Statistical Analysis Plan: TC-BC-12

Study Title:	A Phase 2b, Single-Arm, Multicenter Trial to Evaluate the Efficacy and Safety of UGN-102 as Primary Chemoablative Therapy in Patients with Low Grade (LG) Non-Muscle-Invasive Bladder Cancer (NMIBC) at Intermediate Risk of Recurrence
Study Number:	TC-BC-12
Study Phase:	2b
Sponsor:	UroGen Pharma 499 Park Avenue, Suite 1200 New York, NY 10022
Version	Final
NCT #:	NCT03558503
Date:	24 July 2020

Confidentiality Statement

1 TABLE OF CONTENTS

LIST OF TABLES

LIST OF APPENDICES

2 SIGNATURE PAGE

Study Title: A Phase 2b, Single-Arm, Multicenter Trial to Evaluate the Efficacy and Safety of UGN-102 as Primary Chemoablative Therapy in Patients with Low Grade (LG) Non-Muscle-Invasive Bladder Cancer (NMIBC) at Intermediate Risk of Recurrence

Study Number: TC-BC-12

Prepared by: DocuSigned by: Hung Lam — Date: 7/27/2020
Hung Lam, PhD
Principal Statistician I
MMS Holdings, Inc.

Reviewed by: DocuSigned by: Harry Haber — Date: 7/27/2020
Harry Haber, MPH
Principal Statistician I
MMS Holdings, Inc.

Approved by: DocuSigned by: Soumi Lahiri — Date: 7/27/2020
Soumi Lahiri, PhD
Sr. Director, Biostatistics
UroGen Pharm, Inc.

Approved by: DocuSigned by: Sunil Raju — Date: 7/27/2020
Sunil Raju, MD
Sr. Director, Clinical Development & Pharmacovigilance
UroGen Pharm, Inc.

Approved by: DocuSigned by: Elyse Seltzer — Date: 7/27/2020
Elyse Seltzer, MD
SVP Clinical Development
UroGen Pharma Inc.

3 LIST OF ABBREVIATIONS AND DEFINITIONS OF TERMS

AE	adverse event
AESI	adverse event of special interest
ALP	alkaline phosphatase
ALT	alanine aminotransferase
AST	aspartate aminotransferase
ATC	anatomical therapeutic class
AUC	area under the time-concentration curve
BMI	body mass index
CI	confidence interval
CDISC	clinical data interchange standards consortium
C_{max}	maximum plasma drug concentration
COVID-19	coronavirus disease 2019
CR	complete response
CTCAE	common terminology criteria for adverse events
DBL	direct bilirubin
DBP	diastolic blood pressure
DCR	durable complete response
DFS	disease-free survival
DOR	duration of CR
eCRF	electronic case report form
EOS	end of study
HRQoL	health related quality of life
IP	investigational product
ITT	intent-to-treat
LFT	liver function test
LG	low grade
LLN	lower limit of the normal range
LOCB	last observation carried backward

LOCF	last observation carried forward
MedDRA	Medical Dictionary for Regulatory Activities
mITT	modified intent-to-treat
MMC	Mitomycin C
NMIBC	non-muscle-invasive bladder cancer
PCS	potentially clinically significant
PK	pharmacokinetic
PP	per protocol
PT	preferred term
QLQ-NMIBC24	24-item survey to assess quality of life for patients with NMIBC
SAE	serious adverse event
SAP	statistical analysis plan
SAS	Statistical analysis system
SBP	systolic blood pressure
SMQ	standardized MedDRA queries
SOA	schedule of activities
SOC	system organ class
TBL	total bilirubin
TEAE	treatment-emergent adverse event
TLFs	tables, listings, figures
t_{max}	time to maximum plasma concentration
TTR	time to recurrence
ULN	upper limit of the normal range

4 INTRODUCTION

The purpose of this statistical analysis plan (SAP) is to describe the framework for the reporting, summarization, and statistical analysis methodology of the safety and efficacy parameters measured throughout the study. It is based on Protocol TC-BC-12 version 5.0 dated 15 April 2019.

5 TRIAL OBJECTIVES

5.1 Primary Objective

To evaluate the tumor ablative effect of UGN-102 in patients with Low Grade (LG) Non-Muscle-Invasive Bladder Cancer (NMIBC).

5.2 Secondary Objectives

- To evaluate the durability of response in patients with LG NMIBC who achieve Complete Response (CR) at the 3-MONTH Visit.
- To evaluate the safety and tolerability of intravesical UGN-102 instillations in patients with LG NMIBC.

5.3 Tertiary Objective

To assess the pharmacokinetic (PK) profile of the first Mitomycin C (MMC) instillation in 6 patients.

5.4 Exploratory Objective

To examine patient reported outcome of health-related quality of life (HRQoL) as compared to baseline.

6 STUDY DESIGN CONSIDERATIONS

6.1 Study Design

This study is a prospective, proof-of-concept, open-label, single-arm, multicenter Phase 2b trial designed to assess the efficacy and safety of UGN-102 treatment instilled in patients diagnosed with LG NMIBC, including newly-diagnosed patients, and determined to have intermediate risk of progression, defined as 1 or 2 of the following: multiple tumors, tumors >3 cm, or recurrence (≥1 occurrence of LG NMIBC within 1 year of the current diagnosis). Eligible patients will be treated with 6 weekly instillations of UGN-102.

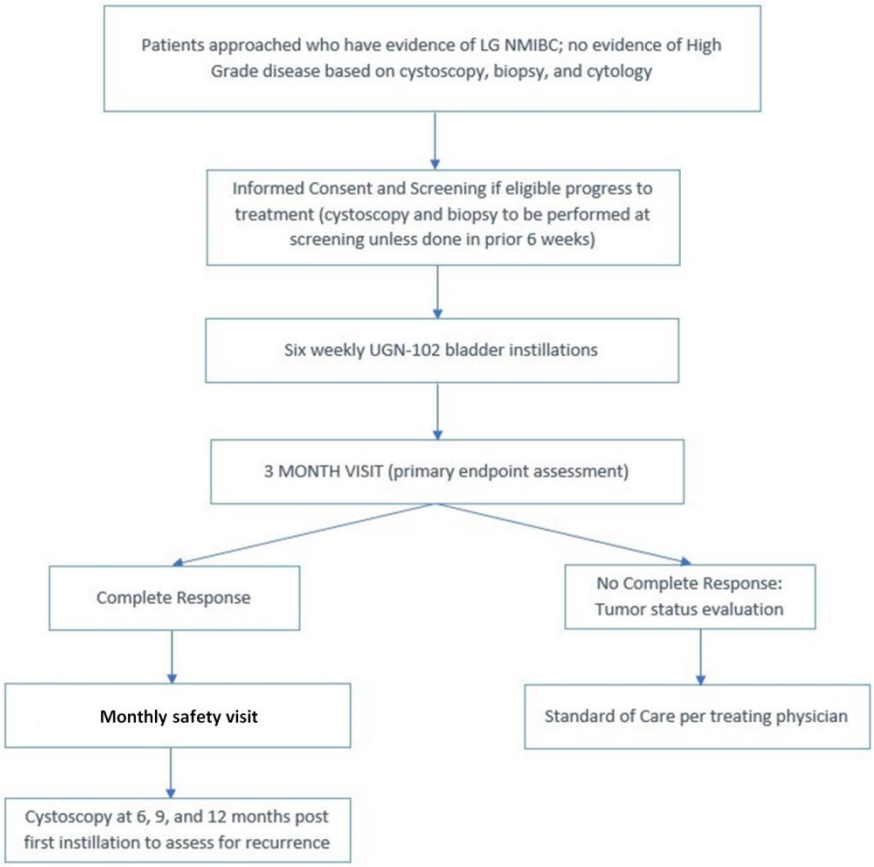
Upon signing of informed consent, the patients will undergo a Screening Visit for an eligibility evaluation. The screening period is 2 weeks if the patient's cold cup biopsy to confirm LG tumor is historic (taken within 6 weeks of the screening visit) or 4 weeks if the patient's cold cup biopsy to confirm LG tumor is current (taken at the screening visit).

Starting at the Baseline Visit (Day 1), eligible patients will be treated with UGN-102 once weekly for a total of 6 doses; the UGN-102 concentration to be used in this trial will be 1.33 mg MMC per 1 mL. The volume of UGN-102 to be instilled will be 56 mL (75 mg of MMC).

The ablative effect of UGN-102 will be evaluated at the 3-month assessment, which will take place 5 weeks ±1 week after the last weekly instillation (3 months after initiation of treatment). Response will be determined based on visual evaluation by cystoscopy (appearance, number, and size of the lesions) and, if there are remaining lesions, by histopathology of the remaining lesions. CR is defined as having no detectable disease and will be assessed visually during cystoscopy and also upon urine cytology. In the event that the investigator is not sure, and there is suspect tissue, a small biopsy will be taken from the suspect tissue to confirm CR in addition to cystoscopy and urine cytology. Patients who achieve a CR will continue to have monthly telephone contacts to document any adverse events and changes in concomitant mediations and will be assessed at 6, 9, and 12 months after the first instillation of UGN-102 for evidence of disease recurrence. The group of patients considered to be non-responders (non-CR) will discontinue the study and continue with standard of care according to their treating physician.

Safety will be determined based on physical examination, laboratory assessments, and a review of adverse events (AEs) including adverse events of special interest (AESIs). All safety data will be reviewed on an ongoing basis, and qualified per National Cancer Institute Common Terminology Criteria for Adverse Events (CTCAE) as Grade 3 or 4.

6.1.1 Study Schema



LG = low grade; NMIBC = non-muscle-invasive bladder cancer

Schedule of activities for Screening to the 3-MONTH Visit and for CRs at the 3-MONTH Visit: Visit 8 to end of study (EOS) are shown in Appendix 1 and Appendix 2, respectively.

6.2 Statistical Hypotheses

No formal hypotheses will be tested.

6.3 Sample Size Determination

Approximately 66 patients will enter the study, with 60 expected to complete, assuming 10% dropout rate. Given the patient population under study, i.e., those at intermediate risk of recurrence, a more conservative CR rate than previous studies of UGN-102 in the treatment of LG NMIBC has been assumed. The primary objective will be successfully achieved if complete responses from 36 out of 60 patients with CR rate of 60% will be observed, which would yield a 95% confidence interval of (46.5%, 72.4%) based on the exact (Clopper-Pearson) binomial distribution.

6.4 Efficacy Measures

There are 2 efficacy measures. The first one corresponds to primary objective. The second one corresponds to the secondary efficacy objective (refer to Protocol Section 5).

1. CR rate at the 3-MONTH defined as the proportion of patients with CR at the 3-MONTH Visit using the ITT Analysis Set
2. Durable CR (DCR) rate at 6, 9, and 12 months after the first instillation of UGN-102. DCR rate is defined as the percentage of patients who continue to display CR at Study Visits 10, 13, and 16, respectively, after achieving a CR at the 3-MONTH using the 3-MONTH$_{CR}$ Analysis Set.

6.5 Safety Measures

The safety of intravesical UGN-102 instillations will be assessed by the following:

- Evaluation (frequency, seriousness, and severity) of treatment emergent AEs and treatment-emergent AESIs.

  AESIs are defined as:

  - Overdose of an IP
  - Suspected abuse/misuse of IP
  - Inadvertent or accidental exposure to IP
  - Medication error involving IP (with or without patient/patient exposure to the Sponsor medicinal product
  - Identified significant safety issue either from an individual case report or review of aggregate data
  - Allergic reaction to MMC
  - Unexpected adverse drug reaction
  - Voiding interruption due to urethral/penile edema (unrelated to prostatic hypertrophy
  - Indication of bone marrow suppression
  - Lower urinary tract symptoms
- Changes from baseline in laboratory values and incidence of measurements defined as potentially clinically significant (PCS)
- Changes in physical examination findings (including urology-oriented physical examination)
- Changes from baseline in vital sign values and incidence of measurements defined as PCS changes

6.6 Pharmacokinetic Parameters

The pharmacokinetic (PK) profile of the first MMC instillation during the 6 hours following the first administration of UGN-102 will be based on PK sampling at 0 (pre-instillation), 0.5, 1, 2, 3, 4, 5, and 6 hours post instillation. PK will be performed with just 6 patients who provide informed consent for PK testing. The following PK parameters will be provided:

- AUC_{0-6} – Area under the concentration-time curve from time zero to 6 hours post-instillation on Day 1, calculated using the linear trapezoidal rule
- C_{max} – Observed maximum plasma concentration.
- t_{max} – Time to reach maximum plasma concentration.

7 STUDY POPULATIONS

7.1 Analysis Populations

7.1.1 Intent-to-Treat (ITT) Analysis Set (Safety Analysis Set)

The ITT analysis set (Safety analysis set), hereafter referred to as ITT analysis set, will consist of all patients who have been enrolled into the trial and who have received any instillation of UGN-102 preparation. ITT analysis set, which will include all data captured in the database for these patients, will serve for all efficacy and safety analyses. With respect to safety displays, a footnote will be added to emphasize ITT and Safety analysis sets are identical for this study.

7.1.2 Modified ITT (mITT) Analysis Set

The mITT analysis set will consist of all patients who have been enrolled into the trial, who have a confirmed diagnosis of LG NMIBC based on the local pathology reading, who received at least 1 instillation of UGN-102 preparation, and who have undergone the 3-MONTH Visit (since it is the first visit for the outcome evaluation). This analysis set will serve as the supportive analysis set for the primary efficacy endpoint analysis.

7.1.3 Per Protocol (PP) Analysis Set

The PP analysis set will consist of all patients who have been enrolled into the trial, who have a confirmed diagnosis of LG NMIBC based on the local pathology reading, who received all 6 weekly instillations of UGN-102 preparation, have undergone the 3-MONTH Visit, and do not have any major protocol deviations that are deemed to affect efficacy outcomes. This analysis set, which will include all data captured in the database for these patients, will serve as the supportive analysis set for the primary efficacy endpoint analysis.

7.1.4 CR at Primary Outcome Analysis Set

The CR at 3-MONTH Visit ($3\text{-}MONTH_{CR}$) analysis set will consist of all patients who achieved CR at the 3-MONTH Visit (Study Visit 7). This analysis set, which will include all data captured in the database for these patients, will serve as the primary analysis set for the secondary efficacy endpoint analysis.

7.1.5 PK Analysis Set

The PK analysis set will consist of all patients who consent to provide blood samples for PK analysis and have at least one plasma sample for evaluation of PK properties of MMC following an initial instillation of UGN-102.

8 CHANGES IN CONDUCT OR PLANNED ANALYSES FROM THE PROTOCOL

1. Definition for Per Protocol analysis set is modified to exclude any patients who have any major protocol deviations that are deemed to affect efficacy outcomes.
2. Analysis of CR at the 3-MONTH visit using mITT analysis set will be performed as an additional supportive analysis.
3. Addition of Kaplan-Meier analysis for DOR, DFS, and TTR. Addition of sensitivity analyses for DOR, DFS, and TTR to address delayed or missed efficacy visits due to COVID-19.
4. Some patients had disruptions to their visit schedules due to COVID-19 situation after completing their 3-MONTH visit. All missed / delayed study visits following the 3-MONTH Visit will be summarized. All protocol deviations due to visits outside the protocol window will be recorded and summarized. Appropriate changes will be made those analyses which are typically summarized by scheduled visits (e.g. laboratory measures, vital signs, QOL etc.).

9 OVERALL STATISTICAL CONSIDERATIONS

9.1 General Conventions

Data will be listed and summarized according to Clinical Data Interchange Standards Consortium (CDISC) standards. SAS software will be used to perform all data analyses, generate tables, figures, and listings.

All measured variables and derived parameters will be listed individually and, if appropriate, tabulated using descriptive statistics. For categorical variables, summary tables will be provided giving sample size, absolute and relative frequency, and 95% confidence interval (CI) for proportions using exact approach. For continuous variables, summary tables will be provided giving sample size, arithmetic mean, standard deviation, coefficient of variation (if appropriate), median, minimum and maximum, percentiles. For variables with non-normal distribution, medians and interquartile range will be calculated, and the appropriate non-parametric tests will be applied.

Change from baseline for a measured variable at a particular post-baseline time point will be computed as the value at the post-baseline time point minus the baseline value. If either the baseline or visit value is missing, the change from baseline is set to missing as well.

Generally, only pre-specified planned times will be used in the summaries, statistical analyses and calculations of any derived parameters of safety data; unscheduled readings will be listed. Summaries of safety data will include data from unscheduled visits only in sections labelled “worst case on-therapy”.

Percentages will be displayed with one decimal place and percentages for zero counts will be omitted from the presented results.

Day 1 will be defined as the first date on which UGN-102 was administered. Positive study days will be counted forward from Day 1. Day -1 will be the date immediately preceding Day 1, and negative study days will be counted backward from Day 1.

Deviations from the analyses in the SAP will be identified in the CSR.

9.2 Reference Date

There are three reference dates:

- Informed consent date will be used as a reference date for calculating age.
- The primary efficacy endpoint reference date is the primary disease assessment at the 3-MONTH visit, and will be used to calculate duration of CR.
- The safety and other efficacy reference date is the treatment start date, and will be used to calculate study day for safety and efficacy measures.

9.3 Study Day for Safety Measures

If the date of interest occurs on or after the safety reference date then the safety study day will be calculated as (date of interest - safety reference date) + 1. If the date of interest occurs before the safety reference date then the safety study day will be calculated as (date of interest – safety reference date). There is no safety study day 0.

9.4 Study Day for Efficacy

Same as study day for safety measures (Section 9.3), since the safety and efficacy reference dates are the same for this study.

9.5 Calculation of Duration

Durations (e.g. duration of adverse event, duration of response, etc.) will be calculated as stop date minus start date plus one.

When reporting time to event durations (e.g., DOR, DFS, TTR) for efficacy and safety measures in months, divide the number of days by 30.4375; to report in weeks divide the number of days by 7; to report in years divide the number of days by 365.25.

The algorithm below will be used to compute integer age accounting for the actual number of days in months or years between date of birth (date1) and informed consent date (date2)

intck('year', date1, date2+1) - (month(date2+1)<month(date1) or
(month(date2+1)=month(daste1) and day(date2+1)<day(date1)))

9.6 Baseline definition

Baseline will be defined as the most recent non-missing value prior to the first instillation of UGN-102.

Parameter	Baseline Days Collected			Baseline To Be Used in Analysis / Summaries
	Day -2	Day -1	Day 1	
Safety :				
Laboratory		X		Day -1
Vital Signs			X	Day 1[1]

1. Use the mean of replicate assessments at any given time point as the value for that time point in all summaries, figures and statistical analyses.

9.7 Handling of Missing Data

Imputation rules for partial or missing start/stop date for AEs and medical history are detailed in Appendix 3. Imputations done for partial or missing start/stop dates will only be used to classify events as treatment-emergent, while the listings will display the dates as collected on the case report form. For these date imputations, an estimated study day will be determined based on the imputation rules and will be flagged as such in the listings.

Imputation rules for partial or missing start/stop date for prior and concomitant medications and surgical procedures are detailed in Appendix 4. Imputations done for partial or missing start/stop dates will only be used to classify medications or surgical procedures as concomitant, while the listings will display the dates as collected on the case report form.

Other than above and the specific rules for imputing missing disease assessment for durable CR as described in Section 11.3.1, no imputation of missing data will be performed.

9.8 Interim Analysis

The purpose of the interim analysis is to obtain interim complete response (CR) and non-complete response data for informational purposes. The trial will not be stopped as a result of this analysis. Details of the interim analysis will be described in a separate interim analysis plan.

9.9 Pooling Strategy for Study Sites

Data will be pooled across all study sites.

10 STATISTICAL ANALYSIS METHODS

10.1 Patient Disposition

The number and percentages of patients for each analysis set (ITT, mITT, PP, 3-MONTH$_{CR}$, and PK) will be summarized. For computing percentages, the denominator will be the number of patients in the ITT analysis set.

Patient disposition will be summarized for the following study periods:

- **Screening Period**:

The following patient disposition as recorded on the Inclusion/Exclusion Criteria and Enrollment eCRF will be tabulated using the number and percentages of patients:

- Screened
- Screen failures
- Primary reasons for screen failure comprise failed to meet the inclusion/exclusion criteria, consent withdrawn, adverse event/death and other

For computing percentages, the denominator will be the number of patients screened.

- **Treatment Period:**

  The following patient disposition as recorded on the 1-6 Weeks Treatment Early Termination eCRF will be tabulated using the number and percentages of patients:

  - Completed treatment
  - Discontinued treatment
  - Primary reasons for discontinuation of treatment comprises adverse event, consent withdrawal, investigator discretion, pregnancy, noncompliance, lost to follow-up, death and other

  For computing percentages, the denominator will be the number of patients in the ITT analysis set.

- **End of Study:**

  The following patient disposition as recorded on the Study Completion eCRF will be tabulated using the number and percentages of patients by non-CR, CR, and Total (non-CR and CR combined):

  - Completed study
  - Discontinued study
  - Primary reasons for discontinuation of study comprises adverse event, consent withdrawn, investigator discretion, pregnancy, noncompliance, lost to follow-up, death and other

  Non-CR completers are patients who completed the end of study at the 3-MONTH visit as dictated by the protocol. CR completers are patients who completed the follow-up period or had recurrence at 6 or 9 months after the first instillation of UGN-102 as recorded on eCRF. For computing percentages, the denominator will be the number of non-CR patients, CR patients, and all patients in the ITT analysis set for the respective column.

A data listing of patient disposition will be provided.

Duration of Follow-up

Descriptive statistics (mean, standard deviation, median, 1st and 3rd quartile, minimum, and maximum) for duration, expressed in months, of study follow-up computed as [End of study date – first dose date + 1]/30.4375. This analysis will be based on ITT and 3-MONTH$_{CR}$ analysis sets.

10.2 Demographics and Baseline Characteristics

Demographics and baseline characteristics will be summarized using the ITT analysis set. Continuous variables such as age, BMI, weight, and height will be summarized using descriptive statistics (n, mean, standard deviation, and median, minimum, maximum). Categorical variables including but not limited to age group (<65, 65 to <75, ≥75 years), BMI category (<30 kg/m^2, ≥30 kg/m^2), sex, race, ethnicity, treatment course (full course - 6 instillations, partial course <6 instillations) , tumor size (≤3 cm, >3 cm), tumor type, visual appearance of the lesions, and previous LG NMIBC episodes within 1 year of the current diagnosis (Yes, No) will be summarized using counts and percentages.

10.3 Medical History

Medical history reported terms will be coded to a SOC and PT using the most recent version of the Medical Dictionary for Regulatory Activities (MedDRA) dictionary version 21.0 or later. Number and percent of patients reporting any medical history by SOC and PT for the ITT analysis set will be provided. A patient with multiple medical conditions will be counted once per SOC and PT. For computing percentages, the denominator will be the number of patients in the ITT analysis Set. The summary table will be sorted by descending order of frequency of SOC (then alphabetically for ties), then by descending order of frequency of PT within each SOC (then alphabetically for ties).

10.4 Prior and Concomitant Medications

All medications will be coded using the World Health Organization Drug Dictionary, March 1, 2018 version or later. The Anatomical Therapeutic Class (ATC) Level 3 and preferred name will be used to list and summarize the data. Analysis of prior and concomitant medications use will be performed in the following manner:

- **Pre-instillation (prior) medications use:** Medications that were administered prior to the first instillation of UGN-102, regardless if stopped after the first instillation, will be included for the summary of prior medications. The number and percentage of patients reporting the use of prior medications by ATC Level 3 and preferred name will be summarized for the ITT analysis set. For computing percentages, the denominator will be the number of patients in the ITT analysis set.
- **Concomitant medications use:** Concomitant medications that were administered following the first instillation of UGN-102 through end of study regardless if concomitant medication initiation was before or after the first instillation will be included for the summary of concomitant medications. The number and percentage of patients reporting the use of concomitant medications by ATC Level 3 and preferred name will be summarized for the ITT analysis set. For computing percentages, the denominator will be the number of patients in the ITT analysis Set.

All prior and concomitant medications will be included in a data listing.

10.5 Prior and Concomitant Surgical Procedures

All surgical procedures will be coded to a SOC and PT using the most recent version of the MedDRA dictionary version 21.0 or later. Prior and concomitant surgical procedures will be summarized similar to that described for prior and concomitant medication use in Section 10.4. In addition, incidence of prior TURBT interventions will be summarized.

All prior and concomitant surgical procedures will be included in a data listing.

10.6 Analysis of Pharmacokinetics

The plasma concentration versus time profiles collected for MMC will be summarized using descriptive statistics for predose, and 0.5, 1, 2, 3, 4, 5, and 6 hours after the first instillation of UGN-102 for patients in the PK analysis set. The following statistics will be calculated for plasma concentrations at each time point: arithmetic mean, standard deviation and coefficient of variation, geometric mean, minimum, median, maximum value and number of measurements. When calculating the descriptive statistics for the plasma concentration data, values for data points below the lower limit of quantitation will be set to zero.

The pharmacokinetic PK analysis will be performed using a non-compartmental analysis to determine MMC C_{max}, t_{max}, and AUC_{0-6}, where C_{max} is the observed maximum plasma concentration, t_{max} is the time C_{max} occurs, and AUC_{0-6} is the area under the plasma concentration-time profile from time 0 to 6 hours after the first instillation of UGN-102. When available, actual sampling and dose times, rather than scheduled times, will be used for all computations.

10.7 Treatment Exposure

Treatment exposure will be categorized as follows:

- 1 instillation of UGN-102
- 2 instillations of UGN-102
- 3 instillations of UGN-102
- 4 instillations of UGN-102
- 5 instillations of UGN-102
- 6 instillations of UGN-102

Categorical treatment exposure will be summarized using patient counts and percentages. For computing percentages, the denominator will be the number of patients in the ITT analysis set.

Descriptive statistics of UGN-102 volume (mL) and mitomycin C dose (mg) instilled at week 1 through week 6 will be presented for the ITT analysis set. Mitomycin C dose in mg will be calculated as 1.33 mg/mL concentration of mitomycin C times volume of UGN-102 (mL) instilled at week 1 through week 6.

A data listing of treatment exposure will be presented.

10.8 Protocol Deviations

Reported protocol deviations will be categorized to a deviation category, either as major or minor; and mapped to a deviation term. Patients having protocol deviations will be summarized using counts and percentages by deviation category and coded deviation term for the ITT analysis set. Patients will be only counted once within each deviation category and coded deviation term. For computing percentages, the denominator will be the number of patients in the ITT analysis set.

Impact of COVID-19 Situation on Protocol Deviations

Because some patients had disruptions to their visit schedules due to the COVID-19 after completing their 3-MONTH Visit, scheduled visits beyond the 3-MONTH Visit were delayed (outside the protocol window). All protocol deviations due to visits outside the protocol window will be recorded and summarized.

10.9 Delayed or Missed Study Visits Due to COVID-19

Any disruptions to study visits due to COVID 19 occurred after the 3-Month Visit. All missed / delayed study visits following the 3-MONTH Visit will be summarized for the 3-MONTH$_{CR}$ analysis set. Patients with missed study visit following the 3-MONTH Visit will be tabulated using frequency and percentages by study visit. A listing of patients reporting missed study visits due to COVID-19 will be presented.

11 EFFICACY PARAMETERS

11.1 Primary Analysis

Assessment of CR and non-CR at the 3-MONTH Visit for patients who undergo treatment with intravesical UGN-102 (3 months after the first instillation of UGN-102). CR and non-CR will be determined based on the following:

- Cystoscopy
- Biopsy of remaining lesions, if applicable
- Voiding urine cytology

The analysis of the primary endpoint will assess the CR rate defined as the proportion of patients with CR at the 3-MONTH Visit using the ITT analysis set. A patient will be considered to have a CR at the 3-MONTH Visit if CR is recorded on the Evaluation of Response electronic case report form (eCRF). The date of disease assessment associated to CR or non-CR will be determined using the date of cystoscopy, biopsy, or cytology, whichever occurs first. If response cannot be evaluated for a patient at the 3-MONTH Visit, the patient will be considered as non-CR for the purpose of the analysis and will be included in the denominator.

A test for binomial proportions (SAS PROC FREQ with binomial option) will be used to derive the exact two-sided 95% CI for the CR rate and non-CR rate using the Clopper-Pearson method. In addition, reasons for non-CR comprised “Indeterminate” and “Evidence of

Persistence/Worsening of Disease" will be tabulated using the number and percentages of patients.

11.2 Supportive Analysis

Sensitivity analyses for the primary endpoint similar to that described in Section 11.1 will be performed using the mITT analysis set and the PP analysis set. The results will also be displayed by a forest plot.

11.3 Secondary Analysis

11.3.1 Durable Complete Response

Assessment of durable CR (DCR) at 6, 9, and 12 months after the first instillation of UGN-102 for patients who achieved CR at the 3-MONTH Visit will be based on the following:

- Cystoscopy
- Biopsy of remaining lesions, if applicable
- Voiding urine cytology

Two types of analyses will be performed to estimate the visit level durable CR rates.

Analysis 1

Point estimates and two-sided exact 95% binomial CIs [Clopper and Pearson 1934] for durable CR rates (i.e. proportion of patients who maintained CR) at 6, 9, and 12 months after the first instillation of UGN-102 (i.e. 3, 6, and 9 months post CR at the 3-MONTH Visit, respectively) will be presented. The denominator to calculate the proportion will include all patients who were CR at 3-month visit.

Analysis 2

The above analysis will be repeated using denominator with number of patients who reached at a follow-up visit (i.e. 3, 6, and 9 months post CR at the 3-MONTH Visit) with conclusive disease evaluation (i.e. either CR or recurrence). Patients who terminated the study early due to the reasons as per study termination page of the eCRF or patients who have indeterminate response (which are not confirmed by subsequent visits) will be excluded from the denominator.

Illustrating cases of patients who would be excluded from the denominator for computing proportion of patients who maintained CR at months 6, 9, and 12.

Case	Exclusion from the denominator		
	Month 6 Visit	Month 9 Visit	Month 12 Visit
Patient A: discontinued due to AE prior to Month 9 visit	No	Yes	Yes
Patient B: had indeterminate response at Month 12 visit	No	No	Yes
Patient C: had recurrence response at Month 6 visit	No	No	No

11.3.1.1 Impact of COVID-19 Situation on Analysis of Durability

Because some patients had disruptions to their visit schedules due to COVID-19 situation after completing their 3-MONTH Visit, scheduled visits beyond the 3-MONTH Visit were delayed (outside the protocol window) or missed. The CRF has been updated to record such efficacy data in unscheduled forms. In some cases, if the unscheduled visit was closer to the next scheduled visit (still outside the protocol defined window), investigator decided not to bring the patient until the following scheduled visits. For example, if a patient missed 6-month scheduled visit, had a delayed visit at 8 month, investigator may decide not to bring the patient for the 9-month scheduled visit, but to bring the patient in for the 12-month visit.

The following rules will be used to impute missing or indeterminate visit level responses:

Scheduled Visits (in Month)	**Converted Scheduled Visits (in days)**
3 months +/-2 weeks	Day 57 – Day 85
6 months +/-2 weeks	Day 141 – Day 169
9 months +/-2 weeks	Day 225 – Day 253
12 months +/-2 weeks	Day 309 – Day 337

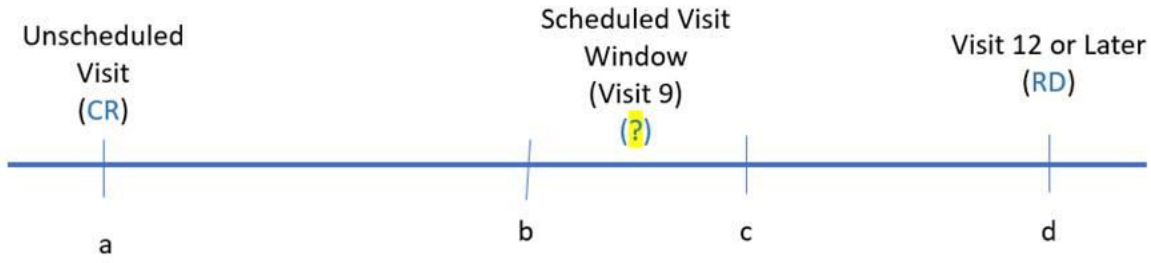
1. If any recorded response is CR, then all previous missing or unknown scheduled visits responses are CR (Patient A).
2. If any recorded response is RD, then all previous missing or unknown scheduled visits responses are RD (Patient B; Patient C) unless there is a previous CR response closer in time (days) to the scheduled visit in which case the visit response will be CR (Patient D). In the case of ties the scheduled visit will be RD (Patient E).
3. Otherwise, responses will not be carried forward to scheduled visits that have not occurred (Patient C; Patient E)

Visit	**Patient A**	**Patient B**	**Patient C**	**Patient D**	**Patient E**
3-MONTH Visit (Days 57- 85)	CR	CR	CR	CR	CR
Month 6 Visit (Days 141 -169)	U/CR	U/ RD	U/RD	CR	CR
Unscheduled Visit (Day 190)					CR
Unscheduled Visit (Day 220)				CR	
Month 9 Visit (Days 225 – 253)	CR	RD	RD	CR	RD
Unscheduled Visit (Day 288)					RD
Unscheduled Visit (Day 300)			RD		
Month 12 Visit (Days 309 – 337)				RD	

RD=Recurrence disease; U=Unknown/Indeterminate; Red implicates imputed response.

In order to further explain 'closer in time (days) to the scheduled visit', the following variables are defined:

a= Study day of the CR unscheduled visit
b=Lower limit of next scheduled visit (days)
c= Upper limit of the above scheduled visit (days)
d= Study day of the following RD visit



If (b-a) < (d-c) then the response at the scheduled visit will be CR; otherwise, RD.

For example, Patient D has a CR response on Study Day 220, five days before the beginning of the Visit 9 window. This patient also has a RD response on study day 309 (say), 56 days (309-253) after the end of the visit 9 window. Since the CR response is closer to the visit 9, the response will be CR.

11.3.2 Duration of Response

Duration of CR (DOR) will be evaluated only for CR patients in the 3-MONTH$_{CR}$ analysis set. DOR is defined as time from the date of evidence of CR at the 3-MONTH Visit to the earliest date of recurrence as determined using the date of cystoscopy, biopsy, or cytology, whichever occurs first. If a patient has not had an event (recurrence), DOR will be censored at the date of the last adequate disease assessment (i.e. assessment where visit level response is CR) or date of death. The rules for defining recurrence and censoring are described in Table 1.

Table 1. Assignments for Event and Censoring Dates for DOR Analysis

Situation	Date of event/censoring	Outcome
Recurrence	Date of visit when first recurrence was documented	Event
Death prior to recurrence	Date of death	Censored
No disease assessment after the first determination of CR at 3-	Date of 3-MONTH assessment visit	Censored
Early discontinuation due to AE or any other reason	Date of the last adequate assessment visit.	Censored
Response evaluation 'indeterminate'	Date of last adequate assessment visit	Censored
Completed study without recurrence or death	Date of last adequate assessment visit	Censored

The distribution of DOR will be estimated using the Kaplan-Meier method. Median times to DOR, first and third quartiles along with 95% CI [Brookmeyer and Crowley 1982] will be estimated, if there are sufficient number of recurrences. A figure and listing of DOR time will also be provided.

A supplementary table showing the details of the number at risk, number of events, and number censored with the event time will also be displayed.

11.3.2.1 Impact of COVID-19 Situation on DOR Analysis

The following sensitivity analyses will be performed to address delayed visits under this situation. If a recurrence is observed after two or more missing disease assessments (i.e. after an extended period), then DOR will be censored at the last adequate assessment visit prior to recurrence. If a recurrence is observed after a single missing disease assessment, the actual date of recurrence will be used. There will be no change to rest of the event/censoring rules as defined in Table 1.

Example (protocol defined schedule of disease assessments is every 3 months): A patient had the disease assessments at 3 months - 6 months - missing - missing - Recurrence. Then the DOR status of this patient will be censored and the censoring date will be the date of 6 month visit (if the patient had an adequate assessment at 6 month) or the date of 3 month visit (if the patient did not have an adequate assessment at 6 month).

A patient will be identified as an extended loss to follow up if the patient did not have an adequate assessment during the time period of 196 days (6 months + 2 weeks) and then had a recurrence.

The above sensitivity analysis will only be performed if there is at least one patient who is an extended lost to follow-up.

11.3.3 Pathological Evaluation

Summary of pathological evaluation comprises number of lesions, visual appearance of the lesions, staging, grading, and urine cytology. Frequency distribution of each classification by visit will be provided.

11.3.3.1 Impact of COVID-19 on Pathological Evaluation

Because some patients had disruptions to their visit schedules due to COVID-19 situation after completing their 3-MONTH Visit, scheduled visits beyond the 3-MONTH Visit were delayed (outside the protocol window) or missed. The summary of pathological evaluation will only be presented at baseline, 3-MONTH, and end of study visit.

11.4 Other Efficacy Analysis

11.4.1 Disease-free Survival (DFS)

Disease-free Survival (DFS) will be evaluated only for CR patients in the 3-MONTH$_{CR}$ analysis set.

DFS is defined as the time from first dose to earliest date of disease recurrence as determined using the date of cystoscopy, biopsy, or cytology, or death from any cause, whichever occurs first. If a patient has not had an event (recurrence or death), DFS will be censored at the date of the last adequate disease assessment (i.e. assessment where visit level response is CR). The rules for defining event and censoring are described in Table 2.

Table 2. Assignments for Event and Censoring Dates for DFS Analysis

Situation	Date of event/censoring	Outcome
Recurrence	Date of visit when first recurrence was documented	Event
Death prior to recurrence	Date of death	Event
No disease assessment after the first determination of CR	Date of 3-MONTH assessment visit	Censored
Early discontinuation due to AE or any other reason	Date of the last adequate assessment visit.	Censored
Response evaluation 'indeterminate'	Date of last adequate assessment visit	Censored
Completed study without recurrence or death	Date of last adequate assessment visit	Censored

The distribution of DFS will be estimated using the Kaplan-Meier method. Median times to DFS, first and third quartiles along with 95% CI [Brookmeyer and Crowley 1982] will be estimated, if there are sufficient number of recurrences or deaths. A figure and listing of DFS time will also be provided.

A supplementary table showing the details of the number at risk, number of events, and number censored with the event time will also be displayed.

11.4.1.1 Impact of COVID-19 Situation on DFS Analysis

A similar sensitivity analyses as of DOR will be performed by censoring extended loss to follow-up patients. For details please see Section 11.3.2.

11.4.2 Time to Recurrence (TTR)

Time to Recurrence (TTR) will be evaluated only for CR patients in the 3-MONTH$_{CR}$ analysis set.

TTR is defined as the time from first dose to earliest date of disease recurrence as determined using the date of cystoscopy, biopsy, or cytology whichever occurs first. If a patient has not had

a recurrence, TTR will be censored at the date of the last adequate disease assessment (i.e. assessment where visit level response is CR). The rules for defining event and censoring are described in Table 3.

Table 3. Assignments for Event and Censoring Dates for TTR Analysis

Situation	Date of event/censoring	Outcome
Recurrence	Date of visit when first recurrence was documented	Event
Death prior to recurrence	Date of death	Censored
Not related death	Date of death	Censored
No disease assessment after the first determination of CR at 3-MONTH Visit	Date of 3-MONTH assessment visit	Censored
Early discontinuation due to AE or any other reason	Date of the last adequate assessment visit.	Censored
Response evaluation 'indeterminate'	Date of last adequate assessment visit	Censored

The distribution of TTR will be estimated using the Kaplan-Meier method. Median times to TTR, first and third quartiles along with 95% CI [Brookmeyer and Crowley 1982] will be estimated, if there are sufficient number of recurrences or related deaths. A figure and listing of DFS time will also be provided.

A supplementary table showing the details of the number at risk, number of events, and number censored with the event time will also be displayed.

11.4.2.1 Impact of COVID-19 Situation on TTR Analysis

A similar sensitivity analyses as of DOR will be performed by censoring extended loss to follow-up patients. For details please see Section 11.3.2.

11.5 Exploratory Analyses

The EORTC QLQ-NMIBC24 is a self-reported 24-item NMIBC specific instrument that assesses 11 domains: 2 functional scales or single item (sexual function, sexual enjoyment), and 9 symptom scales or single items (urinary symptoms, malaise, future worries, bloating and flatulence, male sexual problems, intravesical treatment issues, sexual intimacy, risk of contaminating partner, and female sexual problems). The QLQ-NMIBC24 questionnaire module is presented in Appendix 6.

It will be administered at baseline/Day 1 and at 3-MONTH, 6 Month, 9 Month, and 12 Month or end of study visit.

For the functional scales and symptom scales or single items, patients will be assessed as to how true each of the statements has been for them on a 4-point scale,

1= not at all

2= a little

3= quite a bit

4= very much.

For the functional scale and single item, a higher score reflects a more favorable level of functioning. Conversely, symptom scales or single items a higher score represents low QoL. The scale score will be calculated with only those items for which all of the relevant questions have non-missing answers. For example: The scale score for Urinary Symptoms can only be calculated if Questions 31-37 all were answered.

All scoring information specific to the QLQ-NMIBC24 is presented in Table 4. The scoring method of this questionnaire will consist of a calculation of a raw score, followed by a linear transformation to standardize the raw score so that the scores range from 0 to 100. Raw scores from the EORTC QLQ-NMIBC24 questionnaires will be transformed based on the scoring manual.

Table 4. Scoring 11 Domains of QLQ-NMIBC24

	Scale	Number of Items (n)	Item Range*	Item Numbers ($I_1, I_2,\ldots,I_n$)	Raw Score (RS)
Symptom scales / items					
Urinary Symptoms	US	7	3	31- 37	$(I_{31} +\ldots+ I_{37})/7$
Malaise	MAL	2	3	38, 39	$(I_{38} + I_{39})/2$
Intravesical treatment issues	InV	1	3	40	I_{40}
Future worries	FW	4	3	41 - 44	$(I_{11}+\ldots+I_{44})/4$
Bloating and flatulence	BAF	2	3	45, 46	$(I_{45} + I_{46})/2$
Male sexual problems	SXmen	2	3	49, 50	$(I_{49} + I_{50})/2$
Sexual intimacy [a]	SXI	1	3	51	I_{51}
Risk of contaminating partner [a]	SXCP	1	3	52	I_{52}
Female sexual problems [a]	SXfem	1	3	54	I_{54}
Functional scales / items					
Sexual function	SX	2	3	47, 48	$(I_{47} + I_{48})/2$
Sexual enjoyment [a]	SXEN	1	3	53	I_{53}

* "Item range" is the difference between the possible maximum and the minimum response to individual items. All items are scored 1 to 4, giving range = 3.

[a] Items 51 to 54 are conditional questions and must only be scored if these are applicable to the patient.

Functional scales: Score= {1 - (RS-1)/range} × 100

Symptom scales/items: Score= {(RS-1)/range} × 100

Summary statistics will be provided for 11 domain scores at each assessment time for which there are adequate data. The change from baseline for 11 domain scores will be summarized with summary statistics at each assessment time for which there are adequate data. Analyses will be based on the ITT analysis set. A listing of the 11 domain scores values will be produced for each patient.

11.5.1 Impact of COVID-19 on QLQ Analysis

Scheduled visits during follow up, i.e. after the 3 month visit, were delayed or missed in some patients due to COVID-19. Therefore, at the time of final analysis, summary of QLQ score and change from baseline will be presented at baseline, 3-MONTH, End of Study visit, and worst-case post baseline (includes both scheduled and unscheduled assessments) as time points.

Worst-case score for an individual scale is derived as follows:

- For a symptom scale and single item, the worst-case score is the highest score.
- For a functional scale and single item, the worst-case score is the highest score.

11.6 Interim Analysis

A separate Interim Analysis Plan was created. The purpose of this interim analysis is to inform on the broader program. No decision directly related to this study was made from the results of this interim analysis.

11.7 Subgroup Analyses

The following subgroups will be explored in the analysis of key efficacy data provided subgroups have a minimum of 5 patients in each category, to result in meaningful analyses.

- Age group (<75, ≥75 years)
- BMI category (<30 kg/m^2, ≥30 kg/m^2)
- Sex (male, female)
- Treatment course (full course - 6 instillations , partial course <6 instillations)
- Tumor size (>3 cm, ≤3 cm)
- Tumor (single, multiple)
- Previous LG NMIBC episodes within 1 year of the current diagnosis (zero versus ≥ 1) based on Urothelial Carcinoma Medical History form.
    - In case of partial dates are reported for onset of LG NMIBC episodes, imputed dates based on the imputation rules described in Appendix 3 will be used. Number of days of previous LB NMIBC episode relative to the date of current diagnosis is calculated as
      DAYS = Date of current diagnosis – Date of onset of previous LB NMIBC +1.
      A previous onset of LB NMIBC occurring within a year of the current diagnosis when DAYS is less than 366. Date of current diagnosis is defined as the most recent date of LG NMIBC documented in the Urothelial Carcinoma Medical History form.

The frequency and percentage of CR (primary endpoint) along with the exact 95%confidence interval will be presented for each of the above covariates (subgroups) using the ITT analysis set. The summary will also be displayed by a forest plot.

The frequency and percentage of durable CR (at 12 month or later) will be presented along with the exact 95% confidence interval for each of the above covariates (subgroups) using the 3-MONTH CR analysis set. The summary will also be displayed by a forest plot.

12 SAFETY AND TOLERABILITY

Summaries of AEs, serious adverse events (SAEs), laboratory assessments, vital signs, physical examination, urology-oriented physical examination, and concomitant medications will be presented for the ITT analysis set. ITT and Safety analysis sets are identical for this study.

12.1 Adverse Events

AEs will be recorded from the time a patient has signed the informed consent until end of study. Each verbatim AE term will be coded to a system organ class (SOC) and preferred term (PT) using the most recent version of the MedDRA dictionary version 21.0 or later.

The intensity of an AE will be graded by the investigator according to CTCAE version 5.0 as follows:

- Grade 1: Mild
- Grade 2: Moderate
- Grade 3: Severe or medically significant
- Grade 4: Life-threatening consequences
- Grade 5: Death related to AE (SAE)

12.1.1 General Definitions and Conventions for AEs

Any AEs that occur on or after the day of first instillation of UGN-102 (Visit 1) will be defined as treatment-emergent AEs (TEAEs). If the start date of an AE is incomplete or missing, the event will be assumed to be treatment emergent, unless the incomplete start date (month and/or year) or the stop date (complete or incomplete) clearly indicates that the event started before study Day 1.

AE incidence will be presented as the number and percentage of patients with a specific AE. Percentages will be calculated as the number of patients with the AE divided by the number of patients in ITT analysis set.

Unless otherwise specified, at each level of patient summarization (i.e. by SOC and PT), a patient will be counted only once. If there is more than one occurrence of an event, the event of the worst severity or the most causality category will be summarized.

Summaries of AEs by SOC and PT will be sorted by descending order of frequency of SOC (then alphabetically for ties), then by descending order of frequency of PT within each SOC (then alphabetically for ties). Summaries of AEs by PT will be sorted by descending order of frequency of PT, then alphabetically for ties.

12.1.2 Overall Summary of AEs

An AE overview containing the number of unique patients and percentages will be provided for the following categories:

- Any AEs
- Any serious AEs
- Any TEAEs
- Any study drug or procedure related TEAEs
- Any study drug related TEAEs
- Any procedure related TEAEs
- Any TEAEs leading to treatment discontinuation
- Any TEAEs leading to study discontinuation
- Any study drug or procedure related TEAEs leading to treatment discontinuation
- Any TEAEs by worst severity (mild, moderate, severe or medically significant, life-threatening, or death)
- Any serious TEAEs
- Any study drug or procedure related serious TEAEs
- Any study drug related serious TEAEs
- Any procedure related serious TEAEs
- Any TEAEs leading to death
- Any Treatment-emergent suspected unexpected adverse reactions (SUSAR)
- TEAEs of special interest
- Any Serious TEAEs by worst severity (mild, moderate, severe or medically significant, life-threatening, or death)

12.1.3 Summary of AEs by SOC and PT

The following will be summarized by SOC and PT:

- Incidence of TEAEs
- Incidence of study drug or procedure related TEAEs
- Incidence of study drug related TEAEs
- Incidence of procedure related to TEAEs
- Incidence of TEAEs leading to treatment discontinuation
- Incidence of TEAEs leading to study discontinuation.
- Incidence of study drug or procedure related to TEAEs leading to treatment discontinuation
- Incidence of TEAEs by worst severity
- Incidence of TEAEs by worst outcome
- Incidence of Serious TEAEs
- Incidence of study drug or procedure related serious TEAEs
- Incidence of study drug related serious TEAEs
- Incidence of procedure related serious TEAEs
- Incidence of TEAEs leading to death
- Incidence of treatment-emergent suspected unexpected adverse reactions (SUSAR)
- Incidence of serious TEAEs by worst severity
- Incidence of TEAEs of special interest
- Incidence of TEAEs of special interest by worst outcome

Data listings of all AEs (including TEAEs and non-TEAEs), all serious AEs (including serious TEAEs and serious non-TEAEs), TEAE leading to treatment discontinuation, and TEAEs leading to death will be provided.

12.1.4 Summary of AEs by PT

The following will be summarized by PT

- Incidence of TEAEs
- Incidence of study drug or procedure related TEAEs
- Incidence of TEAEs leading to treatment discontinuation
- Incidence of TEAEs by worst severity
- Incidence of TEAEs by worst outcome
- Incidence of Serious TEAEs
- Incidence of study drug or procedure related serious TEAEs
- Incidence of serious TEAEs by worst severity
- Incidence of TEAEs of special interest
- Incidence of TEAEs of special interest by worst outcome

12.1.5 Summary of AEs by SOC, PT, and Age

The following will be summarized by SOC, PT, and age (<65 vs >=65):

- Incidence of TEAEs
- Incidence of serious TEAEs

12.1.6 Summary of AEs by SOC, PT, and Sex

The following will be summarized by SOC, PT, and sex:

- Incidence of TEAEs
- Incidence of Serious TEAEs

In addition, all deaths occurring any time from the time of informed consent to the clinical cut-off date will be summarized based on the number and percentage of patients. This summary will classify patients by time of death relative to the last instillation of medication (> 30 days or ≤ 30 days) and primary cause of death.

12.1.7 Adverse Event of Special Interest

A comprehensive list of reported AE terms based on clinical review will be used to identify MedDRA PT terms for each of the AESI categories listed below. For some events several AE preferred terms may be ‘collapsed’. The list of terms to be used for each category of adverse events of interest may be based on the Safety Review Team (SRT) agreements in place at the time of reporting. The SRT agreements are based on a review of the MedDRA dictionary.

The following categories of AE of special interest will be presented in table summary as permitted by data:

- Overdose of an IP

- Suspected abuse/misuse of IP
- Inadvertent or accidental exposure to IP
- Medication error involving IP (with or without patient/patient exposure to the Sponsor medicinal product
- Identified significant safety issue either from an individual case report or review of aggregate data
- Allergic reaction to MMC
- Voiding interruption due to urethral/penile edema (unrelated to prostatic hypertrophy)
- Indication of bone marrow suppression
- Lower urinary tract symptoms

Additional AE of special interests may be identified through manual review of all AEs by PT prior to database lock.

Summaries of the number and percentage of patients with these events will be provided by category and PT; and SOC and PT. The summary of event characteristics will also be provided, including number of patients with any event, number of events, number of patients with any event that is serious, number of patients with any event that is related to study treatment, the worst outcome of the event, maximum grade and the action taken for the event. The worst case approach will be applied at patient level for the event outcome, maximum grade and the action taken, i.e. a patient will only be counted once as the worst case from all the events that patient had. In addition, onset and duration of the first occurrences for each type of events will be summarized if data permits.

Other safety analyses may be performed if deemed necessary during the data review.

12.2 Laboratory Assessments

Laboratory parameters of chemistry, hematology, and urinalysis will be collected at screening, baseline, Day 8, Day 15, Day 22, Day 29, Day 36, 3-MONTH Visit, 6 Months, 9 Months, and at 12 Months. Laboratory results of chemistry and hematology will be graded according to laboratory CTCAE version 5.0 for each laboratory parameter defined in Appendix 5. Laboratory CTCAE grading has 5 levels (grade 0 to grade 4).

Summaries of laboratory parameters will include chemistry and hematology using conventional units.

A complete listing of chemistry, hematology, and urinalysis parameters is described in Table 5.

Table 5. Laboratory Parameters

Chemistry		Hematology	Urinalysis
Liver Function Tests:	**Kidney Function Tests:**	Red blood cell count	Specific gravity
Aspartate aminotransferase	Creatinine	Hemoglobin	pH
Alanine aminotransferase	Blood urea nitrogen	Hematocrit	Bilirubin
Alkaline phosphatase	Uric acid	WBC count	Glucose
Total bilirubin	Sodium	Neutrophils	Urobilinogen
Direct bilirubin	Potassium	Neutrophils absolute	Ketones
Albumin	Phosphorus	Lymphocytes	Blood
Total protein	Calcium	Lymphocytes absolute	Protein
	Bicarbonate	Monocytes	Nitrites
	Chloride	Monocytes absolute	Leukocyte
	Estimated glomerular filtration rate (calculated in electronic data capture)	Eosinophils	Bacteriuria
		Eosinophils absolute	
		Basophils	
		Basophils absolute	
		Platelet count	
		Mean corpuscular volume	
		Mean corpuscular hemoglobin concentration	

Patients having laboratory results that are outside the laboratory normal ranges, a complete laboratory test profile along with CTCAE grades and classification relative to the laboratory normal ranges for those patients will provided in a listing.

12.2.1 Impact of COVID-19 on Laboratory Assessment

Due to some missing/delayed scheduled visits beyond the 3-MONTH Visit, at the time of final analysis, the summary of laboratory parameters will be presented at baseline, 3-MONTH, end of study, worst-case treatment-period, and worst-case post-baseline (includes both scheduled and unscheduled assessments) as time points. Summaries of worst-case of observed lab values for a planned interval time point will be presented based on worst minimum and worst maximum lab values.

Summaries of worst-case grade increase from baseline grade will be provided for all the lab tests that are gradable by CTCAE version 5.0. These summaries will display the number and percentage of subjects with a maximum post baseline grade increasing from their baseline grade. Any increase in grade from baseline will be summarized along with any increase to a maximum grade of 3 and any increase to a maximum grade of 4. Missing baseline grade will be assumed as grade 0. For laboratory tests that are graded for both low and high values, summaries will be done separately and labeled by direction, e.g., sodium will be summarized as hyponatremia and hypernatremia.

For lab tests that are not gradable by CTCAE version 5.0, summaries of worst-case changes from baseline with respect to normal range will be generated. Decreases to low, changes to normal or

no changes from baseline, and increases to high will be summarized at each planned time. If a subject has a decrease to low and an increase to high during the same time interval, then the subject is counted in both the "Decrease to Low" categories and the "Increase to High" categories.

12.2.2 Potentially Clinically Significant Laboratory Values

Patients with PCS laboratory values will be assessed via the incidence of patients meeting the PCS criteria by time point. For computing percentages, the denominator will be the number of patients with a post-baseline value for the specific laboratory parameter, except for hemoglobin that requires patients with a baseline and a post-baseline value, and the respective time point. The PCS laboratory criteria are provided in Table 6.

Table 6. PCS Laboratory Criteria

Laboratory Parameter	Conventional Unit	Lower Limit	Upper Limit
Chemistry			
Creatinine	mg/dL		>2.2
Sodium	mEq/L	≤130	>150
Potassium	mEq/L	<3.0	>5.5
Total bilirubin	mg/dL		>1.5 × ULN
ALT	U/L		>3 × ULN
AST	U/L		>3 × ULN
GGT	U/L		>2.5 × ULN
Hematology			
Hemoglobin	g/dL	<0.8 × LLN and >20% decrease from baseline	>1.3 × ULN and >30% increase from baseline
Leukocytes	$\times 10^3/\mu L$	≤ 2.8	≥ 16.0
Lymphocytes	$\times 10^3/\mu L$	<0.5	>20
Neutrophils	$\times 10^3/\mu L$	<1.0	
Platelets	$\times 10^3/\mu L$	<75	≥700

ALT = alanine aminotransferase; AST = aspartate aminotransferase; GGT = gamma-glutamyltransferase; PCS = potentially clinically significant; LLN=Lower limit of normal; ULN=Upper limit of normal

12.2.3 Abnormal Liver Function Tests

The number and percentage of patients meeting criteria for abnormal liver function tests (LFTs) by time point will be summarized and listed by patient. For computing percentages, the denominator will be the number of patients with a post-baseline value for the specific laboratory

parameter and the respective time point. Threshold values of interest for LFTs are provided in Table 7.

Table 7. Liver Function Tests Criteria

Parameter	Criterion
ALT	>3×ULN; >5×ULN; >10×ULN
AST	>3×ULN; >5×ULN; >10×ULN
ALT and TBL	ALT 3×ULN and TBL >2×ULN
AST and TBL	AST 3×ULN and TBL >2×ULN
ALT = alanine aminotransferase; AST = aspartate aminotransferase; TBL = total bilirubin; ULN = upper limit of normal	

For a combined criterion to be fulfilled, all conditions have to be fulfilled by the same laboratory measurement. Only patients with newly occurring values (at least one post-baseline measurement and meeting the criterion but not meeting the criterion at baseline) will be counted.

12.3 Vital Signs

12.3.1 Measurements of Vital Signs

Systolic blood pressure (SBP), diastolic blood pressure (DBP), pulse rate, respiration rate, temperature, and weight will be measured at baseline, Day 8, Day 15, Day 22, Day 29, Day 36, 3-MONTH, 6 Months, 9 Months, and 12 Months/End of Study.

12.3.2 Impact of COVID-19 on Vital Signs

Due to some missing/delayed scheduled visits beyond the 3-MONTH Visit, at the time of final analysis, the summary of vital signs will be presented at baseline, 3-MONTH, end of study, worst-case treatment-period, and worst-case post-baseline (includes both scheduled and unscheduled assessments) as time points. Summaries of worst-case of observed vital sign measurements for a planned interval time point will be presented based on worst minimum and worst maximum vital sign measurements.

12.3.3 Potentially Clinically Significant Vital Sign Abnormalities

PCS criteria will be used to assess PCS vital sign abnormalities. The incidence of patients meeting PCS criteria for pulse rate, SBP, and DBP at 3-MONTH, end of study, worst-case treatment-period, and worst-case post-baseline will be provided. For computing percentages, the denominator will be the number of patients with a post-baseline value for the specific vital signs parameter and the respective time point. The PCS criteria are shown in Table 8.

Table 8. Vital Signs PCS Criteria

Parameter	PCS Criterion
Pulse rate	≤50 bpm ≤50 bpm and decrease of ≥15 bpm from Baseline ≥120 bpm ≥120 bpm and increase of ≥15 bpm from Baseline
Systolic blood pressure	≤90 mmHg ≤90 mmHg and decrease of ≥20 mmHg from Baseline ≥180 mmHg ≥180 mmHg and increase of ≥20 mmHg from Baseline
Diastolic blood pressure	≤50 mmHg ≤50 mmHg and decrease of ≥15 mmHg from Baseline ≥105 mmHg ≥105 mmHg and increase of ≥15 mmHg from Baseline

12.3.4 Physical Examination

Physical examination findings will be summarized at Baseline (Screening Visit) and at 12 Months/End of Study. The body system of physical examination to be summarized are listed in Table 9. Each body system is assessed with the following examination findings: 'normal', 'abnormal – not clinically significant', or 'abnormal – clinically significant'.

Shift tables for each body system, showing changes of examination findings from baseline to post-baseline. For a specific body system, patients with both baseline and a post-baseline value will be included for shift summaries. For computing percentages, the denominator will be the number of patients with a baseline and a post-baseline value for the specific body system.

Table 9. Physical Examination

Body System
General appearance
Cardiovascular system
Respiratory system
Head, eyes, ears, nose, throat, and neck
Abdomen
Extremities
Neurologic system
Skin

12.4 Urology-Oriented Physical Examination

Any clinically relevant changes occurring from Day 1 until the last trial visit will be recorded and summarized as part of adverse event data. A listing of patients with abnormal urology-oriented physical examination findings either clinically significant or not clinically significant will be provided.

12.5 MMC Levels (Pharmacokinetics) Profiling

The Pharmacokinetic (PK) profile of MMC during the 6 hours following the first instillation of UGN-102 will be based on PK sampling at 0 (pre-instillation), 0.5, 1, 2, 3, 4, 5, and 6 hours post instillation. PK analysis will be performed using the PK analysis set. Plasma concentration data, as well as PK parameters (C_{max}, t_{max}, and AUC_{0-6}) of MMC, will be provided using descriptive statistics. The PK parameters obtained, specifically C_{max}, will be compared to the threshold of MMC plasma concentration (400 ng/mL) known to be associated with myelosuppression, as well as to the known PK profile of MMC obtained with different routes/methods of administration (e.g., intravesical instillation with water for injection).

Additionally, these data, combined with data emerging from other trials, may be used in a population PK model development and analysis that will be described in a separate analysis plan.

13 REFERENCES

BCG-Unresponsive Nonmuscle Invasive Bladder Cancer: Developing Drugs and Biologics for Treatment – Guidance for Industry
https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM529600.pdf

Clopper CJ and Pearson ES (1934). The use of confidence or fiducial limits illustrated in the case of the binomial. Biometrical, 26, 404-413.

Brookmeyer R and Crowley J (1982). A Confidence Interval for the Median Survival Time. Biometrics, 38, 29 - 41.

14 APPENDICES

Appendix 1 Schedule of Activities: Screening to 3-MONTH Visit

Procedures	Screening Day -14 to Day -1 or Day -28 to Day -1§	Enrollment and Baseline Study Visit 1 Day 1	Study Visit 2 Day 8 -1/ +3 days*	Study Visit 3 Day 15 -1/ +3 days*	Study Visit 4 Day 22 -1/ +3 days*	Study Visit 5 Day 29 -1/ +3 days*	Study Visit 6 Day 36 -1/ +3 days*	3-MONTH Study Visit 7 5 weeks ±1 week after last instillation (3 months post initiation)†	Unscheduled Visit^
Informed consent	X								
Inclusion/exclusion criteria	X								
Demographics	X								
Medical history	X								
Administer study medication		X	X	X	X	X	X		
Concomitant medication review	X	X	X	X	X	X	X	X	X
Directed physical examination‡	X	X			X			X	
Vital signs	X	X	X	X	X	X	X	X	X
Height	X								
Weight	X	X		X		X		X	
Hematology	X	X	X	X	X	X	X	X	X
Serum chemistry	X	X	X	X	X	X	X	X	X
Urinalysis#	X	X	X	X	X	X	X	X	X
Pregnancy test	X							X	
Administer QlQ – NMIBC24		X						X	X
PK (in a subset of 6 patients)		X							
Adverse event review and evaluation	X	X	X	X	X	X	X	X	X
Cystoscopy/urine cytology***	X							X***	X
Cold cup biopsy****	X								
CT Urogram*****	X								
Complete eCRFs	X	X	X	X	X	X	X	X	X

3-MONTH = Primary endpoint assessment; CT = computerized tomography; eCRF = electronic Case Report Form; Non-CR = nonresponder.
§ The screening period is up to 14 days for patients that do not need a screening biopsy and up to 28 days for patients that need a screening biopsys (Refer to Section 1.0.1.1 of the protocol).
* Windows are provided to accommodate patient logistics in scheduling. Instillations should not occur more frequently than 6 days apart.

† Patients who discontinue before the 3-MONTH Visit or who are discontinued as Non-CR at the MONTH 3 Visit should have all assessments specified for the 3-MONTH Visit as their end-of study assessments. Patients discontinued as Non-CR will be managed according to best practice according to their treating physician (Refer to Section 10.2.2 of the protocol).
^ If an unscheduled visit is required, assessments should be performed as appropriate to the needs of the visit, although safety assessments should be performed, if feasible.
‡ Full physical examination at Screening; urology-oriented physical examination at Visits 1, 4, and 7 (Refer to Section 10.5.4: Table 5 of the protocol).
#To include culture and sensitivity at Screening.
** No cytology needed if results from a prior cytology within 6 weeks of screening are available)
*** Biopsies may be needed at 3 months if any remaining lesions are visualized during cystoscopy (Refer to Sections10.2.1, 10.2.2)
**** Cold cup biopsy performed only if not already performed within 6 weeks of screening
***** Acceptable if performed within 6 months prior to Screening

Appendix 2 Schedule of Activities: Complete Responder at 3-MONTH Visit: Visit 8 to End of Study

Procedures	Study Visit 8 4 months post initiation ±1 week	Study Visit 9 5 months post initiation ±1 week	Study Visit 10 6 months post initiation ±2 weeks	Study Visit 11 7 months post initiation ±1 week	Study Visit 12 8 months post initiation ±1 week	Study Visit 13 9 months post initiation ±2 weeks	Study Visit 14 10 months post initiation ±1 week	Study Visit 15 11 months post initiation ±1 week	END OF STUDY Study Visit 16 12 months post initiation ±2 weeks	Unscheduled Visit^
Concomitant medication review	X	X	X	X	X	X	X	X	X	X
Full physical examination									X	
Vital signs			X			X			X	X
Weight						X				
Hematology			X			X			X	X
Serum chemistry			X			X			X	X
Urinalysis			X			X			X	X
Administer QLQ - NMIBC24			X			X			X	X
Adverse event review and evaluation	X	X	X	X	X	X	X	X	X	X
Cystoscopy/urine cytology[§]			X			X			X	X
Complete eCRFs	X	X	X	X	X	X	X	X	X	X

CR = Complete Response; eCRF = electronic Case Report Form.

^ If an unscheduled visit is required, assessments should be performed as appropriate to the needs of the visit, although safety assessments should be performed, if feasible.

[§] Any patient with evidence of disease recurrence post-CR will be documented as such, discontinued from the study, and should have all assessments specified for the END OF STUDY Visit as their end-of-study assessments. Patients discontinued with disease recurrence will be managed according to best practice according to their treating physician (Refer to Section 10.2.4 of the protocol).

Appendix 3 Imputation Rules for Missing or Partial Dates for AEs and Medical History

Date	Situation	Imputation Rule
AE Start Date	Only month and year are known and month and year are prior to first dose date	Use the first day of the month
	Only month and year are known and month and year are the same as first dose date	Use the first dose date
	Only month and year are known and month and year are after first dose date	Use the first day of the month
	Only year is known and year is before first dose date	Use Jan 1 of that year
	Only year is known and year is after first dose date	Use Jan 1 of that year
AE End Date	Only month and year are known and month and year are prior to last dose date	Use the last day of the month
	Only month and year are known and month and year are the same as last dose date	Use the last dose date
	Only month and year are known and month and year are after last dose date	Use the last day of the month
	Only year is known and year is before last dose date	Use Dec 31 of that year
	The estimated stop date is before a complete or imputed AE start date	Use the last day of the month of the AE start date

AE = adverse event

Note: The imputation of end date must be later than start date.

Appendix 4 Imputation Rules for Missing or Partial Dates for Prior and Concomitant Medications and Surgical Procedures

Imputation rules for missing or partial dates (D=day, M=month, Y=year, T=time)			
Parameter	**Missing**	**Additional conditions**	**Imputation Rule**
Start date	D only	M and Y same as M and Y of first dose of study drug	Date of first dose of study drug
		M and/or Y not same as M and Y of first dose of study drug	First day of month
	M and D	Y same as Y of first dose of study drug	Date of first dose of study drug
		Y not same as Y of first dose of study drug	Use Jan 01 of Y
	M, D and Y	None--date completely missing	Date of first dose of study drug
Stop date	D only	M and Y same as M and Y of last dose of study drug	Date of last dose of study drug
		M and/or Y not same as M and Y of last dose of study drug	Last day of month
	M and D	Y same as Y of last dose of study drug	Date of last dose of study drug
		Y not same as Y of last dose of study drug	Use Dec 31 of Y
	M, D and Y	None--date completely missing	Date of last dose of study drug

Appendix 5 Laboratory CTCAE Grade Version 5.0 Criteria

Lab Parameter	Conventional Unit	CTCAE Grade v5.0				
		Grade 0	Grade 1	Grade 2	Grade 3	Grade 4
Hemoglobin decreased (Anemia)	g/dL	≥LLN	10 – <LLN	8 – <10	< 8	
Hemoglobin increased	g/dL	≤ULN	>ULN – 2 + ULN	>2 + ULN – 4 + ULN	>4 + ULN	
Hypoglycemia (Glucose decreased)	mg/dL	≥LLN	55 – <LLN	40 – <55	30 – <40	<30
Glucose (hyperglycemia)	mg/dL	LLN – ULN	>ULN – 160	>160 – 250	>250 –500	>500
Albumin	g/dL	≥LLN	<LLN – 3	<3 – 2	<2	Life-threatening consequences; urgent intervention indicated
Alkaline phosphatase		≤ULN	>ULN – 2.5 × ULN	>2.5 – 5.0 × ULN	>5.0 – 20.0 × ULN	>20.0 × ULN
Alanine aminotransferase increased	U/L	≤ULN	>ULN – 3 × ULN	>3 × ULN – 5 × ULN	>5 × ULN – 20 × ULN	> 20 × ULN
Aspartate aminotransferase increased	U/L	≤ULN	>ULN – 3 × ULN	>3 × ULN – 5 × ULN	>5 × ULN – 20 × ULN	> 20 × ULN
Blood bilirubin increased	mg/dL	≤ULN	>ULN – 1.5 × ULN	>1.5 × ULN – 3 × ULN	>3 × ULN – 10 × ULN	>10 × ULN
Creatinine increased	mg/dL	≤ULN	>ULN – 1.5 × ULN	>1.5 × ULN – 3 × ULN	>3 × ULN – 6 × ULN	>6 × ULN
Calcium (hypocalcemia)	mg/dL	LLN – ULN	<LLN – 8.0	<8.0 – 7.0	<7.0 – 6.0	<6.0
Calcium (hypercalcemia)	mg/dL	LLN – ULN	>ULN – 11.5	>11.5 – 12.5	>12.5 – 13.5	>13.5
Gamma-glutamyl transferase increased	U/L	≤ULN	>ULN – 2.5 × ULN	>2.5 × ULN – 5 × ULN	>5 × ULN – 20 × ULN	> 20 × ULN
Eosinophils increased	$\times 10^3/\mu L$	≤ULN	>ULN and >Baseline			
Lymphocyte count decreased	$\times 10^3/\mu L$	≥LLN	0.8 – < LLN	0.5 – <0.8	0.2 – <0.5	<0.2
Lymphocyte count increased	$\times 10^3/\mu L$	≤4		>4 - 20	>20	

Neutrophil count decreased	$\times 10^3/\mu L$	≥LLN	1.5 – <LLN	1.0 – <1.5	0.5 – <1.0	<0.5
Platelet count decreased	$\times 10^3/\mu L$	≥LLN	75 – <LLN	50 – <75	25 – <50	<25
White blood cell decreased	$\times 10^3/\mu L$	≥LLN	3.0 – <LLN	2.0 – <3.0	1.0 – <2.0	<1.0
White blood cell increased (eukocytosis)	$\times 10^3/\mu L$	≤100			>100	
Hyperkalemia (Potassium increased)	mEq/L	≤ULN	>ULN – 5.5	>5.5 – 6.0	>6.0 – 7.0	>7.0
Hypokalemia (Potassium decreased)	mEq/L	≥LLN	3.0 – <LLN	Symptomatic with 3.0 – <LLN	2.5 – <3.0	<2.5
Hypernatremia (Sodium increased)	mEq/L	≤ULN	>ULN – 150	>150 – 155	>155 – 160	>160
Hyponatremia (Sodium decreased)	mEq/L	≥LLN	130 – <LLN	125 – <130	120 – <125	<120
Notes: ULN = upper limit of normal; LLN = lower limit of normal						

Appendix 6 QLQ – NMIBC24

		Not at all	A little	Quite a bit	Very much
31.	Have you had to urinate frequently **during the day**?	1	2	3	4
32.	Have you had to urinate frequently **at night**?	1	2	3	4
33.	When you felt the urge to pass urine, did you have to hurry to get to the toilet?	1	2	3	4
34.	Was it difficult for you to get enough sleep, because you needed to get up frequently at night to urinate?	1	2	3	4
35.	Have you had difficulty going out of the house, because you needed to be close to a toilet?	1	2	3	4
36.	Have you had any unintentional release (leakage) of urine?	1	2	3	4
37.	Have you had pain or a burning feeling when urinating?	1	2	3	4
38.	Did you have a fever?	1	2	3	4
39.	Did you feel ill or unwell?	1	2	3	4
40.	Did you have trouble arranging your life around the repeated bladder treatment appointments (cystoscopies or instillations)?	1	2	3	4
41.	Did you worry about having repeated bladder treatments (cystoscopies or instillations)?	1	2	3	4
42.	Were you worried about your health in the future?	1	2	3	4
43.	Did you worry about the results of examinations and tests?	1	2	3	4
44.	Did you worry about possible future treatments?	1	2	3	4
45.	Did you have a bloated feeling in your abdomen?	1	2	3	4

46.	Have you had flatulence or gas?	1	2	3	4
47.	To what extent were you interested in sex?	1	2	3	4
48.	To what extent were you sexually active (with or without sexual intercourse)?	1	2	3	4
49.	**For men only:** Did you have difficulty gaining or maintaining an erection?	1	2	3	4
50.	**For men only:** Did you have ejaculation problems (e.g. dry ejaculation)?	1	2	3	4
51.	Have you felt uncomfortable about being sexually intimate?	1	2	3	4
52.	Have you worried that you may contaminate your partner during sexual contact with the bladder treatment you have been receiving?	1	2	3	4
53.	To what extent was sex enjoyable for you?	1	2	3	4
54.	**For Women only:** did you have a dry vagina or other problems during intercourse?	1	2	3	4