Cover Page for Statistical Analysis Plan

Sponsor name	Sesen Bio Inc
NCT Number	NCT02449239
Sponsor trial ID:	VB4-845-02-IIIA
Official title of study	Open-Label, Multicenter, Ph 3 [Phase 3] Study to Evaluate the Efficacy and Tolerability of Intravesical Vicinium™ in Subjects With Non Muscle-Invasive Carcinoma in Situ and/or High-Grade Papillary Disease of the Bladder Treated With BCG
Document Date	27 October 2020

STATISTICAL ANALYSIS PLAN

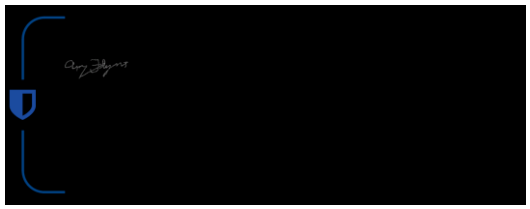
Study Title:	An Open-Label, Multicenter, Phase 3 Study to Evaluate the Efficacy and Tolerability of Intravesical Vicinium™ in Subjects with Non Muscle-Invasive Carcinoma in Situ (CIS) and/or High-Grade Papillary Disease of the Bladder Previously Treated with Bacillus Calmette-Guérin (BCG)
Phase:	3
Protocol No.:	VB4-845-02-IIIA
Protocol Date	Original Protocol: 3 April 2015 Amendment 1: 8 September 2015 Amendment 2: 22 June 2016 Amendment 3: 26 January 2017
Analysis Plan Version and Date	Version 1, Final 18 September 2019 Version 2, Final 27 October 2020
Prepared By:	[redacted]
Prepared For:	Sesen Bio, Inc. 245 First Street Suite 1800 Cambridge, MA 02142

CONFIDENTIAL AND PROPRIETARY INFORMATION

STATISTICAL ANALYSIS PLAN REVIEW AND APPROVAL

This Statistical Analysis Plan has been prepared in accordance with team reviewers' specifications.

Prepared by:



28-Oct-2020 | 16:34 EDT

[redacted]

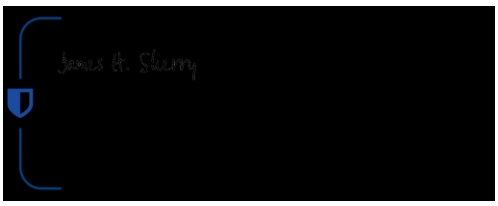
Date

Review:

29-Oct-2020 | 06:57 EDT

[redacted]

Date



18-Nov-2020 | 02:41 P T

[redacted]

Date

Sesen Bio, Inc.

TABLE OF CONTENTS

1 INTRODUCTION

This document describes the statistical methods and data presentations to be used in the summary and planned analysis of data from Protocol VB4-845-02-IIIA. Background information is provided for the overall study design and objectives. The reader is referred to the study protocol and case report forms (CRFs) for details of study conduct and data collection. Analyses of pharmacokinetic, EpCAM, and anti-Vicinium antibody data will be described in a separate Analysis Plan.

This analysis plan will serve as the final documentation for the planned statistical analyses for study VB4-845-02-IIIA and will supersede the protocol if there are any discrepancies.

1.1. STUDY OVERVIEW

This is an open-label, non-randomized, multicenter, multiple-dose, study of Vicinium in approximately 134 subjects with non-muscle invasive bladder cancer (NMIBC). Subjects will have histologically-confirmed non muscle-invasive carcinoma in situ (CIS) and/or papillary disease (high grade non-invasive papillary carcinoma (Ta) or any grade in which the tumor invades lamina propria (T1)) of the bladder who failed previous treatment (i.e., not those who are intolerant) with bacillus Calmette-Guérin (BCG).

All subjects must have received "adequate BCG treatment", defined as at least 2 courses of full dose BCG i.e., at least one induction course of at least 5 doses given within 7 weeks, and a second course (maintenance or second induction) of at least 2 doses given within 6 weeks. A full dose BCG is one of the following

- TICE® BCG: one vial of TICE containing 1 to 8 x 10^8 colony-forming units (CFU) equivalent to approximately 50 mg (wet weight) or equivalent.
- ImmuCyst® BCG: one vial of ImmuCyst containing 81 mg (dry weight) equivalent to approximately 10.5 ± 8.7 x 10^8 CFU or equivalent

The "5+2" doses must be completed within 1 year and given to treat the same disease episode with which the subject is enrolling. All subjects must have disease for which the investigator would not treat with additional BCG at this time.

Three cohorts of subjects will enroll:

- Cohort 1: Subjects with CIS with or without associated papillary disease whose disease is determined to be refractory or relapsed within 6 months of the last dose of adequate BCG treatment.
- Cohort 2: Subjects with CIS with or without associated papillary disease whose disease is determined to be refractory or relapsed more than 6 months but within 11 months of the last dose of adequate BCG treatment.
- Cohort 3: Subjects with high-grade Ta or any grade T1 papillary disease (without CIS) whose disease is determined to be refractory or relapsed within 6 months of the last dose of adequate BCG treatment.

For eligibility and cohort assignment 6 months is defined as 28 weeks i.e., 26 weeks (6 months) plus an additional 2 weeks for diagnostic work-up and 11 months is defined as 50 weeks i.e., 48 weeks (11 months) plus an additional 2 weeks for diagnostic work-up.

Of the 134 subjects, ≥77 will accrue to Cohort 1. Up to 57 additional subjects in total may accrue to Cohorts 2 and 3. This is an outpatient study. All treatments may be administered in the study clinic.

All diagnostic pathology (bladder tissue and urine cytologies and if performed, tissue from prostatic urethral biopsies) will be reviewed and graded by an independent pathologist and cytopathologist with experience in NMIBC diagnoses. The details of sample handling and logistics will be provided in a separate laboratory manual.

The sample used for determining subject eligibility for inclusion into the trial (post-BCG biopsy) must be reviewed and signed off by the pathologist prior to study enrollment and start of study treatment. All other samples will be read in a timely fashion and the results will be available to study investigators via the electronic CRF.

Each subject's study course consists of a Screening Period, a 12-week Induction Phase, a Maintenance Phase and a Post-Treatment Evaluation.

Following the 12-week Induction Phase, subjects will be evaluated for disease status as described in Protocol Section 14. Subjects without a treatment failure as specifically defined in Protocol Section 14.2 or other criteria necessitating discontinuation of study treatment (see Protocol Section 11.3) will begin the Maintenance Phase. Subjects will receive treatment for up to 2 years in total if they are tolerating treatment and do not meet the definition of protocol-defined treatment failure or other criteria for study discontinuation. Following the Post-Treatment Evaluation, additional subject data will be collected as needed for the evaluation of the Secondary Endpoints.

1.2. VISIT SCHEDULE AND ASSESSMENTS (SCREENING & INDUCTION)

Assessment	Screening and Baseline Weeks		Induction Phase Weeks 1- 12					Post-Induction Phase
Week	-8 through -1	-1	1	2-5	6	7-11	12	As soon as possible after Week 12
Informed consent	X							
Inclusion/Exclusion criteria	X							
Cystoscopy	X[1,2]							X[2]
Bladder biopsy/TURBT	X[1]							X[3]
EpCAM determination, Biomarker evaluations[15]	X[15]							X[15]
Urine Cytology			X[8]				X[10]	
Demographics	X							
Medical and medication history (including previous anti-tumor therapy and surgical procedures)	X	X[6]						
Physical examination		X	X[8]	X[11]		X[11]	X[11]	
Weight and height		X[12]		X[12]		X[12]	X[12]	
Radiological imaging	X[4]							
Triplicate ECG	X[5]						X[11]	
Vital signs	X		X[9]	X[9]	X[9]	X[9]	X[9]	
Hematology		X	X[8]	X[11]	X[11]	X[11]	X[11]	
Chemistry		X	X[8]	X[11]	X[11]	X[11]	X[11]	
Urinalysis		X	X[8]		X[11]		X[11]	
Anti-Vicinium antibody test		X					X	
Plasma Vicinium Levels (PK)[14]			X		X		X	
Serum or urine pregnancy test		X[7]		X[7]		X[7]	X	
Concomitant medication			X[10]	X[10]	X[10]	X[10]	X[10]	X
Adverse events (AEs / SAEs)			X[9]	X[9]	X[9]	X[9]	X[9]	X
Drug administration (BIW)[13]			X	X	X			
Drug administration (once/week)						X	X	

[1] Must have documented non-muscle invasive urothelial bladder cancer that is histologically confirmed by the independent central pathology reviewer prior to the first dose of study drug. All papillary disease must be completely resected. In subjects with T1 disease, a second (i.e., restaging) TURBT must be performed approximately 2 - 6 weeks after the first. A biopsy may be more than 8 weeks prior to the first dose of study drug as long as the last biopsy that confirms eligibility is within 8 weeks.

[2] White-light cystoscopy. Any subject who undergoes fluorescence cystoscopy for bladder biopsy/TURBT during the screening period must continue having this modality for all subsequent biopsies (however, white-light cystoscopy will suffice for all assessments if a biopsy is not required (Protocol Section 14).

[3] Directed biopsies of all areas suspicious for disease. Random biopsies are required if cystoscopy is negative and urine cytology is positive or there are two suspicious cytology results (Protocol Section 14)

[4] CT urogram or MRI urogram. If intravenous contrast is contraindicated, retrograde ureteropyelography, or CT or MRI without intravenous contrast may be performed. Required during Screening for all subjects with T1 disease. Required for all other subjects if not previously performed within 13 weeks before the start of dosing.

[5] Single tracing only at Screening unless QTc is prolonged (Protocol Section 16.4); thereafter, triplicate ECGs are to be performed.

[6] Update medication history.

[7] Performed every 4 weeks (i.e., week 4, week 8, week 12, etc.) while on treatment in females of childbearing potential only.

[8] Cytology sample pre-dose on Day 1. Note: Chemistry, hematology (CBC and differential), urinalysis and physical examination need to be performed pre-dose Day 1 only if not performed in the Baseline Period (≤7 days before the first dose of study medication).

[9] Evaluation performed pre-dose and post-dose prior to leaving the clinic.
[10] Evaluation performed pre-dose. Concomitant medications required post-dose only if additional medications (not including Vicinium) were administered during or after the start of the instillation.
[11] Weeks 2-5: Chemistry and hematology (CBC and differential) performed at any time once weekly (Protocol Section 10.3.2). Week 4: Targeted physical examination. Week 6: Urinalysis performed pre-dose at the second dose of the week; Hematology (CBC and differential) and chemistry (Protocol Section 10.3.3). Weeks 7-11: Chemistry and hematology (CBC and differential) performed at any time once weekly (Protocol Section 10.3.4). Week 8: Targeted physical examination. Week 12: Urinalysis performed pre-dose. Hematology (CBC and differential), chemistry (Protocol Section 10.3.5). Targeted physical examination and triplicate ECG performed post-dose prior to leaving the clinic.
[12] Height and weight at Baseline. Weight performed Week 4, Week 8 and Week 12.
[13] Twice weekly (BIW) dosing with each dose separated by at least 48 hours (and no more than 2 doses within any calendar week (Sunday through Saturday)). As an example, while a Monday/Thursday or Tuesday/Friday would be optimal, a Monday/Wednesday or Tuesday/Thursday, or Wednesday/Friday dosing schedule could be used.
[14] Levels are to be taken pre-dose of first instillation. Additional sampling is at the end of the dwell time as soon as possible (within 30 minutes) after the subject voids (or bladder is drained via catheter) following the first instillation and after the second of the twice-weekly instillations during Weeks 6 and 12.
[15] For all bladder biopsy/TURBT samples positive for urothelial carcinoma (Protocol Section 16.7).

1.3. VISIT SCHEDULE AND ASSESSMENTS (MAINTENANCE PHASE)

Assessment	Maintenance Phase					Post-Treatment (30 – 37 days after last dose)
	Treatment Days	Every 2 Weeks	Every 4 Weeks	Every 13 Weeks	Between Weeks 23 and 27	
Cytology				X[1,10]		X[1,6]
Cystoscopy				X[1,2,10]		X[2,3,6]
Biopsy/TURBT				X[1,3,10,13]		X[2,3,6,13]
Physical examination			X[7]			X[7]
Triplicate ECG						X
Vital signs	X[4]					X
Weight			X			X
Anti-Vicinium antibody test					X	X
Chemistry		X[12]				X
Hematology and urinalysis			X[9]			X
Urine or serum pregnancy test			X[5]			X
Concomitant medication	X[4]					X[8]
Adverse events (AEs / SAEs)	X[4]					X
Drug administration (once every other week)	X					
Contact to assess for progression and survival				X[11]		

[1] Window of 1 week before to 2 weeks after to start evaluation

[2] White-light cystoscopy. Any subject who underwent fluorescence cystoscopy for bladder biopsy/TURBT during the screening period must continue having this modality for all subsequent biopsies (however, white-light cystoscopy will suffice for all assessments if a biopsy is not required (Protocol Section 14).

[3] Directed biopsies of all areas suspicious for disease. Random biopsies are required if cystoscopy is negative and urine cytology is positive or there are two suspicious cytology results (Protocol Section 14)

[4] Evaluation performed pre-dose and post-dose prior to leaving the clinic. Concomitant medications required post-dose only if additional medications (not including Vicinium) were administered during or after the start of the instillation

[5] Performed every 4 weeks during study participation in females of childbearing potential only.

[6] Evaluations performed if the subject has not discontinued treatment due to documented disease recurrence, progression and if they have not been performed within the past 8 weeks.

[7] Targeted physical examination.

[8] Additional follow-up will occur for subjects with ongoing clinically significant AEs (e.g., those leading to study drug discontinuation) or SAEs assessed as related to study drug.

[9] Urinalysis performed pre-dose if on a treatment day.

[10] Weeks 26, 39, 52, 65, 78, 91, 104

[11] Starting after subject discontinues study drug. Progression will be assessed if the subject did not have disease progression as the reason for discontinuation of study drug and has not had a cystectomy.

[12] Protocol Section10.5.2.

[13] EpCAM and biomarker evaluation on all bladder biopsy/TURBT samples positive for urothelial carcinoma (Protocol Section 16.7).

1.4. LIST OF ABBREVIATIONS

Abbreviation	
AE	adverse event
BCG	bacillus Calmette-Guérin
BIW	twice weekly
CFU	colony forming units
CIS	carcinoma in situ
CRF	case report form
CT	computerized tomography
CTCAE	Common Terminology Criteria for Adverse Events
DSMB	data safety monitoring board
ECG	electrocardiogram
FDA	Food and Drug Administration
EFS	event-free survival
EpCAM	epithelial cell adhesion molecule
MedDRA	Medical Dictionary for Regulatory Activities
mITT	modified intent-to-treat
MRI	magnetic resonance imaging
NCI	National Cancer Institute
NMIBC	non-muscle invasive bladder cancer
OS	overall survival
PFS	progression-free survival
PK	pharmacokinetics
QTc	corrected QT interval
QTcB	corrected QT interval - Bazett's correction
QTcF	corrected QT interval - Fridericia correction
SAE	serious adverse event
SD	standard deviation
SOC	system organ class
Ta	non-invasive papillary carcinoma
T1	tumor invades lamina propria
T2	tumor invades muscularis propria
TEAE	treatment-emergent adverse event
TTP	time to progression
TURBT	transurethral resection of bladder tumor(s)

2. OBJECTIVES

To assess the efficacy and tolerability of Vicinium when administered as a monotherapy intravesical instillation in subjects with non-muscle invasive bladder cancer - CIS, high-grade Ta or any grade T1 papillary disease or CIS plus papillary disease who failed previous treatment (i.e., not those who are intolerant) with BCG.

Primary endpoints:
Complete response rate at 3 months in subjects with CIS (with or without papillary disease) whose disease is refractory or relapsed in 6 months or less or relapsed from 6 months to 11 months following adequate BCG treatment.

Duration of response (DoR) will be estimated (Kaplan-Meier Estimate) for those subjects with CIS (with or without papillary disease) who experience a complete response.

Secondary endpoints:
- Complete response rate in subjects at 6, 9, 12, 15, 18, 21, and 24 months in subjects with CIS
- Event free survival (EFS) in all subjects
- Cystectomy-Free Days in all subjects
- Time to disease recurrence in papillary subjects
- Progression-free survival (PFS) in all subjects
- Overall survival (OS) in all subjects
- Safety and tolerability of Vicinium therapy in all subjects

Exploratory:
To evaluate biomarkers that may be associated with response or disease progression or treatment failure, which may include, for example, EpCAM status, tumor subtype morphology, furin levels in tumor cell endosomes, presence of a glycosaminoglycan coat, and presence of receptors that could impede a host anti-tumor immune response such as PD-L1.

3. GENERAL STATISTICAL CONSIDERATIONS

3.1. SAMPLE SIZE AND POWER

Assuming the true complete response rate at 12 months is 30%, sample sizes associated with various widths of a two-sided, exact (Clopper-Pearson) 95% confidence intervals are provided in the table below:

Sample Size (N)	Target Distance from P to Lower Limit	Proportion (P)	Lower Limit
121	0.080	0.300	0.220
95	0.090	0.300	0.210
76	0.100	0.300	0.200

Under the provided assumptions, a sample size of ≥77 evaluable subjects with early relapsed/ refractory CIS with or without papillary disease is sufficient to estimate complete response rate with an exact 95% confidence interval that excludes 0.2 (PASS 2008).

The secondary endpoint of event-free survival (EFS) will be used to assess the treatment effect in all subjects, i.e., those with CIS (with or without papillary disease) and those with papillary-only disease. An estimate of EFS at the 18 month time-point will be calculated. The total sample size of 134 subjects will provide at least 80% power to test the null hypothesis that the event-free survival rate at 18 months is 20% versus the alternative hypothesis that the event-free survival rate at 18 months is ≥30%. This calculation is based on the assumptions of a nonparametric distribution for event-free survival and the length of follow-up is 24 months.

3.2. RANDOMIZATION AND BLINDING

This is an open label, non-randomized study; therefore, blinding considerations are not applicable.

3.3. HANDLING OF DATA

3.3.1. Strata and Covariates

No adjustments will be made for strata.

3.3.2. Examination of Subject Subsets

Analyses of the mITT population will be summarized by cohort at study entry and an overall CIS group.

To assess consistency of efficacy endpoints, the primary and secondary efficacy analyses will be performed for the following subgroups:

- Age (<75, ≥75)
- Gender (Male, Female)
- Region (US, Canada)
- Race (White, Black/African American, Asian, Other)
- Number of prior BCG therapies (≤2, >2)
- Number of prior BCG instillations (≤9, ≥10)

In each subgroup, CR rate and the associated 95% CI will be calculated, and median values and associated 95% CIs will be calculated for all time-to-event endpoints for the appropriate cohort (EFS, cystectomy-free days, time to disease recurrence, time to progression, PFS, and OS). Results will be displayed in forest plots and Kaplan-Meier curves when indicated.

Adverse events will also be analyzed by subgroup (adverse events, serious adverse events, treatment-related adverse events treatment-related, serious adverse events, and Grade 3 or higher adverse events).

3.3.3. Multiple Testing and Comparisons

All analyses will be conducted without adjustments for multiple comparisons.

3.3.4. Missing Data and Outliers

Every effort will be made to obtain required data at each scheduled evaluation from all subjects who have been enrolled.

See section 3.3.5 for details of imputing missing dates. Unless otherwise specified, all other missing data will not be imputed.

3.3.5. Imputation of Incomplete Dates

An incomplete date occurs when the exact date an event occurred or ended cannot be obtained from a subject. The database contains data fields for month, day, and year. A date is incomplete if at least one of these three fields is not known. For many of the planned analyses, a complete date is necessary in order to determine if the event should be included in the analysis (i.e., if the event is treatment-emergent) or to establish the duration of an event. In such cases, incomplete dates will be imputed.

For purposes of imputation of dates for adverse events and concomitant medications, the start and stop dates of study drug will be utilized to impute missing event start and stop dates. All events with an incomplete start/onset date are assumed to have started at the initiation of the study drug. All events with an incomplete end/resolution date are assumed to have ended on the last date of study drug. The missing or incomplete dates will be imputed using the algorithm described in Table 1.

In an effort to minimize bias, the project statistician will impute missing or incomplete dates from forms other than the adverse events or concomitant medications in a systematic, reasonable manner. In these cases, missing date portions will be imputed using available data as described in Table 2.

Table 1. Algorithm for Imputing Missing/Incomplete Onset or Resolution Dates for Adverse Events or Concomitant Medications

Description of Missing/Incomplete Data Piece	Algorithm	Example
Completely Missing Start Date	Use Study Drug Start Date	Original Event Start Date = " " Study Drug Start Date = "04-16-2014" Imputed Event Start Date = "04-16-2014"
Incomplete Start Date - Missing Day Part	Use Day Part of Study Drug Start Date	Original Event Start Date="04- -2014" Study Drug Start Date="04-16-2014" Imputed Event Start Date="04-16-2014"
Incomplete Start Date - Missing Day/Month Part	Use Day/Month Part of Study Drug Start Date	Original Event Start Date="2014" Study Drug Start Date="04-16-2014" Imputed Event Start Date="04-16-2014"
Completely Missing Stop Date	Use Study Drug Stop Date	Original Event Stop Date = " " Study Drug Stop Date = "04-16-2014" Imputed Event Stop Date = "04-16-2014"
Incomplete Stop Date - Missing Day Part	Use Day Part of Study Drug Stop Date	Original Event Stop Date="04- -2014" Study Drug Stop Date="04-16-2014" Imputed Event Stop Date="04-16-2014"
Incomplete Stop Date - Missing Day/Month Part	Use Day Part of Study Drug Stop Date	Original Event Stop Date="2014" Study Drug Stop Date="04-16-2014" Imputed Event Stop Date="04-16-2014"

Table 2. Algorithm for Imputing Missing/Incomplete Dates Not on the Adverse Events or Concomitant Medications CRF Pages

Description of Missing/Incomplete Data Piece	Algorithm	Example
Completely Missing Date	Use CRF Visit Date, if available	Original Date = " " CRF Visit Date = "04-16-2014" Imputed Date = "04-16-2014"
Incomplete Date - Missing Day Part	Use Day Part of CRF Visit, if available	Original Date="04- -2014" CRF Visit Date="04-16-2014" Imputed Date="04-16-2014"
Incomplete Date - Missing Day/Month Part	Use Day/Month Part of CRF Visit, if available	Original Date="2014" CRF Visit Date="04-16-2014" Imputed Date="04-16-2014"

If the CRF Visit Date is not available, dates of related assessments may be used to determine the date of the assessment. For instance, missing dates for the centrally read pathology or centrally read cytology results where there is not a corresponding visit date for the given visit, the date of the TURBT or date of the cystoscopy associated with the biopsy for the visit will be substituted.

A list of incomplete and imputed dates will be prepared by the project statistician or statistical programmer(s) and will be submitted for review by the sponsor.

3.3.6. Derived and Transformed Data

Transformations for variables with skewed distributions may be performed as appropriate.

3.3.7. Definitions and Terminology

Adverse Event
An adverse event (AE) is defined as any unfavorable and unintended sign (including an abnormal laboratory finding), symptom or disease temporally associated with the use of a medical treatment or procedure regardless of whether it is considered related to the medical treatment or procedure (attribution of unrelated, unlikely, possible, probable, or definite).

Any medical condition that is present at the time the subject is screened, but does not deteriorate, should not be reported as an AE. However, if it deteriorates at any time during the study it should be recorded as an AE. Hospitalization for elective treatment of a preexisting condition that did not worsen from baseline is not considered an AE or serious AE (SAE).

The descriptions and grading scales found in the revised National Cancer Institute Common Terminology Criteria for Adverse Events (NCI-CTCAE) version 4.03 will be utilized for AE reporting.

Serious Adverse Event
A serious adverse event is any AE that is fatal, life-threatening, requires or prolongs hospital stay, results in persistent or significant disability or incapacity, results in a congenital anomaly or birth defect, or is an otherwise important medical event.

Treatment-emergent AE
Treatment-emergent AEs (TEAEs) are all AEs occurring on or after the date of first dose of study drug through 30 days post the last dose of study drug.

Baseline Value
For purposes of analysis, the baseline value is defined as the last value obtained prior to the first dose of Vicinium.

Study Drug
For the purposes of this study, the term study drug refers to Vicinium.

Day 1 (Baseline)
Day 1 is the day that the first dose of Vicinium is received.

Age
The age of a subject is defined as (Date of Informed Consent – Date of Birth + 1)/365.25.

Study Day
Study Day is defined relative to Baseline (Day 1). Thus, the study day of an event is calculated as:

Study Day = event date – date of Day 1 + 1.

Change from Baseline
Change from baseline for a given endpoint is defined as the Study Day X value minus the Baseline Value.

Cohort 1 Subjects
Subjects with CIS with or without associated papillary disease whose disease is determined to be refractory or relapsed within 6 months of the last dose of adequate BCG treatment. For cohort assignment, 6 months is defined as 28 weeks i.e., 26 weeks (6 months) plus an additional 2 weeks for diagnostic work-up. This group of subjects is identified as Cohort 1 within the eCRF.

Cohort 2 Subjects
Subjects with CIS with or without associated papillary disease whose disease is determined to be refractory or relapsed within 6 months to 11 months of the last dose of adequate BCG treatment. For cohort assignment, 11 months is defined as 50 weeks i.e., 48 weeks (11 months) plus an additional 2 weeks for diagnostic work-up. This group of subjects is identified as Cohort 2 within the eCRF.

Cohort 3 Subjects
Subjects with papillary disease whose disease is determined to be refractory or relapsed within 6 months of the last dose of adequate BCG treatment. For cohort assignment, 6 months is defined as 28 weeks i.e., 26 weeks (6 months) plus an additional 2 weeks for diagnostic work-up. This group of subjects is identified as Cohort 3 within the eCRF.

Adequate BCG Treatment
Adequate BCG Treatment is defined as at least 2 courses of full dose BCG i.e., at least one induction course of at least 5 doses given within 7 weeks, and a second course (maintenance or second induction) of at least 2 doses given within 6 weeks. Full-dose BCG is:

- TICE® BCG: one vial of TICE containing 1 to 8 x 10^8 colony-forming units (CFU) equivalent to approximately 50 mg (wet weight)
- ImmuCyst® BCG: one vial of ImmuCyst containing 81 mg (dry weight) equivalent to approximately 10.5 ± 8.7 x 10^8 CFU or equivalent.

The "5+2" doses must be completed within 1 year and given to treat the same disease episode with which the subject is enrolling. All subjects must have disease for which the investigator would not treat with additional BCG treatment at the time of enrollment. The determination of adequate BCG will be based on the data in the eCRF.

Treatment-emergent Laboratory Toxicity
A treatment-emergent laboratory toxicity is defined as an increase of at least one toxicity grade from the baseline assessment at any post baseline assessment. For defining treatment-emergent laboratory events, it is assumed that a laboratory assessment obtained on Day 1 with missing collection times occurred prior to the initiation of study drug. If the relevant baseline assessments are missing, then any graded abnormality (i.e., at least Grade 1) is considered as treatment-emergent. For the grading of severity of the laboratory toxicity, the CTCAE v 4.03 will be used.

Concomitant Medications
Concomitant medications are those medications taken on or after the initiation of study drug.

Previous Medications
Previous medications are those medications taken and stopped prior to initiation of study drug.

Treatment Failure

Treatment failure is defined as an event or events that demonstrate the study drug, Vicinium, is not able to achieve or maintain a response in the subject's bladder cancer. Specifically, these events include:

- Subjects whose diagnosis at study entry was CIS or high-grade papillary disease will be considered to have a treatment failure if at the time of any disease status evaluation (per protocol every 13 weeks or any unscheduled evaluation) there is centrally read evidence of high-grade disease (CIS, high-grade Ta or any grade T1 disease) or disease progression (e.g., to muscle invasive disease).
- Low-grade Ta disease is not considered a treatment failure in these subjects and they may remain on study treatment following TURBT. Low-grade Ta disease at the time of evaluation must have the low-grade Ta disease resected.
- Subjects with positive centrally-assessed urine cytology without histologic, cytologic or locally read radiographic evidence of upper or lower urinary tract malignancy will be considered as no longer in complete response and hence, as having a treatment failure.
- Subjects with histologic, cytologic or radiographic evidence of upper or lower urinary tract malignancy will not be considered to have treatment failure but will have study therapy discontinued.
- Subjects who are discontinued due to treatment failure or for recurrence will be considered as treatment failure

Complete Response

Subjects with CIS (Cohort 1 or Cohort 2) will be considered to have a complete response if at the time of any disease status evaluation (per protocol every 13 weeks or any unscheduled evaluation) there is no evidence of high-grade disease (CIS, high-grade Ta or any grade T1 disease) or disease progression (e.g., to muscle invasive disease). Subjects whose initial diagnosis was CIS who are found to have low-grade Ta disease at the time of an evaluation must have the low-grade disease resected by TURBT. Low-grade Ta disease is not considered a treatment failure in these subjects and they may remain on study treatment following TURBT.

Date of Complete Response or Treatment Failure

Given that the cystoscopy, biopsy, and cytology assessments may be conducted on different dates, the date of complete response or treatment failure will be defined as the earliest assessment date of these assessments.

Duration of Response

Duration of Response will be analyzed for subjects with CIS with or without papillary disease (Cohort 1 + Cohort 2) who achieve complete response using a Kaplan-Meier (KM) product-limit method. It is defined as the number of days from the date of first occurrence of complete response to the date of documented treatment failure or death, whichever occurs first. Subjects who do not experience treatment failure or death will be censored at their last assessment of response.

Event

For the purposes of defining Event-Free Survival, an event is defined as treatment failure or death on or prior to treatment discontinuation. Because disease assessments are not recorded after treatment discontinuation, a death after treatment discontinuation will not be counted as an event so as not to artificially inflate EFS. Presence of upper or lower urinary tract urothelial carcinoma and having a cystectomy are not treated as events for analysis but do trigger special data handling in the event-free survival analysis.

Event-Free Survival

EFS is defined as the interval from the date of first dose of study treatment to an event. Subjects who have a treatment failure or die following the post induction visit will be treated as having an event on the date of treatment failure or death. Subjects without evidence of a treatment failure who are diagnosed with upper or lower urinary tract urothelial carcinoma by histology, cytology or radiographic evidence will be censored at the date of diagnosis. Subjects without evidence of a treatment failure who undergo cystectomy will be censored at the date of cystectomy. All other subjects will be censored at the time of last assessment including long-term follow up.

Censoring rules for the analysis of EFS are presented in the table below.

Table: Censoring scheme for EFS

Situation	Date of Event or Censoring	Outcome
Treatment failure	Date of treatment failure	Event
Death without treatment failure and occurring prior to treatment discontinuation	Date of death	Event
Death without treatment failure and occurring after treatment discontinuation	Date of last evaluable disease assessment on or prior to treatment discontinuation[a]	Censored[a]
No baseline disease assessment	Date of first dose	Censored
No on-study disease assessments and no death	Date of first dose	Censored
No treatment failure and no death	Date of last evaluable disease assessment	Censored
New anti-cancer therapy received without treatment failure and reported prior to or on the same day as disease assessment	Date of last evaluable disease assessment on or prior to the date of initiation of subsequent therapy	Censored
Second non-NMIBC cancer reported prior to or on the same day as disease assessment	Date of last evaluable disease assessment on or prior to the date of diagnosis of non-NMIBC cancer	Censored
a This censoring rule acknowledges that disease assessments were not recorded after treatment discontinuation due to an AE.		

Cystectomy-Free Days

Cystectomy-free days is defined as the time from the date of first dose of study treatment to physical removal of the bladder. Subjects who are alive and do not have their bladder removed will be censored at the time of last contact.

Censoring rules for the analysis of cystectomy-free days are presented in the table below.

Table: Censoring scheme for Cystectomy

Situation	Date of Event or Censoring	Outcome
Cystectomy	Date of cystectomy	Event
Death without cystectomy	Date of death	Censored
No cystectomy	Date of last contact	Censored

Time to Disease Recurrence

Time to disease recurrence is defined for subjects in Cohort 3 only as follows:

Time to disease recurrence is defined as the number of days from first dose of study treatment to the first occurrence of treatment failure or death on or prior to treatment discontinuation. Because disease assessments are not recorded after treatment discontinuation, a death after treatment discontinuation will not be counted as an event so as not to artificially inflate time to disease recurrence. Subjects who do not experience treatment failure will be censored at the last assessment on treatment.

Censoring rules for the analysis of time to disease recurrence are presented in the table below.

Table: Censoring scheme for time to disease recurrence in Cohort 3

Situation	Date of Event or Censoring	Outcome
Disease recurrence	Date of disease recurrence	Event
Death without disease recurrence and occurring prior to treatment discontinuation	Date of death	Event
Death without disease recurrence and occurring after treatment discontinuation	Date of last evaluable disease assessment on or prior to treatment discontinuation[a]	Censored[a]
No baseline disease assessment	Date of first dose	Censored
No on-study disease assessments and no death	Date of first dose	Censored
No disease recurrence and no death	Date of last evaluable disease assessment	Censored
New anti-cancer therapy received without treatment failure and reported prior to or on the same day as disease assessment	Date of last evaluable disease assessment on or prior to the date of initiation of subsequent therapy	Censored
Second non-NMIBC cancer reported prior to or on the same day as disease assessment	Date of last evaluable disease assessment on or prior to the date of diagnosis of non-NMIBC cancer	Censored
a This censoring rule acknowledges that disease assessments were not recorded after treatment discontinuation due to an AE.		

Progression-free Survival (PFS)

Progression-free survival (PFS) will be analyzed for subjects in Cohorts 1, 2 and 3 and is defined as the time from the date of first dose of study treatment to the date of disease progression (e.g., T2 or more advanced disease) or death on or prior to treatment discontinuation, whichever occurs first. Because disease assessments are not recorded after treatment discontinuation, a death after treatment discontinuation will not be counted as an event so as not to artificially inflate PFS. Subjects without disease progression or death will be censored at the last assessment.

Censoring rules for the analysis of PFS are presented in the table below.

Table: Censoring scheme for PFS in Cohorts 1, 2, and 3

Situation	Date of Event or Censoring	Outcome
Disease progression	Date of disease progression	Event
Death without disease progression and occurring prior to treatment discontinuation	Date of death	Event
Death without disease progression and occurring after treatment discontinuation	Date of last evaluable disease assessment on or prior to treatment discontinuation[a]	Censored[a]
No baseline disease assessment	Date of first dose	Censored
No on-study disease assessments and no death	Date of first dose	Censored
No disease progression and no death	Date of last evaluable disease assessment	Censored
New anti-cancer therapy received without treatment failure and reported prior to or on the same day as disease assessment	Date of last evaluable disease assessment on or prior to the date of initiation of subsequent therapy	Censored
Second non-NMIBC cancer reported prior to or on the same day as disease assessment	Date of last evaluable disease assessment on or prior to the date of diagnosis of non-NMIBC cancer	Censored
a This censoring rule acknowledges that disease assessments were not recorded after treatment discontinuation due to an AE.		

Overall Survival (OS)

Overall Survival (OS) is defined as the time from the date of first dose of study treatment to death due to any cause. Subjects who are alive will be censored at the date of last contact.

Fridericia Corrected QT (QTcF)

$$QTcF = \frac{QT}{\sqrt[3]{RR}}$$

Bazett Corrected QT (QTcB)

$$QTcB = \frac{QT}{\sqrt{RR}}$$

3.3.8. By-Study Visit Displays

When data are collected serially over time, individual data presentations may include by-visit displays. Visits will be presented according to the nominal visit as obtained from the CRF or laboratory. If assessments are collected multiple times within a given visit, the value from the scheduled visit will be chosen over the value from the unscheduled visit. If there are two or more assessments with the same nominal visit label, the value closest to the nominal visit will be summarized. If two assessments have the exact same distance from the nominal visit by date and time, if applicable, then the last visit will be chosen for summarization. Unscheduled, End of Treatment, and Post Treatment Follow-Up Visits will be windowed according to visit date so that these data are appropriately analyzed. The following visit windows will be applied according to the following table:

Analysis Visit	**Study Days**	**Target Study Day for Visit**
Baseline	-8< Study Day <=1	
Week 1	1 < Study Day <= 7	1
Week 2	7 < Study Day <= 14	8
Week 3	14 < Study Day <= 21	15
Week 4	21 < Study Day <= 28	22
Week 5	28 < Study Day <= 35	29
Week 6	35 < Study Day <= 42	36
Week 7	42 < Study Day <= 49	43
Week 8	49 < Study Day <= 56	50
Week 9	56 < Study Day <= 63	57
Week 10	63 < Study Day <= 70	64
Week 11	70 < Study Day <= 77	71
Week 12	77 < Study Day <= 84	78
Week 13	84 < Study Day <= 91	85

Analysis Visit	Study Days	Target Study Day for Visit
Maintenance Month 1	91 < Study Day <= 121	115
Maintenance Month 2	121 < Study Day <= 152	146
Maintenance Month 3	152 < Study Day <= 182	176
Maintenance Month 4	182 < Study Day <= 212	206
Maintenance Month 5	212 < Study Day <= 243	237
Maintenance Month 6	243 < Study Day <= 273	267
Maintenance Month 7	273 < Study Day <= 303	297
Maintenance Month 8	303 < Study Day <= 334	328
Maintenance Month 9	334 < Study Day <= 364	358
Maintenance Month 10	364 < Study Day <= 394	388
Maintenance Month 11	394 < Study Day <= 425	419
Maintenance Month 12	425 < Study Day <= 455	449
Maintenance Month 13	455 < Study Day <= 485	479
Maintenance Month 14	485 < Study Day <= 516	510
Maintenance Month 15	516 < Study Day <= 546	540
Maintenance Month 16	546 < Study Day <= 576	570
Maintenance Month 17	576 < Study Day <= 607	601
Maintenance Month 18	607 < Study Day <= 637	631
Maintenance Month 19	637 < Study Day <= 667	661
Maintenance Month 20	667 < Study Day <= 698	692
Maintenance Month 21	698 < Study Day <= 773	722
Long-term Follow up Month 3	773 < Study Day <= 864	819
Long-term Follow up Month 6	864 < Study Day <= 955	910
Long-term Follow up Month 9	955 < Study Day <= 1046	1001

Analysis Visit	Study Days	Target Study Day for Visit
Long-term Follow up Month 12	1046 < Study Day <= 1137	1092
Long-term Follow up Month 15	1137 < Study Day <= 1228	1183
Long-term Follow up Month 18	1228 < Study Day <= 1319	1274
Long-term Follow up Month 21	1319 < Study Day <= 1410	1365
Long-term Follow up Month 24	1410 < Study Day <= 1501	1456

3.3.9. Combining Data Across Panels

When data are combined across panels to create composite endpoints, like complete response, information will be combined in the following manner:

1. Data will be merged by visit for each subject.
2. Unscheduled and early termination data will be windowed according to the visit windows below. Additionally, data from visits where pathology, cytology, and cystoscopy are unexpected (i.e., Maintenance Month 4) will be windowed using the table below.
3. Data within a given visit window will be collapsed to the worst-case result. For example, if a subject has two central pathology assessments for a given visit one with 'CIS' and the other with 'No Tumor', the analysis result for the collapsed visit will be considered as 'CIS'.
4. The date corresponding to the collapsed visit will be the earliest observed date for the visit.

Analysis Visit	Study Days
Screening	-8< Study Day <1
Week 1	1 <= Study Day <= 49
Post Induction Phase	49 < Study Day <= 137
Maintenance Month 3	137 < Study Day <= 227
Maintenance Month 6	227 < Study Day <= 318
Maintenance Month 9	318 < Study Day <= 409
Maintenance Month 12	409 < Study Day <= 500
Maintenance Month 15	500 < Study Day <= 591
Maintenance Month 18	591 < Study Day <= 682

Maintenance Month 21	682 < Study Day <= 773
Long-term Follow up Month 3	773 < Study Day <= 864
Long-term Follow up Month 6	864 < Study Day <= 955
Long-term Follow up Month 9	955 < Study Day <= 1046
Long-term Follow up Month 12	1046 < Study Day <= 1137
Long-term Follow up Month 15	1137 < Study Day <= 1228
Long-term Follow up Month 18	1228 < Study Day <= 1319
Long-term Follow up Month 21	1319 < Study Day <= 1410
Long-term Follow up Month 24	1410 < Study Day <= 1501

3.4. TIMING OF ANALYSES

A Data Safety Monitoring Board (DSMB) will review aggregated safety and limited efficacy data on a biannual basis.

An analysis of safety and limited efficacy data will be conducted once all subjects have completed 3, 6, and 12 months of follow-up.

A final analysis of all data will be conducted once all subjects enter the off-treatment period and the resulting data have been cleaned and quality assured, and the database locked.

4. ANALYSIS POPULATIONS

All reasonable efforts will be made to obtain complete assessment of all outcome measures for all subjects in the protocol. The following analysis populations will be evaluated:

4.1. ALL ENROLLED POPULATION

All enrolled population includes subjects who signed an informed consent form and were registered into the central enrollment system.

4.2. MODIFIED INTENT TO TREAT POPULATION

Modified Intent-to-Treat population includes subjects who received at least one dose of study medication. This is the primary population for efficacy and safety.

4.3. PRIMARY EFFICACY POPULATION

The primary efficacy population includes any subject in the mITT population with CIS. This will be the primary population utilized for the evaluation of complete response.

4.4. EVALUABLE POPULATION

The evaluable population includes any subject in the mITT population who completes the induction phase. This population will be used for secondary analyses of the efficacy endpoints.

4.5. INDUCTION PHASE DOSING COMPLIANT POPULATION

The induction phase dosing compliant population includes any subject in the evaluable population who completes at least 85% of the dosing the induction phase. This population will be used for secondary analyses of efficacy endpoints.

4.6. STUDY DOSING COMPLIANT POPULATION

The study dosing compliant population includes any subject in the mITT population who completes at least 85% of the dosing for visits completed prior to discontinuation. This population will be used for secondary analyses of efficacy endpoints.

5. STATISTICAL METHODS

Descriptive statistical methods will be used to summarize the data from this study, with hypothesis testing performed for the primary and other selected efficacy endpoints. Unless stated otherwise, the term "descriptive statistics" refers to number of subjects (n), mean, median, standard deviation (SD), minimum, and maximum for continuous data and frequencies and percentages for categorical data.

All data collected during the study will be included in data listings. Unless otherwise noted, the data will be sorted first by subject number and then by phase (induction, maintenance), cycle/visit, and date.

Unless otherwise specified, all statistical testing will be two-sided, and will be performed using a significance (alpha) level of 0.05.

The statistical analyses will be conducted with the SAS® software package version 9.2 or higher. All analyses will be subject to formal verification procedures. Specifically, results will be verified utilizing independent programming prior to issuance of the draft statistical report. All documents will be verified by the lead statistician to ensure accuracy and consistency of analyses.

5.1. SUBJECT DISPOSITION, DEMOGRAPHIC, AND BASELINE CHARACTERISTICS

The number of subjects enrolled into the study, the number in each analysis population, major protocol violations, number of subjects who discontinue early, reasons for discontinuation, and duration of treatment will be summarized in tabular format. Demographic and background characteristics (including baseline medical history and cancer history) will be summarized descriptively.

5.2. EFFICACY ANALYSIS

5.2.1. Primary Efficacy Endpoint

The primary outcome measure is the complete response rate in subjects with refractory /relapsed CIS or CIS plus papillary disease (i.e., cohorts 1 and 2) following the initiation of Vicinium therapy.

Patients who discontinue or die before Month 3 will be included in the denominator of the primary analysis.

For the primary efficacy endpoint analysis, the following conventions will be utilized in determining treatment failure or complete response at a given visit.

1. Subjects that have centrally read pathology findings ('CIS','any-grade T1', 'high-grade Ta') or centrally read cytology results ('High Grade Urothelial Carcinoma') that indicate treatment failure for a given visit will be considered to have treatment failure.
2. Subjects who met a treatment failure criterion at a previous visit are treatment failures at all subsequent visits.
3. Subjects who have discontinued the study on or prior to the visit and have missing data that would be necessary to generate response are considered non-CRs.
4. Subjects who do not meet criterion 1-3 are complete responders.

5.2.2. Primary Efficacy Analysis

The number and percentage of subjects with CIS or CIS plus papillary disease who have a complete response will be summarized separately by cohort (Cohort 1 or Cohort 2) and overall (Cohort 1 + Cohort 2) at each assessment. Ninety-five percent (95%) confidence intervals around the complete response rate at each assessment will be calculated using the Clopper-Pearson method.

The calculations of the lower (P_L) and upper (P_U) confidence limits using the Clopper-Pearson method are below:

$$P_L = \left(1 + \frac{n - n_1 + 1}{n_1 F\left(1 - \frac{\alpha}{2}, 2n_1, 2(n - n_1 + 1)\right)}\right)^{-1}$$

$$P_U = \left(1 + \frac{n - n_1}{(n_1 + 1)F\left(\frac{\alpha}{2}, 2(n_1 + 1), 2(n - n_1)\right)}\right)^{-1}$$

Where

n = the number of subjects with CIS or CIS plus papillary disease

n_1 = the number of subjects with CIS or CIS plus papillary disease who have a complete response

$F(\alpha, b, c)$ is the αth percentile of the F distribution with b and c degrees of freedom.

Duration of response will be estimated (Kaplan-Meier Estimate) for those subjects who experience a complete response. Subjects who do not experience treatment failure or death will be censored at the time of the last non-missing response assessment.

5.2.3. Sensitivity Assessments for the Primary Efficacy Analysis

Given that complete response is assumed in cases where there is no evidence of treatment failure, it is important to evaluate the impact, if any, of missing data on the primary efficacy analysis. To this end, a sensitivity analysis will be conducted where subjects who have a missing centrally read pathology and/or centrally read cytology at a visit where there was an abnormal cystoscopy will be imputed as a treatment failure for that visit if the subsequent visit indicates treatment failure. A second sensitivity analysis will be performed where subjects with low grade Ta are treated as a failure.

5.2.4. Additional Efficacy Endpoints

Additional efficacy endpoints include:

- Event-free survival (EFS) in all subjects
- Cystectomy-free days in all subjects
- Time to disease recurrence in papillary subjects
- Progression-free survival (PFS) in all subjects
- Overall survival (OS) in all subjects

5.2.5. Additional Efficacy Analyses

The following analyses will be conducted in the mITT population, the evaluable population, the induction phase dosing compliant population, and the study dosing compliant population.

EFS will be estimated using the method of Kaplan-Meier. Subjects without evidence of a treatment failure who are diagnosed with upper or lower urinary tract urothelial carcinoma by histology, cytology or radiographic evidence will be censored at the date of diagnosis. Subjects without evidence of a treatment failure who undergo cystectomy will be censored at the date of cystectomy. All other subjects will be censored at the time of the last assessment. The ninety-five percent (95%) confidence interval around the EFS rate at 18 months will be presented.

Time to disease recurrence will be estimated separately for each cohort using the method of Kaplan-Meier. For subjects in Cohort 3, time to disease recurrence is defined as the number of days from first dose of study treatment to the first occurrence of treatment failure. Subjects who do not experience treatment failure will be censored at the last assessment.

Cystectomy-free days, time to progression, progression-free survival, and overall survival will be estimated based on the method of Kaplan-Meier.

5.2.6. Subgroup Analyses

The planned efficacy analyses will be repeated for each subgroup.

5.3. SAFETY

5.3.1. Adverse Events

Adverse events will be mapped to a Medical Dictionary for Regulatory Activities (MedDRA) version 18.0 or higher preferred term and system organ classification. The severity of all AEs will be assessed

by the investigator according to the protocol and the CTCAE v 4.03 criteria. If a subject experiences multiple events that map to a single preferred term, the greatest severity grade and strongest investigator assessment of relation to study medication will be assigned to the preferred term for the appropriate summaries. Should an event have a missing severity or relationship, it will be classified as having the highest severity and/or strongest relationship for the purposes of summary analyses. For presentation in the data listings, the missing values will be noted. The occurrence of TEAEs will be summarized by phase (induction vs maintenance) using preferred terms, system organ classifications, and severity. Separate summaries of treatment-emergent SAEs, TEAEs related to Vicinium, TEAEs leading to discontinuation of study or study drug, serious TEAEs, and Grade 3 or higher TEAEs will be generated. All AEs reported will be listed for individual subjects showing both verbatim and preferred terms. Missing onset dates will be imputed as previously outlined in Section 3.3.5 as required to determine treatment-emergent events.

Summaries of TEAEs, TEAEs related to Vicinium, Serious TEAEs, and Grade 3 or higher TEAEs will also be presented by each subgroup.

5.3.2. Study Drug Exposure and Concomitant Medications

Exposure to study drug will be summarized with quantitative descriptive statistics.

Previous and concomitant medications will be coded using the World Health Organization (WHO) dictionary (Version: Q4 March 2015). Concomitant medications will be summarized by frequency of drug classification and generic drug name. Previous and concomitant medications will be presented in a data listing

5.3.3. Clinical Laboratory Assessments

Descriptive summaries of all quantitative clinical laboratory results will be presented by phase and visit, as appropriate. Laboratory abnormalities will be graded according to the NCI CTCAE version 4.03 grading system (http://evs.nci.nih.gov/ftp1/CTCAE/CTCAE_4.03_2010-06-14_QuickReference_5x7.pdf). The number and percentage of subjects experiencing treatment-emergent graded laboratory toxicities will be summarized. A laboratory abnormality is considered a treatment-emergent toxicity if it is an increase of 1 grade or more from the baseline toxicity value. Laboratory toxicity shifts from baseline to worst post-treatment toxicity will also be presented.

A spaghetti plot of clinical assessment by visit for subjects with a treatment emergent Grade 3 or higher lab abnormality. In addition scatter plots of lab values versus baseline will be produced for select labs by visit.

All clinical laboratory toxicities that occurred before the initiation of study drug will be excluded from the tables but will be included in the listings.

5.3.4. Other Safety Analyses

Descriptive summaries of vital signs and their change from baseline will be presented by visit. The following values of vital signs will be considered potentially significant:

- Diastolic blood pressure
  - $\leq$ 55 mmHg and/or decrease of $\geq$ 20 mmHg from baseline

  - ≥ 90 mmHg and/or increase of ≥ 30 mmHg from baseline
- Systolic blood pressure
  - ≤ 90 mmHg and/or decrease of ≥ 30 mmHg from baseline
  - ≥ 160 mmHg and/or increase of ≥ 40 mmHg from baseline
- Pulse
  - ≤ 50 bpm and/or decrease of ≥ 30 bpm from baseline
  - ≥ 120 bpm and/or increase of ≥ 30 bpm from baseline
- Temperature > 38.3°C

The number and percentage of subjects with potentially significant vital signs will be summarized by visit.

Results from the triplicate 12-lead ECGs will be averaged and summarized by visit. Corrected QT values will be calculated using the Fridericia and Bazett corrections. The number and percentage of subjects with QTcF values > 470 msec for females and > 450 msec for males will be summarized by visit.

Medical history, physical examination findings, and pregnancy test results will be presented in the data listings.

6. PROTOCOL DEVIATIONS

All protocol deviations will be reviewed by the project team prior to database lock to identify subjects with major protocol deviations.

Reported deviations from the protocol will be listed by subject and date. The type of deviation along with a description and any additional comments about the deviation will be listed.

7. CHANGES IN THE PLANNED ANALYSES

Based on FDA Guidance an analysis of Cohort 1 + Cohort 2 is now planned as the primary efficacy population. In particular, the FDA has stated that recurrence of bladder cancer within 12 months of BCG is an acceptable study population for investigational agents in this setting.

No other changes to the protocol planned analyses have been described. Should changes arise, these changes will be fully described in the Clinical Study Report.

Following the completion of the protocol-specified analyses, the sponsor may choose to conduct additional exploratory analyses. Such analyses will be detailed in a separate document.

8. REVISION HISTORY

Date	Revision	Rationale
27OCT2020	Updated Section 3.3.8 by study visit windows for maintenance month and long-term follow-up windows	Updated for 3 months to cover 91 days (i.e. 13 weeks) instead of 12 weeks for consistency with response assessments which occur every 13 weeks during maintenance.
27OCT2020	Removed third sensitivity analysis for efficacy endpoints	Data collection did not allow for determination of upper and lower tract disease.

9. PROGRAMMING CONVENTIONS

- Page orientation, margins, and fonts: Summary tables, listings, and figures will appear in landscape orientation. There should be a minimum of a 1.25" boundary on the upper (bound) edge, and a minimum of a 1.0" boundary on the remaining three edges. Output should be printed in Courier New with a point size of 8. Titles may be printed using a larger font (e.g., Arial point size 10).
- Identification of analysis population: Every summary table and figure should clearly specify the analysis population being summarized. Listings will be prepared for all enrolled subjects.
- Group headers: In the summary tables, the group headers will identify the cohort and the within-group sample size for the indicated analysis population. Of note, the header's sample size does not necessarily equal the number of subjects actually summarized within any given summary module; some subjects in the analysis population may have missing values and thus may not be summarized.
- Suppression of percentages corresponding to null categories: When count data are presented as category frequencies and corresponding percentages, the percent should be suppressed when the count is zero in order to draw attention to the non-zero counts.
- Presentation of sample sizes: Summary modules should indicate, in one way or another, the number of subjects actually contributing to the summary statistics presented in any given summary module. As mentioned above, this may be less than the number of subjects in the analysis population.
    - In the quantitative modules describing continuous variables (and thus presenting sample size, means, and standard deviations), the sample size should be the number of non-missing observations
    - For categorical variables that are presented in frequency tables, the module should present the total count in addition to the count in each category. Percentages should be calculated using this total as the denominator, and the percentage corresponding to the sum itself (that is, 100%) should be presented to indicate clearly to a reviewer the method of calculation.
- Sorting: Listings will be sorted first by subject number and then by phase (induction, maintenance), visit, and date, if applicable. If a listing is sorted in a different manner, a footnote will indicate as such.
- General formatting rules: Rounding for all variables will occur only as the last step, immediately prior to presentation in listings, tables, and figures. No intermediate rounding will be performed on derived variables. The standard rounding practice of rounding numbers ending in 0-4 down and numbers ending in 5-9 up will be employed.
- The presentation of numerical values will adhere to the following guidelines:
    - Raw measurements will be reported to the number of significant digits as captured electronically or on the CRFs.
    - Standard deviations will be reported to one decimal place beyond the number of decimal places the original parameter is presented.
    - Means will be reported to the same number of significant digits as the parameter.
    - Calculated percentages will be reported with no decimals.

- Dates will be formatted as DDMMMYYYY. Partial dates will be presented on data listings as recorded on CRFs.
    - Time will be presented according to the 24-hour clock (HHMM).

10. PROPOSED TABLES, LISTINGS, AND FIGURES DSMB ANALYSES

Summary Tables

Table 1 Eligibility, Enrollment, and Treatment Status, All Consented Subjects

Table 2 Subject Termination from Study, Safety Population

Table 3 Demographics, Safety Population

Table 4 Baseline Disease Characteristics, Safety Population

Table 5 Treatment-Emergent Adverse Events by System Organ Classification, Preferred Term, and Greatest Severity, Safety Population

Table 6 Treatment-Emergent Adverse Events Related to Vicinium by System Organ Classification, Preferred Term, and Greatest Severity, Safety Population

Table 7 Treatment-Emergent Serious Adverse Events by System Organ Classification, Preferred Term, and Greatest Severity, Safety Population

Table 8 Clinical Laboratory Assessments by Visit, Safety Population

Table 9 Graded Laboratory Toxicities, Safety Population

Table 10 Worst Post-Baseline Laboratory Shifts, Safety Population

Table 11 Vital Signs, Weight, and Height by Visit, Safety Population

Table 12 Electrocardiogram by Visit, Safety Population

Table 13 Summary of Complete Response by Visit, Ongoing Subjects Are Excluded as 'Missing', mITT Population

Data Listings

Listing 1 Study Completion, Safety Population

Listing 2 Demographics, Safety Population

Listing 3 Adverse Events, Safety Population

Listing 3.1 Serious Adverse Events, Safety Population

Listing 4.1.1 Select Chemistry Results for Subjects with a Post-Baseline Abnormality, Safety Population

Listing 4.1.2 Clinical Chemistry Laboratory Data, Safety Population

Figures

11. TABLES, LISTINGS, AND FIGURES 6 MONTH ANALYSIS

Summary Tables

Table 1 Eligibility, Enrollment, and Treatment Status, All Consented Subjects

Table 2 Subject Termination from Study, Safety Population

Table 3 Demographics, Safety Population

Table 4 Baseline Disease Characteristics, Safety Population

Table 5 Treatment-Emergent Adverse Events by System Organ Classification, Preferred Term, and Greatest Severity, Safety Population

Table 6 Treatment-Emergent Adverse Events Related to Vicinium by System Organ Classification, Preferred Term, and Greatest Severity, Safety Population

Table 7 Treatment-Emergent Serious Adverse Events by System Organ Classification, Preferred Term, and Greatest Severity, Safety Population

Table 8 Clinical Laboratory Assessments by Visit, Safety Population

Table 9 Graded Laboratory Toxicities, Safety Population

Table 10 Worst Post-Baseline Laboratory Shifts, Safety Population

Table 11 Vital Signs, Weight, and Height by Visit, Safety Population

Table 12 Electrocardiogram by Visit, Safety Population

Table 13 Summary of Complete Response by Visit, Ongoing Subjects Are Excluded as 'Missing', mITT Population

Table 14 Cystectomy-Free Days, mITT Population

Data Listings

Listing 1 Study Completion, Safety Population

Listing 2 Demographics, Safety Population

Listing 3 Adverse Events, Safety Population

Listing 3.1 Serious Adverse Events, Safety Population

Listing 4.1.1 Select Chemistry Results for Subjects with a Post-Baseline Abnormality, Safety Population

Figures

12. PROPOSED TABLES, LISTINGS, AND FIGURES FINAL ANALYSIS

Summary Tables

Accountability and Baseline Characteristics

14.1.1 Eligibility, Enrollment and Treatment Status
14.1.2 Subject Termination from Study
14.1.3 Demographics
14.1.4 Baseline Disease Characteristics

Efficacy

14.2.1.1.1 Summary of Complete Response by Visit – Primary Efficacy Population
14.2.1.1.2 Summary of Complete Response by Initial Disease Group and Visit – Primary Efficacy Population
14.2.1.1.3 Summary of Complete Response by Subgroup and Visit – Primary Efficacy Population
14.2.1.2.1 Summary of Complete Response by Visit, Sensitivity Assessments – Primary Efficacy Population
14.2.1.2.2 Summary of Complete Response by Visit, Sensitivity Assessments – Low Grade Ta as Failure - Primary Efficacy Population
14.2.1.4 Summary of Complete Response by Visit – Evaluable Population
14.2.1.5 Summary of Complete Response by Visit – Induction Phase Dosing Compliant Population
14.2.1.6 Summary of Complete Response by Visit – Study Dosing Compliant Population
14.2.1.7.1 Summary of Recurrence-free Subjects by Visit – mITT Population - Cohort 3
14.2.1.7.2 Summary of Recurrence-free Subjects by Initial Disease Group and Visit – mITT – Cohort 3
14.2.1.8 Summary of Recurrence-free Subjects by Visit – Evaluable Population
14.2.1.9 Summary of Recurrence-free Subjects by Visit – Induction Phase Dosing Compliant Population
14.2.1.10 Summary of Recurrence-free Subjects by Visit – Study Dosing Compliant Population
14.2.2.1.1 Duration of Response – Primary Efficacy Population
14.2.2.1.2 Duration of Response by Age Subgroup – Primary Efficacy Population
14.2.2.1.3 Duration of Response by Gender Subgroup – Primary Efficacy Population
14.2.2.1.4 Duration of Response by Region Subgroup – Primary Efficacy Population
14.2.2.1.5 Duration of Response by Race Subgroup – Primary Efficacy Population
14.2.2.1.6 Duration of Response by Prior BCG Therapies Subgroup – Primary Efficacy Population
14.2.2.1.7 Duration of Response by Instillation Subgroup – Primary Efficacy Population
14.2.2.2.1 Duration of Response, Sensitivity Assessments – Primary Efficacy Population
14.2.2.2.2 Duration of Response, Sensitivity Assessments - Low Grade Ta as Failure – Primary Efficacy Population
14.2.2.6 Duration of Response – Evaluable Population
14.2.2.7 Duration of Response – Induction Phase Dosing Compliant Population
14.2.2.8 Duration of Response – Study Dosing Compliant Population
14.2.3.1.1 Event-free Survival – mITT Population
14.2.3.1.2 Event-free Survival by Age Subgroup – mITT Population
14.2.3.1.3 Event-free Survival by Gender Subgroup – mITT Population
14.2.3.1.4 Event-free Survival by Region Subgroup – mITT Population
14.2.3.1.5 Event-free Survival by Race Subgroup – mITT Population
14.2.3.1.6 Event-free Survival by Prior BCG Therapies Subgroup – mITT Population
14.2.3.1.7 Event-free Survival by Instillation Subgroup – mITT Population

Safety

14.3.1.1.1 Treatment-Emergent Adverse Events by System Organ Classification, Preferred Term, and Greatest Severity
14.3.1.1.2 Treatment-Emergent Adverse Events By System Organ Classification, Preferred Term, and Greatest Severity – Age Subgroup
14.3.1.1.3 Treatment-Emergent Adverse Events By System Organ Classification, Preferred Term, and Greatest Severity – Gender Subgroup
14.3.1.1.4 Treatment-Emergent Adverse Events By System Organ Classification, Preferred Term, and Greatest Severity – Region Subgroup
14.3.1.1.5 Treatment-Emergent Adverse Events By System Organ Classification, Preferred Term, and Greatest Severity – Race Subgroup
14.3.1.1.6 Treatment-Emergent Adverse Events By System Organ Classification, Preferred Term, and Greatest Severity – Prior BCG Therapies Subgroup
14.3.1.1.7 Treatment-Emergent Adverse Events By System Organ Classification, Preferred Term, and Greatest Severity – Instillation Subgroup
14.3.1.2.1 Treatment-Emergent Adverse Events Related to Vicinium By System Organ Classification, Preferred Term, and Greatest Severity
14.3.1.2.2 Treatment-Emergent Adverse Events Related to Vicinium By System Organ Classification, Preferred Term, and Greatest Severity – Age Subgroup
14.3.1.2.3 Treatment-Emergent Adverse Events Related to Vicinium By System Organ Classification, Preferred Term, and Greatest Severity – Gender Subgroup
14.3.1.2.4 Treatment-Emergent Adverse Events Related to Vicinium By System Organ Classification, Preferred Term, and Greatest Severity – Region Subgroup
14.3.1.2.5 Treatment-Emergent Adverse Events Related to Vicinium By System Organ Classification, Preferred Term, and Greatest Severity – Race Subgroup
14.3.1.2.6 Treatment-Emergent Adverse Events Related to Vicinium By System Organ Classification, Preferred Term, and Greatest Severity – Prior BCG Therapies Subgroup
14.3.1.2.7 Treatment-Emergent Adverse Events Related to Vicinium By System Organ Classification, Preferred Term, and Greatest Severity – Instillation Subgroup
14.3.1.3 Treatment-Emergent Adverse Events Leading to the Discontinuation of Study Drug by System Organ Classification, Preferred Term, and Greatest Severity
14.3.1.4.1 Treatment-Emergent Serious Adverse Events by System Organ Classification, Preferred Term, and Greatest Severity
14.3.1.4.2 Treatment-Emergent Serious Adverse Events by System Organ Classification, Preferred Term, and Greatest Severity – Age Subgroup
14.3.1.4.3 Treatment-Emergent Serious Adverse Events by System Organ Classification, Preferred Term, and Greatest Severity – Gender Subgroup
14.3.1.4.4 Treatment-Emergent Serious Adverse Events by System Organ Classification, Preferred Term, and Greatest Severity – Region Subgroup
14.3.1.4.5 Treatment-Emergent Serious Adverse Events by System Organ Classification, Preferred Term, and Greatest Severity – Race Subgroup
14.3.1.4.6 Treatment-Emergent Serious Adverse Events by System Organ Classification, Preferred Term, and Greatest Severity – Prior BCG Therapies Subgroup
14.3.1.4.7 Treatment-Emergent Serious Adverse Events by System Organ Classification, Preferred Term, and Greatest Severity – Instillation Subgroup
14.3.1.5.1 Treatment-Emergent Adverse Events by System Organ Classification, Preferred Term, and Severity Group
14.3.1.5.2 Treatment-Emergent Adverse Events Occurring in at Least 5% of Subjects by System Organ Classification, Preferred Term, and Severity Group

Data Listings

16.2.1 Study Completion
16.2.2.1 Informed Consent, Eligibility, and Inclusion/Exclusion Criteria
16.2.2.2 Protocol Deviations
16.2.4 Demographics
16.2.5.1.1 Study Drug Administration
16.2.5.1.2 Plasma Vicinium Sample Collection
16.2.5.2.1 Previous and Concomitant Medications
16.2.5.2.2 Medical History
16.2.5.2.3 Bladder Cancer History
16.2.5.2.4 Cancer Treatment History
16.2.5.2.5 BCG Treatment History
16.2.5.2.6 Non-Drug Therapy Procedures
16.2.5.2.7 Anesthesia
16.2.6.1.1 Cystoscopy Procedure
16.2.6.1.2 Cystoscopy Findings
16.2.6.2 Central Pathology Reader
16.2.6.3.1 Urine Cytology
16.2.6.3.2 Urine Cytology Reader
16.2.6.4 Biopsy
16.2.6.5.1 Efficacy Endpoint Summary
16.2.6.5.2 Response Summary – Primary Efficacy Population
16.2.6.5.3 Time to Event Summary – mITT Population
16.2.6.5.4 Time to Event Summary – Evaluable Population
16.2.6.5.5 Time to Event Summary – Induction Phase Dosing Compliant Population
16.2.6.5.6 Time to Event Summary – Study Dosing Compliant Population
16.2.7.1 Adverse Events
16.2.7.2 Serious Adverse Events
16.2.7.3 Adverse Events by Age Group, System Organ Classification, and Preferred Term
16.2.7.4 Related Adverse Events by Age Group, System Organ Classification, and Preferred Term
16.2.7.5 Serious Adverse Events by Age Group, System Organ Classification, and Preferred Term
16.2.8.1.X Laboratory Values
- ♦ Select Chemistry Results for Subjects with a Post-Baseline Abnormality
- ♦ Chemistry
- ♦ Select Hematology Results for Subjects with a Post-Baseline Abnormality
- ♦ Hematology
- ♦ Select Urinalysis Results for Subjects with a Post-Baseline Abnormality
- ♦ Urinalysis (Includes Pregnancy Test)

16.2.8.2 Vital Signs, Height and Weight
16.2.8.3 Physical Examinations
16.2.8.4 12-Lead Electrocardiogram
16.2.8.5 Transurethral Resection of Bladder Tumors (TURBT)
16.2.8.6 Prostatic Urethral Biopsy
16.2.8.7 Radiological Imaging
16.2.8.8 Immunohistochemistry (EpCam) Sample Collection
16.2.8.9 Extravesical Disease Evaluation
16.2.8.10 Post-Treatment Follow-Up
16.2.8.11 Anti-Vicinium Antibodies Collection

Figures