be used for sensitivity and specificity evaluation.
ii. For margin assessment analysis: for the calculation of the SOC margin status versus the margin status after Lumicell, the margin of that SOC shave will be excluded.

g. In the case of a patient having bilateral lumpectomy, only the data collected from the side of the breast that was imaged will be included in the device performance evaluation.
h. In the case of an unknown dimension of a shave, the volume of the shave will be assigned as the average volume of that of all the existing shaves of the corresponding type (SOC or therapeutic shaves).

4.12 Analysis by additional diagnostic subgroups of interest

Primary and secondary endpoints and associated two-sided 95% confidence intervals will be presented by grouping the subjects to those having the comprehensive shaves, having selective shave or no SOC shaves. The purpose of these analyses is to assess consistency within subgroups. Since the study was not powered to detect differences in subgroups, this aim will be considered exploratory in nature and no formal statistical testing will be performed.

4.13 Exploratory Data Analyses

Patient reported outcome measures (PROMs) will be reported as an exploratory endpoint. Also, the type of tissue found in therapeutic shaves will be reported as an exploratory endpoint.

In this protocol, the definition of positive margins is described in section 3.2.2. In addition, Lumicell will collect margin distance information such that alternative definitions (e.g. DICS on ink) can be tested in an exploratory manner.

As an additional exploratory analysis, the sensitivity and specificity in a per-tissue basis for the outermost surface of the SOC procedure will be calculated with their respective 95% confidence interval. For this analysis, only tissue that has a subsequent therapeutic shave or second surgery will be included in the analysis.

5 References

1. Moran, M.S., et al., *Society of Surgical Oncology-American Society for Radiation Oncology consensus guideline on margins for breast-conserving surgery with whole-breast irradiation in stages I and II invasive breast cancer.* Int J Radiat Oncol Biol Phys, 2014. **88**(3): p. 553-64.
2. Morrow, M., et al., *Society of Surgical Oncology-American Society for Radiation Oncology-American Society of Clinical Oncology Consensus Guideline on Margins for Breast-Conserving Surgery With Whole-Breast Irradiation in Ductal Carcinoma In Situ.* J Clin Oncol, 2016.
3. Bartelink, H., et al., *Impact of a higher radiation dose on local control and survival in breast-conserving therapy of early breast cancer: 10-year results of the randomized boost versus no boost EORTC 22881-10882 trial.* J Clin Oncol, 2007. **25**(22): p. 3259-65.

4. Levey, A.S., et al., *A new equation to estimate glomerular filtration rate.* Ann Intern Med, 2009. **150**(9): p. 604-12.