Statistical Analysis Plan: BL011

Study Title:	A Phase 3, Single-Arm, Multicenter Study 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 (IR) of Recurrence
Study Number:	BL011
Study Phase:	3
Sponsor:	UroGen Pharma Ltd.
Version	2.0
Date:	07 Mar 2024

Confidentiality Statement

SIGNATURE PAGE

PREPARED BY:

Author

Thalis Tsaklas, MSc

Biostatistical Consultant

UroGen Pharma, Ltd.

DocuSigned by:
Thalis Tsaklas
Signer Name: Thalis Tsaklas
Signing Reason: I am the author of this document
Signing Time: 3/7/2024 | 3:35:38 PM EST
D56E419501B74260AAC2AD5251E2F487

Signature

Date: 3/7/2024

APPROVED BY:

Approver

Brent Burger, MS

Senior Director, Biostatistics

UroGen Pharma, Ltd.

DocuSigned by:
Brent Burger
Signer Name: Brent Burger
Signing Reason: I approve this document
Signing Time: 3/8/2024 | 12:18:24 PM EST
A1BA85019ADE45BCBB22E30B76E97A2A

Signature

Date: 3/8/2024

Approver

Michael Louie, MD MPH MSc

SVP, Medical Affairs & Clinical Development

UroGen Pharma, Ltd.

DocuSigned by:
Michael Louie
Signer Name: Michael Louie
Signing Reason: I approve this document
Signing Time: 3/8/2024 | 10:41:58 AM PST
FC2AA91B731C42D189799850AF1C436D

Signature

Date: 3/8/2024

TABLE OF CONTENTS

PAGE

Version History

SAP Version	Date	Rationale
1.0	22 June 2023	Original version for the primary endpoint analysis
2.0	07 Mar 2024	Added Section 4.2 for timing of analyses to describe the planned analyses. Aligned primary analysis of primary endpoint with Protocol. The primary analysis defined in SAP V1.0 is described as sensitivity analysis in SAP V2.0 with additional updates in the imputation method, so that the imputation does not depend on whether the next disease assessment was prior or after the 3-Month analysis database frozen date. Updated the list of pre-specified subgroups for subgroup analyses based on clinically meaningful subgroups. Added DCR Analysis 1 to reflect the worst-case scenario for the calculation of CRR by Visit. Analysis 1 and Analysis 2 in SAP V1.0 are described as Analysis 2 and Analysis 3 respectively in SAP V2.0. Updated description of analyses and listed examples for clarity. Clarified that any missing/indeterminate disease assessments during follow-up are imputed to CR using LOCF, only if the previous conclusive disease assessment was from an Unscheduled Visit and closer in time to the missed/indeterminate disease assessment. Added analyses of key secondary, secondary, exploratory endpoints, safety assessments (clinical laboratory, vital signs). Aligned safety observation periods with updated Protocol.

ABBREVIATIONS

ADaM	Analysis Data Model
AE	adverse event
AESI	adverse event of special interest
ALT	alanine aminotransferase
ASA	American Society of Anesthesiologists
AST	aspartate aminotransferase
ATC	Anatomical Therapeutic Chemical (classification)
BMI	body mass index
CI	confidence interval
CDISC	Clinical Data Interchange Standards Consortium
COVID-19	coronavirus disease 2019
CR	complete response
CRR	complete response rate
CTCAE	Common Terminology Criteria for Adverse Events
DBL	Database lock
DBP	diastolic blood pressure
DFS	disease-free survival
DCR	durable complete response
DOR	duration of response
eCRF	electronic case report form
EOR	evaluation of response
EORTC	European Organisation for Research and Treatment of Cancer
EOS	end of study
FDA	Food and Drug Administration
HR	hazard ratio
ICH	International Council for Harmonisation
IP	investigational product
IR	intermediate risk

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
MMC	Mitomycin C
NCR	non-complete response
NPT	non-protocol therapy
NMIBC	non-muscle invasive bladder cancer
PCS	potentially clinically significant
PP	per protocol
PRO	patient reported outcome
PT	preferred term
QoL	quality of life
QLQ-C30	30-item quality of life questionnaire for cancer patients
QLQ-NMIBC24	24-item quality of life questionnaire for patients with NMIBC
RD	recurrence disease
RS	raw score
SAE	serious adverse event
SAP	statistical analysis plan
SAS	Statistical Analysis System
SBP	systolic blood pressure
SI	international system of units
SoA	schedule of activities
SoC	standard of care
SOC	system organ class
TBL	total bilirubin
TEAE	treatment-emergent adverse event

TLFs	tables, listings, figures
TTR	time to recurrence
TURBT	transurethral resection of bladder tumors
UCMH	urothelial carcinoma related medical history
ULN	upper limit of the normal range
WHO	World Health Organization

1. INTRODUCTION

The purpose of this statistical analysis plan (SAP) is to define all prespecified analyses of study BL011 (ENVISION), a phase 3, single-arm, multicenter study to evaluate the efficacy and safety of UGN-102 as primary chemoablative therapy in patients with recurrent low grade (LG) non-muscle-invasive bladder cancer (NMIBC) at intermediate risk (IR) of recurrence.

This SAP is based on the following study documents:

- Protocol, Version 5.0 dated 28 February 2024
- Case report form (CRF), Version 7.4 dated 22 November 2023

2. STUDY OBJECTIVES AND ENDPOINTS

Table 1 lists the study objectives and corresponding endpoints of the study.

Table 1: Study Objectives and Endpoints

OBJECTIVE	ENDPOINT
Primary	
To evaluate the tumor ablative effect of UGN-102 in patients with LG NMIBC	CRR (complete response rate), defined as the proportion of patients who achieved CR (complete response) at the 3-month Visit (3 months after the first instillation of UGN-102) as determined by cystoscopy, for cause biopsy, and urine cytology
Key Secondary	
To evaluate the durability of response with respect to DOR (duration of response)	DOR in patients who achieved CR at the 3-month Visit, defined as the time from the date of evidence of CR at the 3-month Visit to the earliest date of recurrence or progression as determined using the date of cystoscopy, for cause biopsy, or cytology, or death due to any cause, whichever occurred first
Other Secondary	
To evaluate the durability of response with respect to DCR (durable complete response) rate at scheduled disease assessment time points and DFS (disease-free survival)	• DCR rate at scheduled disease assessment time points, defined as the proportion of patients who achieved CR at the 3-month Visit and maintained CR (i.e., no detectable disease) up to that particular follow-up disease assessment • DFS in patients who achieved CR at the 3-month Visit, defined as the time from first dose to the earliest date of recurrence or progression as determined using the date of cystoscopy, for cause biopsy, or cytology, or death due to any cause, whichever occurred first

OBJECTIVE	ENDPOINT
To evaluate the safety and tolerability of intravesical instillations of UGN-102 in patients with LG NMIBC	The safety profile of UGN-102 will be evaluated through the reporting of AEs (adverse events), including SAEs (serious adverse events) and AESIs (adverse events of special interest), and through standard clinical and laboratory tests (e.g., hematology and chemistry, urinalysis), physical examination, and vital signs
Exploratory	
To evaluate TTR (time to recurrence) following SoC treatment in patients who have NCR (non-complete response) at the 3-month Visit	TTR in patients who were NCR (residual disease) at the 3-month Visit, defined as the time from the date of the first treatment after NCR to the earliest date of recurrence or progression as determined using the date of cystoscopy, for cause biopsy, or cytology, whichever occurred first
To evaluate the preliminary responses following SoC treatment in 3-month CR patients who have disease recurrence or progression during the Follow-up Period	Number (%) of response outcomes evaluated at the first disease assessment visit after SoC
To evaluate the impact of UGN-102 on subsequent SoC treatment in patients who have NCR at the 3-month Visit or disease recurrence or progression during the Follow-up Period	Proportion of patients whose planned NMIBC treatment at baseline was downgraded following treatment with UGN-102 (e.g., from TURBT to biopsy and/or fulguration)
To assess the effect of UGN-102 on PRO (patient reported outcome) measures including disease related symptoms and physical, mental, and social health	Changes from baseline in patient scores on QLQ-C30 (30-item quality of life questionnaire for cancer patients) and QLQ-NMIBC24 (24-item quality of life questionnaire for patients with NMIBC) questionnaires

3. STUDY DESIGN

This Phase 3, multinational, single-arm, multicenter study is designed to evaluate the efficacy and safety of UGN-102 as primary chemoablative therapy in patients with recurrent LG IR NMIBC.

Eligible patients are scheduled to receive 6 once-weekly intravesical instillations of UGN-102.

All patients are scheduled to return to the clinic approximately 3 months after the first

instillation for determination of response to UGN-102 treatment. Assessment of response is based on visual observation (white light cystoscopy), histopathology of any remaining or new lesions by central pathology lab (if applicable), and interpretation of urine cytology by central pathology lab. Any lesions or suspect tissue are required by protocol to be biopsied to evaluate for persistence of disease. Details of the evaluation methods are available in Appendix 1 (Guidance on Evaluation of Response) of the Protocol.

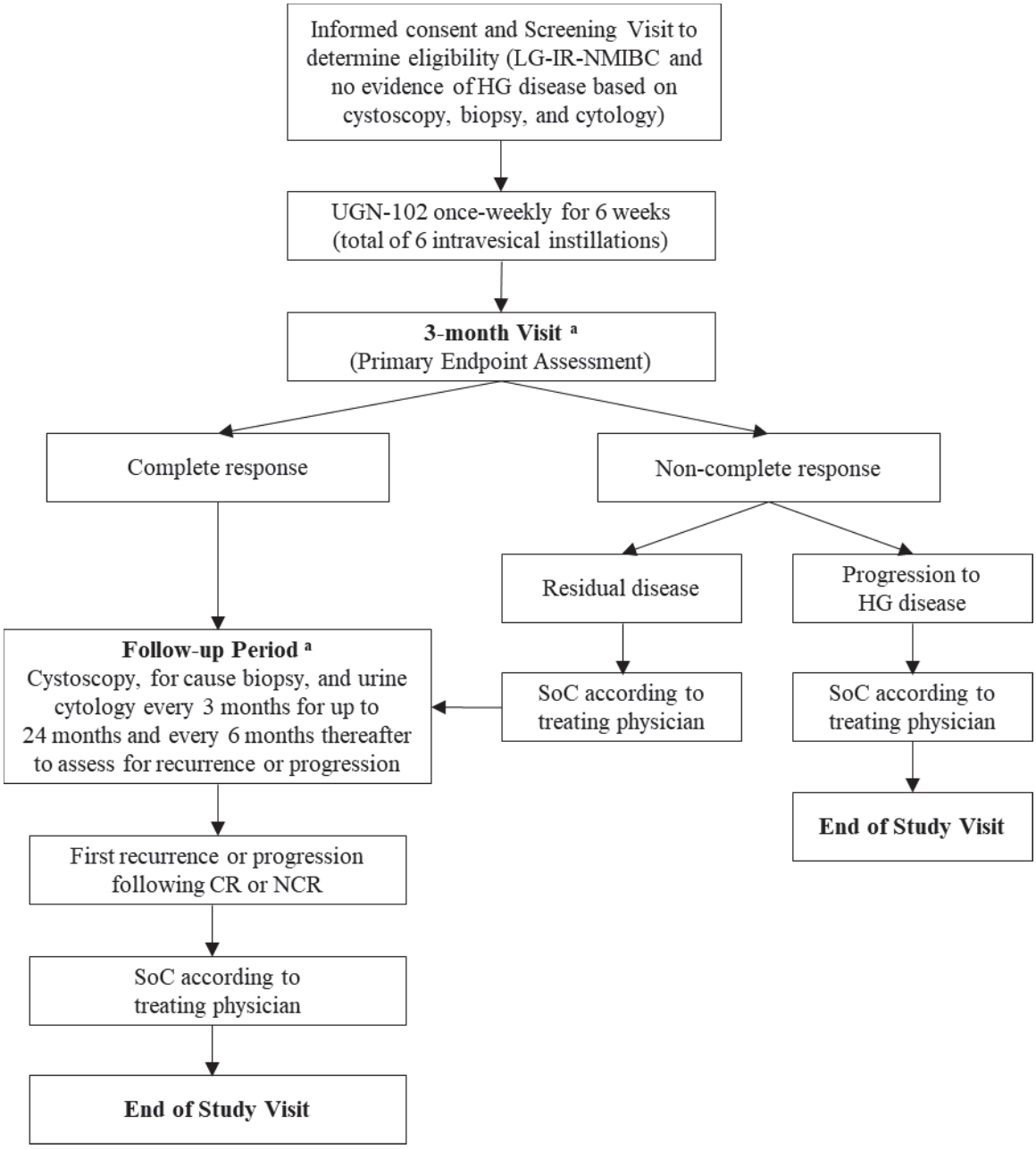
Patients who have a CR at the 3-month Visit, defined as having no detectable disease (NDD) in the bladder, will enter the Follow-up Period of the study. Patients who have an NCR (residual disease) will receive Investigator designated SoC treatment of remaining lesions and then enter the Follow-up Period of the study.

During the Follow-up Period, patients are scheduled to return to the clinic every 3 months for up to 24 months (i.e., 27 months after the first instillation) for evaluation of response. Patients who remain disease free at the 27-month Visit are scheduled to return to the clinic every 6 months for up to 36 months (i.e., 63 months after the first instillation) or until disease recurrence, disease progression, death, or the study is closed by the Sponsor, whichever occurs first.

Patients who have a disease recurrence during the Follow-up Period or a disease progression at the 3-month Visit or during the Follow-up Period will undergo Investigator designated SoC treatment and have a separate End of Study (EOS) Visit performed. The timing of the EOS Visit will be 3 months (±1 week) after SoC treatment of disease recurrence or progression. Details of the SoC provided (e.g., treatment modality, use and type of anesthesia, outcome, associated AEs) will be documented.

The study design is depicted in Figure 1 below. Schedules of activities from Screening to EOS are available in Protocol Section 1.3.

Figure 1: Study Schema



CR = complete response; EOS = end of study; HG = high grade; IR = intermediate risk; LG = low grade; NCR = non-complete response; NMIBC = non-muscle-invasive bladder cancer; SoC = standard of care.

[a] Patients confirmed to have a disease recurrence during the Follow-up Period or a disease progression at the 3-month Visit or during the Follow-up Period will undergo Investigator designated SoC treatment and have a separate EOS Visit performed.

3.1. Statistical Hypothesis

No formal hypotheses will be tested for the confirmation of efficacy. The hypotheses described in Section 3.2 are only for the justification of the sample size of this study. Confirmation of efficacy will be determined by the clinical meaningfulness of the estimates of the primary and key secondary endpoints.

Therefore, no adjustment will be made to the Type-I error rate for the analyses described in Section 4.2.

3.2. Sample Size Determination

The primary endpoint of this study is CRR at the 3-month Visit.

The key secondary endpoint DOR will be used to assess the long-term treatment effect in patients who responded at the 3-month Visit. If a patient has not had an event (recurrence or progression or death), DOR will be censored at the date of the last adequate disease assessment (i.e., a visit when all scheduled disease assessments have been done).

There are limited published data for DOR following TURBT in LG NMIBC. Based on the results published in the European Association of Urology Guidelines on NMIBC (TaT1 and CIS) - 2019 Update (Babjuk et al, 2019), the 1-year DFS rate of TURBT is 50% (i.e., median DFS of 12 months). The clinical hypothesis is that the treatment effect of UGN-102 will be similar to TURBT. The median DOR for patients with recurrent LG IR NMIBC in the TURBT-alone arm is estimated to be 9.1 months in BL006 (ATLAS).

In order to ensure 80% power to test the null hypothesis: 1 year DOR rate = 40% (i.e., median DOR of 9 months), versus the specific alternative hypothesis: 1 year DOR rate = 50% (i.e., median DOR of 12 months), it is calculated that approximately 92 events (recurrence/progression/death) need to be observed. An estimated total of 132 responders would be required at the 3-month Visit in order to achieve 92 events. Assuming a true CRR of 60% at the 3-month Visit, approximately 220 patients would need to be enrolled in this study to observe approximately 132 responders at the 3-month Visit.

If the true CRR of UGN-102 is 60% at the 3-month Visit and an initial 3-month CRR of 45% is considered as the minimum clinical benefit in this patient population, then a minimum of 116 responders out of 220 patients (observed CRR: 52.7%, 95% CI: 45.9, 59.5) will be needed so the exact 2-sided 95% CI excludes 0.45. Table 2 shows the power under different assumptions of true 3-month CRR.

Table 2: Statistical Power Scenarios

Sample Size	True 3-month CRR	Probability that the lower bound of an exact 2-sided 95% CI will exclude 0.45
220	0.60	>0.99
	0.55	0.845
	0.50	0.319

The above minimum clinical benefits of CRR and DOR are used just for the purpose of sample size justification.

Enrollment has been completed and the actual sample size of this study is 240, to account for expectation of approximately 10% early discontinuation.

3.3. Randomization and Blinding

Not applicable.

4. PLANNED ANALYSES

4.1. Interim Analyses

No formal interim analyses were planned in the Protocol.

4.2. Timing of Analyses

3-Month Data Freeze

A preliminary final analysis of the primary endpoint was performed after all patients attended the 3-month Visit for their primary disease evaluation or discontinued from the study prior to the 3-month Visit. This analysis was based on SAP V1.0 and used the imputation method described in Section 11.1.1, which is now referred to as a sensitivity analysis of the primary endpoint in this version of the SAP.

The safety profile of UGN-102 was also evaluated at the 3-month data freeze date through the reporting of AEs, including SAEs and AESIs.

Relevant electronic case report form (eCRF) pages, e.g., Informed Consent, Inclusion/Exclusion Criteria, Enrollment, Demographics, Evaluation of Response, Cystoscopy, Histopathology and Urine Cytology up to the 3-month Visit, and Standard of Care for Disease Recurrence or Progression up to the 3-month Visit were frozen prior to this analysis. No formal database lock (DBL) was conducted. Any changes in the Evaluation of Response of the 3-month Visit due to data cleaning activities that are performed after the 3-month data freeze date will be documented in a Note to File (NTF).

UroGen Statistics&Programming (S&P) team was restricted from accessing patient-level and aggregate-level efficacy data prior to the 3-month data freeze date and all UroGen personnel were restricted from accessing aggregate-level efficacy data prior to 3-month data freeze date. Further details regarding the data restrictions are available in the Trial Integrity Plan (TIP).

15-Month Data Freeze

The final analysis of the primary endpoint using the method described in the Protocol and in this version of the SAP and analyses of the key secondary efficacy endpoint (DOR), other secondary efficacy endpoints (DCR and DFS), as well as exploratory efficacy endpoints will be performed when patients in the 3-month$_{CR}$ analysis set (See Section 5 for detailed definition) have completed a minimum 12-month follow-up duration (i.e., completed the Month 15 Visit) or have early discontinued from the study. The actual date and rationale will be described in the CSR.

The safety profile (including clinical laboratory and vital signs assessments) of UGN-102

will also be updated at the 15-month data freeze date.

All eCRF pages will be frozen prior to this analysis. No formal DBL will be conducted for this analysis. Any changes in the Evaluation of Response at or prior to the 15-month Visit due to data cleaning activities that are performed after the 15-month data freeze date will be documented in a Note to File (NTF).

UroGen S&P team will remain restricted from accessing patient-level and aggregate-level efficacy data collected after the 3-month data freeze date and until approximately 2 months prior to the 15-month data freeze date. All UroGen personnel will be restricted from accessing aggregate-level efficacy data after the 3-month data freeze date and until approximately 2 months prior to the 15-month data freeze date. UroGen S&P team and selected UroGen personnel from other functions will gain access to efficacy data approximately 2 months prior to the 15-month data freeze date. Further details regarding the data restrictions are available in the TIP.

Final DBL

A formal DBL will be conducted at a later date.

The efficacy and safety profile of UGN-102 will be updated at the formal DBL.

UroGen S&P team will remain restricted from accessing patient-level and aggregate-level efficacy data collected after the 15-month data freeze date and until approximately 2 months prior to the formal DBL. All UroGen personnel will be restricted from accessing aggregate-level efficacy data after the 15-month data freeze date and until approximately 2 months prior to the formal DBL. UroGen S&P team and selected UroGen personnel from other functions will gain access to efficacy data approximately 2 months prior to the final DBL. Any changes to this plan will be described in the TIP.

5. ANALYSIS POPULATIONS

For the purposes of analysis, the following analysis sets are defined:

Analysis Set	Description
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. The analysis of the primary endpoint CRR at the 3-month Visit, analyses of PRO measures, and safety analyses will be conducted using the ITT analysis set.
Per Protocol (PP) Analysis Set	The per protocol (PP) analysis set will include the subset of the patients in the ITT analysis set without major protocol deviations that would confound efficacy evaluation. Protocol deviations or conditions leading to exclusion from the PP analysis set will be determined before database freeze. Sensitivity analyses of the primary endpoint (CRR at 3-month Visit) will be performed using data from PP analysis set if the PP and ITT analysis sets differ.

CR at 3-month Disease Assessment Analysis Set (3-month$_{CR}$)	The CR at 3-month disease assessment (3-month$_{CR}$) analysis set will consist of all patients from the ITT analysis set who achieved CR at the 3-month disease assessment. This analysis set will serve as the primary analysis set for analyses of the key secondary endpoint DOR and other secondary efficacy endpoints DFS and DCR rates at specific time points. This analysis set excludes all patients who had either missing or indeterminate response at the 3-month disease assessment.
Imputed CR at 3-month Disease Assessment Analysis Set (3-month Imputed CR)	The imputed CR at 3-month disease assessment (3-month Imputed CR) analysis set will consist of all patients from the ITT analysis set who had either observed CR or imputed CR at the 3-month disease assessment based on the imputation algorithm in Section 11.1.1 and will be used for sensitivity analyses of the key secondary endpoint DOR and DCR rates at specific time points.
NCR at 3-month Disease Assessment Analysis Set (3-month$_{NCR}$)	The NCR at 3-month disease assessment (3-month$_{NCR}$) analysis set will consist of all patients from the ITT analysis set who were NCR with residual disease at the 3-month disease assessment. The exploratory endpoint TTR will be analyzed using this analysis set.
NCR at Follow-up Disease Assessment Analysis Set (FUP$_{NCR}$)	The NCR at follow-up disease assessment (FUP$_{NCR}$) analysis set will consist of all patients from the 3-month$_{CR}$ analysis set who had disease recurrence or progression during the Follow-up Period. The preliminary response of these patients following the SoC treatment will be analyzed using this analysis set.

6. TREATMENT COMPARISONS

Not applicable.

7. CHANGES TO THE PLANNED ANALYSES

Not applicable.

8. GENERAL CONSIDERATIONS FOR DATA ANALYSES

Data will be listed and summarized according to Clinical Data Interchange Standards Consortium (CDISC) standards and International Council for Harmonisation (ICH) E9 Guideline. SAS software Version 9.4 or higher will be used to perform all data analyses, generate tables, figures, and listings.

Unless otherwise stated, all listings will be sorted by patient identification number and then by visit date if applicable.

Unless otherwise stated, continuous variables will be summarized with n, mean, median, standard deviation, minimum and maximum, and categorical variables will be summarized

with frequency counts and percentages.

8.1. Multicenter Studies

It is anticipated that patient accrual will be spread thinly across centers and summaries of data by center would be unlikely to be informative; therefore, data from all participating centers will be pooled prior to analysis.

A summary of the number of patients by country and by site will be provided.

8.2. Other Strata and Covariates

Not applicable.

8.3. Multiple Comparisons and Multiplicity

Not applicable.

9. DATA HANDLING CONVENTIONS

The following sections provide a general description of the derived and transformed variables used to describe and analyze the data. Separate analysis dataset specifications provide full details on all data derivations and transformations including standard Oncology algorithms. The analysis dataset specifications will clearly communicate the content and source of the datasets supporting the statistical analyses.

9.1. Premature Withdrawal and Missing Data

A patient will be considered to have completed the study if the patient completes all follow-up visits in accordance with the Schedule of Activities (SoA) (Protocol Section 1.3), has a recurrence or progression during the follow-up period, dies during the study, or is still in follow-up when EOS is declared.

A patient will be considered to have withdrawn from the study if the patient has not recurred, progressed, or died and is lost to follow-up, has withdrawn consent, or is no longer being followed at the Investigator's discretion.

No time-to-event data were analyzed at the time of the preliminary final analysis of the primary endpoint. For endpoints which determine the percentage of CRs, indeterminate or missing responses will be handled according to the rules specified in the corresponding sections (11.1.1 and 11.3.1.1).

Missing data occurs when any requested data are not provided, leading to blank fields on the collection instrument. These data will be indicated using a "blank" in patient listing displays. Answers such as "Not applicable" and "Not evaluable" are not considered to be missing data and will be displayed as such.

If relationship of AEs or SAEs to study treatment is missing it will be assumed to be "Yes." There will be no other imputation for missing data other than what is described in Appendix 1.

All relevant data will be summarized and/or listed. Summaries of the number of patients who have discontinued study treatment, as well as details on those in follow-up will be provided.

9.2. Derived and Transformed Data

9.2.1. Reference Date

There are two reference dates:

- The reference date for age is the date of Informed Consent as age is an eligibility requirement.
- The safety/efficacy reference date is the date of first instillation of UGN-102 and will be used to calculate study day for safety/efficacy measures.

9.2.2. Study Day

The study day for safety and efficacy measures are the same in this study. If the date of interest occurs on or after the reference date, then the study day will be calculated as (date of interest - reference date) + 1. If the date of interest occurs before the reference date, then the study day will be calculated as (date of interest – reference date). There is no study day 0.

9.2.3. Calculation of Durations

Durations (e.g., duration of adverse event, time to recurrence between prior NMIBC episodes, etc.) will be calculated as stop date minus start date plus one.

When reporting time to event durations (DFS, TTR, DOR) or any other duration 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. These algorithms for time to event return decimal numbers and ignore the actual numbers of days in the months or years between start date and stop date. The "year" used in these algorithms is 365.25 days long, and the "month" is one twelfth of that year.

9.2.4. Imputation of Partial Dates

In general, imputed partial dates will not be used to derive study day or duration (e.g., duration of AEs). In addition, imputed dates are not used for deriving the last contact date.

Imputed dates will not be displayed in listings. However, where necessary, display macros may impute dates as temporary variables for the purpose of sorting data in listings only. In addition, partial dates may be imputed for ‘slotting’ data for specific analysis purposes as outlined below.

The partial date imputation will follow Analysis Data Model (ADaM) conventions. The ADaM approach is to populate the numeric date variables with the imputed date and add a flag variable to the dataset that indicates the level of imputation. The flag variable can contain the values: blank, 'D,' 'M,' 'Y.'

blank: indicates that no imputation was done

D='Day': indicates that the day portion of the date is imputed

M='Month': indicates that the month and day portions of the date are imputed

Y='Year': indicates that the entire date (year, month, and day) is imputed

Details on imputing partial dates for specific datasets are outlined in Appendix 1.

9.2.5. Baseline Definition

In general, baseline will be defined as the most recent non-missing value prior to the first instillation of UGN-102, i.e., prior to Day 1.

For the PRO analyses, baseline is defined as the last PRO assessment on or prior to the first instillation of UGN-102, i.e., on or prior to Day 1.

9.2.6. Change from Baseline

Change from baseline is calculated as:

- For records occurring after baseline: visit value – baseline value.

Percent change from baseline is calculated as:

- For records occurring after baseline: change from baseline/ baseline value * 100.

If either the baseline or visit value is missing, the change from baseline and/or percent change from baseline is set to missing as well.

9.2.7. Analysis Visit Window

A windowing convention will be used to determine the analysis value for a given study visit only for the following visit-based assessments:

- Patient Reported Outcome (PRO)
- Urological exam
- Clinical chemistry, hematology, and urinalysis
- Vital signs

The window for the visits following baseline will be constructed in such a way that the upper limit of the interval falls halfway between the two visits (the lower limit of the first post-Day 1 visit will be Day 2). If an odd number of days exists between two consecutive visits, then the upper limit will be taken as the midpoint value minus 1 day.

Table 3: Analysis windows

Analysis Visit	Target Day	Analysis Window (Study Day)
Baseline	< 1	< 1
Day 1	1	1
Week 2	8	2 - 11

Analysis Visit	Target Day	Analysis Window (Study Day)
Week 3	15	12 - 18
Week 4	22	19 - 25
Week 5	29	26 - 32
Week 6	36	33 - 63
Month 3	90	64 - 134
Month 6	180	135 - 224
Month 9	270	225 - 314
Month 12	360	315 - 404
Month 15	450	405 - 494
Month 18	540	495 - 584
Month 21	630	585 - 674
Month 24	720	675 - 764
Month 27	810	765 - 899
Month 33	990	900 - 1079
Month 39	1170	1080 - 1259
Month 45	1350	1260 - 1439
Month 51	1530	1440 - 1619
Month 57	1710	1620 - 1799
Month 63	1890	≥1800

9.2.8. Multiple Assessments

Endpoints will be reported according to the nominal visit date on which they were reported. Analysis windows will be applied during dataset creation for the summaries of clinical hematology, chemistry, urinalysis, vital signs, urological examination, and the PRO assessments. Both scheduled and unscheduled data will be included in the analysis windows and the summary sections labelled as worst-case.

One or more results for a particular variable may be obtained in the same visit window. In such an event, the result with the date closest to the expected visit date will be used in the analysis. If two observations are equidistant from the expected visit date, the later observation will be used in the analysis.

If multiple assessments are reported on the same date for the same scheduled planned time, then the worst-case result will be analyzed, except for laboratory data reported from both central and local laboratories. If laboratory data is reported from both central and local laboratories with the same date, then the central laboratory data will be analyzed to provide consistency with measurements from other patients.

Data from all assessments (scheduled and unscheduled), including multiple assessments, will be included in the listings.

10. PATIENT DISPOSITION, DEMOGRAPHICS AND OTHER BASELINE CHARACTERISTICS

Unless otherwise stated, all tables and listings in this section will be based on the ITT population and will display the analysis population as 'ITT Analysis Set'.

10.1. Disposition of Patients

A summary of the number of patients who signed informed consent, met all eligibility criteria, not met all eligibility criteria and reasons will be provided.

A summary of the number of patients in each of the analysis sets described in Section 5 will be provided. In addition, the number of patients by country and site will be summarized using the ITT analysis set.

A summary of treatment with UGN-102 will be provided. This display will show the number and percentage of patients who discontinued the treatment with the primary reasons for discontinuation. Reasons for treatment discontinuation will be presented in the order they are displayed in the eCRF.

A summary of patient status of study completion (completed the study versus withdrew/discontinued early) and reasons will be provided. This display will show the number and percentage of patients who completed, discontinued or are ongoing in the study, and primary reasons for study completion/discontinuation. Reasons for study completion or withdrawal will be presented in the order they are displayed in the eCRF.

A combined listing of treatment and study discontinuation will be generated.

10.2. Protocol Deviations

All protocol deviations will be summarized and listed and will include severity of the deviation (e.g., major versus minor) and the categories (e.g., eligibility criteria, protocol implementation, safety etc.) using the ITT population. See the Protocol Deviation Management Plan for details of protocol deviations.

Patients with major protocol deviations that would confound efficacy evaluation, would be excluded from the PP analysis set. All protocol deviations or conditions leading to exclusion from the PP analysis set will be detailed in the protocol deviation plan and will be finalized prior to the data freeze. Sensitivity analyses of the primary endpoint CRR may be performed using data from the PP analysis set if the PP and ITT analysis sets differ.

10.3. Demographic and Baseline Characteristics

Demographic characteristics (e.g., age, sex, race, ethnicity, height, weight and BMI) will be summarized and listed. Age (years), height (cm), weight (kg) and BMI (kg/m^2) will be summarized using the mean, standard deviation, minimum, median (quartiles), and maximum as appropriate. In addition, age will also be categorized and summarized by < 65, ≥ 65 to < 75, ≥ 75 to < 85 and ≥ 85; < 65 and ≥ 65; < 75 and ≥ 75. BMI will also be categorized and summarized by < 30 and ≥ 30.

The count and percentage will be computed for sex, race, ethnicity, attitude about undergoing TURBT prior to participating, and American Society of Anesthesiologists (ASA) physical status.

The following baseline bladder cancer prognostic factors may be used to prespecify subgroup analyses and will be provided by count and percentage:

- Treatment course (instillations): (6, < 6)
- Tumor size (cm): (≤ 3, > 3)
- Tumor count (single, multiple)
- Disease history
  - Previous NMIBC episodes (Yes, No)
  - Previous NMIBC episodes within 1 year of current diagnosis (Yes, No)
  - Number of previous NMIBC episodes (0, 1, 2, > 2)
  - Previous LG NMIBC episodes (Yes, No)
  - Previous LG NMIBC episodes within 1 year of current diagnosis (Yes, No)
  - Number of previous LG NMIBC episodes (0, 1, 2, > 2)
- Number of prior TURBT to treat NMIBC (0, 1, 2, > 2)
- Number of prior TURBT to treat LG NMIBC (0, 1, 2, > 2)
- Smoking history (smoker, non-smoker)

Smoking has been recognized as the strongest established risk factor for bladder cancer. Smoking related data has been collected in the eCRF in great details and will be summarized by:

- Category for substance use (Cigarettes, Cigar, Pipe, Other, and any combination)
- Baseline Status (Current vs Former)
- Duration (< 10, 10 - 20, 21 - 30, > 30 years)
- For cigarettes only: number of cigarettes per day (< 5, 5 – 15, 16 – 25, > 25)

A summary of prior NMIBC along with its standard of care (TURBT) will be provided. Number of prior NMIBC episodes, time to recurrence of baseline NMIBC episode, number of prior TURBT procedures, and time since last TURBT will be summarized using the mean, standard deviation, minimum, median (quartiles), and maximum as appropriate. In addition, the count and percentage will be computed for number of prior TURBT groups (0, 1, 2, ≥ 3) and time since last TURBT procedure at the time of first instillation of UGN-102 (≤ 365 days, >365 days). A listing of prior NMIBC diagnoses and TURBT procedures will also be generated.

10.4. General Medical History and Urothelial Carcinoma Related Medical History (UCMH)

Medical history reported terms will be coded to a system organ class and preferred term (PT) using the most recent version of Medical Dictionary for Regulatory Activities (MedDRA) Version 23.1 or later. Number and percent of patients reporting any medical history by system organ class and PT for the ITT analysis set will be provided. A patient with multiple medical conditions will be counted once per system organ class 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 system organ class (then alphabetically for ties), then by descending order of frequency of PT within each system organ class (then alphabetically for ties). A patient data listing of medical history will be provided.

The UCMH status at baseline will be summarized as:

- Time in months between the date of initial diagnosis and the first instillation of UGN-102
- Number of prior UCMH episodes
- Visual appearance of the lesions at baseline
- Staging of prior UCMH episodes (Noninvasive papillary carcinoma (Ta), Tumor invades lamina propria (T1), Tumor invades muscular propria (T2), Tumor invades perivesical tissue/fat (T3), Tumor invades prostate, uterus, vagina, pelvic wall or abdominal wall (T4), Other
- Grade of the prior UCMH episodes (Papillary urothelial carcinoma - low grade, Papillary urothelial neoplasm of low malignant potential, Urothelial carcinoma in situ, Papillary urothelial carcinoma - high grade, Other)
- Prior UCMH episode treated (Yes, No)
- Treatment given for prior episodes of UC (TURBT, Fulguration, BCG, MMC, Other).

10.5. Prior and Concomitant Medications

All medications will be coded using the World Health Organization (WHO) Drug Dictionary, September 2020 version or later. The Anatomical Therapeutic Chemical (ATC) Level 2, Level 3 and preferred name will be used to list and summarize the medication data. Analysis of prior and concomitant medications use will be performed in the following manner:

- Prior medications: Any medications that were started and stopped prior to the first instillation of UGN-102. The number and percentage of patients reporting the use of prior medications by ATC Level 2, ATC Level 3 and preferred names 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: Any medications that are ongoing or any new medication administered following the first instillation of UGN-102. The number and percentage of patients reporting the use of concomitant medications by ATC Level 2, ATC Level 3 and preferred names 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.

10.6. Prior and Concomitant Surgical Procedures

Prior surgical procedures are surgical procedures that were done prior to the first instillation of UGN-102.

Concomitant surgical procedures are surgical procedures that were done following the first instillation of UGN-102 through end of study.

All surgical procedures will be coded to a system organ class and PT using the most recent version of the MedDRA dictionary (version 23.1 or later). Prior and concomitant surgical procedures will be summarized by SOC and PT in a similar manner to that described for medical history in Section 10.4.

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

10.7. Study Follow-up

Duration of follow-up (months) in the ITT and PP analysis sets is defined as

(last contact date or death date – date of first instillation of UGN-102 + 1) / 30.4375

Duration of follow-up (months) in the 3-Month CR, 3-Month Imputed CR, and 3-Month NCR analysis sets is defined as

(last contact date or death date – 3-Month date + 1) / 30.4375

Duration of follow-up will be summarized using descriptive statistics.

Additionally, the number and percentage of patients who underwent disease assessments will be summarized by Visit.

11. EFFICACY ANALYSES

Efficacy analyses will be based on the ITT population as defined in Section 5 or a subset of the ITT population as described for each analysis. All analyses will be presented by the single UGN-102 arm.

Efficacy assessments are based on cystoscopy, for-cause biopsy, and cytology to assess response status of a patient at each scheduled visit. Investigator assessments will be considered as primary assessments.

11.1. Primary Efficacy Analyses

11.1.1. CRR at 3-Month Visit

The CRR is defined as the proportion of patients in the ITT population who achieved CR at the 3-month Visit (3 months after the first instillation of UGN-102) as determined by cystoscopy, for cause biopsy, and urine cytology. If response cannot be evaluated for a patient at the 3-month Visit, the patient will be considered as an NCR for the purpose of the analysis and will be included in the denominator of the CRR. Patient response will be evaluated by the investigator according to the evaluation of response (EOR) outlined in the Protocol Appendix 1 and will be recorded on the 3-Month EOR eCRF page.

CRR will be presented along with the exact 95% CIs. The exact two-sided 95% CI for the CRR will be derived using the Clopper-Pearson method (Clopper and Pearson, 1934) (SAS

PROC FREQ with binomial option). In addition, reasons for NCR comprised of 'Residual Disease,' 'Progression' and 'Indeterminate' will be tabulated using the number and percentages of patients.

A sensitivity analysis will be performed where the response at the 3-month Visit will be imputed for the following cases:

- If a patient received SoC prior to the 3-month assessment, then the 3-month assessment will be imputed as "NCR".
- If the 3-month response is missing or "Indeterminate", and a patient received SoC prior to the next disease assessment then the 3-month response will be imputed as "NCR".
- If the 3-month response is missing or "Indeterminate", and a patient has not received any further disease assessment visits then the 3-month response will be imputed as "NCR".
- If the 3-month response is missing or "Indeterminate", and the next disease assessment is "NCR" then the 3-month response will be imputed as "NCR".
- If the 3-month response is missing or "Indeterminate", and a patient did not receive SoC prior to the next disease assessment and the next disease assessment is "CR" then the 3-month response will be imputed as "CR".

If response cannot be imputed for a patient at the 3-month Visit (e.g., due to death or lost to follow-up or no follow-up visits beyond the 3-month Visit), the patient will be considered as NCR for the purpose of the analysis and will be included in the denominator of the CRR. The date of disease assessment associated with CR/NCR will be determined using the date of cystoscopy, biopsy, or cytology, whichever occurs first.

Sensitivity analyses may be performed using data from PP analysis set if the PP and ITT differ.

11.1.1.1. Subgroup Analysis

Subgroup analyses will be performed on each level of the following factors using ITT population and a corresponding summary table will be produced. The summary will also be displayed by a forest plot.

Subgroups will include (but not limited to):

- Baseline age group (years): (< 65, ≥ 65 and < 75, ≥ 75)
- Sex: (Male, Female)
- BMI category (kg/m^2): (< 30, ≥ 30)
- Treatment course (instillations): (6, < 6)
- Tumor size (cm): (≤ 3, > 3)
- Tumor count: (Single, Multiple)
- Previous NMIBC episodes (Yes vs No)
- Previous NMIBC episodes within 1 year of current diagnosis (Yes vs No)
- Number of previous NMIBC episodes (≤ 1 vs > 1, ≤ 2 vs > 2)
- Prior TURBT to treat NMIBC (Yes vs No)

- Number of prior TURBT to treat NMIBC (≤ 1 vs > 1, ≤ 2 vs > 2)
- Smoking History (Smoker, Non-smoker)

11.2. Key Secondary Efficacy Analyses

11.2.1. DOR in Patients Who Achieved CR at the 3-month Visit

The key secondary endpoint is DOR in patients who achieved CR at the 3-month Visit, defined as the time from the 1st documented CR to the earliest date of recurrence or progression as determined using the date of cystoscopy, for cause biopsy, or cytology, or death due to any cause, whichever occurred first.

The DOR in months is calculated as (first event date / censored date – date of 1st documented CR + 1) / 30.4375.

The distribution of DOR will be estimated using the Kaplan-Meier method. Median times of DOR, first and third quartiles along with 95% CI (Brookmeyer and Crowley, 1982) will be estimated. The median follow-up time along with 95% CI will be estimated using the inverse Kaplan-Meier method. A figure and listing of DOR time will also be provided. This analysis will be based on the 3-month$_{CR}$ analysis set.

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.

Patients without any adequate disease assessments during the follow-up will be censored at the date of first documented CR at the 3-Month disease assessment. An adequate assessment is defined as an assessment where the visit level response is not missing, and not Indeterminate.

If a patient has not had an event (recurrence or progression or death), DOR will be censored at the date of the last adequate disease assessment (i.e., a visit when all scheduled disease assessments have been done). DOR will be censored if no DOR event is observed before the data cut-off date of analysis or the date when a new anti-cancer therapy or another investigational treatment for cancer is started, whichever occurs earlier. The censoring date will be the date of last adequate disease assessment before either of these two dates.

During the Follow-up Period, patients will return to the clinic every 3 months for up to 24 months (i.e., 27 months after the first instillation) for evaluation of response. Patients who remain disease free at the 27-month Visit will continue to return to the clinic every 6 months for up to 36 months (i.e., 63 months after the first instillation) or until disease recurrence, disease progression, death, or the study is closed by the Sponsor, whichever occurs first.

Patients who recurred or progressed or died after an extended lost to follow-up (period without adequate assessment), will be censored at their date of last adequate assessment prior to recurrence or progression or death, even if subsequent information is available regarding recurrence or progression or death. A patient will be considered to have an extended lost to follow-up if they missed two or more consecutive adequate assessments. For example, during the Follow-up Period up to Month 27, a patient will be considered having an extended lost to follow-up if the patient did not have an adequate assessment of

210 days (6 months + 4 weeks ~ 6*30.4375+28) and then had a recurrence (or progression or death).

If a recurrence or progression or death is observed after a single missing or Indeterminate disease assessment, the actual date of recurrence or progression or death will be used as the event date.

Example (protocol defined schedule of follow-up disease assessments is every 3 months ± 2 week): A patient had the disease assessments at 3 months (CR) - 6 months (CR) – 9 months (missing) – 12 months (missing) - Recurrence. Then the DOR status of this patient will be censored, and the censoring date will be the date of 6-month visit.

Non-Protocol Therapy (NPT) are prohibited concomitant anti-cancer medications or surgical procedures as outlined in Section 6.5 of the Protocol. These medications will be identified via clinical review of the Concomitant Medications and Surgery datasets. If a patient received multiple NPTs, the earliest date will be considered as the start date of NPT. These dates will be used to define event or censoring dates for DOR, DFS, and TTR.

For patients who receive an NPT the following rules will apply:

- If the start date of the NPT is partial (i.e., either missing the day but has the month and year available or missing both day and month), the imputation rules described in Appendix 1 will be applied. No imputation will be made for completely missing dates.
- If the NPT is started prior to documented disease recurrence or progression, then DOR will be censored at the date of the last adequate assessment that is no later than the date of initiation of NPT (i.e., if an assessment occurs on the same day as the start of NPT, the assessment will be used - as it will be assumed the assessment occurred prior to the start of NPT). The date of response at the last adequate assessment will be used as the censoring date.
- If the NPT is started prior to the first disease assessment (i.e., 3-Month Visit), then DOR will be censored at the date of the first instillation (Day 1).
- If a patient does not have any post-baseline adequate assessment, DOR will be censored at the date of the first instillation (Day 1).
- If a patient has neither recurrence or progression nor died nor started NPT, then DOR will be censored at the date of the last adequate assessment.

A summary of the assignments for progression and censoring dates for DOR are specified in the following table.

Table 4: Assignments for Event and Censoring Dates for DOR analysis

Scenario	Date of Event or Censoring	Outcome
No adequate assessments during follow-up[1]	Date of 1st documented CR at the 3-Month disease assessment	Censored
Received NPT after 3-Month disease assessment prior to	Date of last 'adequate' assessment[1,3] on or prior to starting NPT[2]	Censored

documented recurrence or progression		
No recurrence or progression or death	Date of last 'adequate' assessment[1,3]	Censored
No recurrence or progression and the last response evaluation is 'indeterminate'	Date of last 'adequate' assessment[1,3]	Censored
Recurrence or progression or death after 2 or more missed visits	Date of last 'adequate' assessment[1,3] prior to missed assessments	Censored
Disease recurrence or progression documented at or between scheduled visits	Date of first recurrence or progression[3]	Event
Death prior to documented disease recurrence or progression	Date of death	Event

[1] An adequate assessment is defined as an assessment where the visit level response is not missing and not indeterminate.
[2] If recurrence/progression and NPT occur on the same day assume the recurrence/progression was documented first.
[3] The earliest of (i) Date of cystoscopy; or (ii) Date of biopsy, or (iii) Date of cytology

A sensitivity analysis will be performed based on the 3-month Imputed CR analysis set.

Sensitivity analyses may be performed ignoring extended lost to follow-up. That is, patients who recurred or progressed or died after an extended lost to follow-up (i.e., after missing two or more consecutive visits), is considered as an event. Additional sensitivity analyses may also be performed ignoring NPT. That is, patients who recurred or progressed or died after receiving NPT is considered as an event.

11.2.1.1. Subgroup Analysis

Subgroup analyses will be performed on each level of the subgroups defined in Section 11.1.1.1 and a corresponding summary table will be produced. The summary will also be displayed by a forest plot.

11.3. Other Secondary Efficacy Analyses

11.3.1. DCR

DCR rate at scheduled disease assessment timepoints is defined as the proportion of patients who had CR at the 3-Month disease assessment and maintained CR up to that particular follow-up disease assessment. This endpoint will be summarized only for CR patients at the 3-Month visit. Patient response will be evaluated by the investigator

according to the guidance on EOR in Protocol Appendix 1 and will be recorded on the follow-up EOR eCRF page.

The EOR will be based on central laboratory results, when central laboratory results are unavailable local laboratory results will be used. When results from both central and local laboratories are available, central laboratory results will supersede the local laboratory results.

Three types of analyses will be performed to estimate the visit level DCR rate. Patients who have indeterminate response or missed a visit during follow-up will have the response imputed according to the rules specified in Section 11.3.1.1 prior to conducting the analyses.

If a patient receives SoC or any other NPT during the follow-up period, the response status at all subsequent follow-up visits after the SoC (or NPT) start date will be considered as recurrence for these analyses.

Analysis 1

DCR rate will be presented with nominal 95% exact CI (Clopper-Pearson) at each scheduled timepoint (e.g., 6, 9, 12 months etc.). The denominator to calculate the proportion will include all patients who were CR at 3-month visit.

Number and percentage of patients who have indeterminate response that is not imputed (i.e. not confirmed by subsequent visits) and those who have recurrence (LG disease or progression), will also be presented at each scheduled timepoint.

This analysis represents the worst-case scenario where all patients are assumed as not CR for all timepoints beyond their last disease evaluation, including those patients who are still ongoing in the study and have not reached those timepoints, and any patients who completed/discontinued the study and would have not reached those timepoints had they not completed/discontinued the study.

Analysis 2

The above analysis will be repeated using denominator the number of patients who were CR at 3-month visit and reached the follow-up visit or would have reached the follow-up visit at the time of the data read-out had they not discontinued/completed the study.

This analysis represents a scenario where all patients who completed/discontinued the study are assumed as not CR for all time points that they would have reached had they not discontinued/completed the study. Any patients ongoing in the study are excluded from timepoints that they have not reached and any patients who have completed/discontinued the study are excluded from timepoints that they would have not reached had they not discontinued/completed the study.

Illustrating cases of patients who were CR at 3-month visit and would be excluded from the denominator of Analysis 2 for computing proportion of patients who maintained CR at 6, 9, 12, 15, and 18 months are provided below.

Case	Exclusion from the timepoint				
	6-Month	9-Month	12-Month	15-Month	18-Month

Patient A: ongoing, CR up to 15-Month visit, and has not reached 18-Month visit	No	No	No	No	Yes
Patient B: ongoing, CR up to 12-Month visit, indeterminate response at 15-Month visit and has not reached 18-Month visit	No	No	No	No	Yes
Patient C: completed study due to recurrence/progression at 9-Month visit, would have reached up to 15-Month visit had they not recurred/progressed	No	No	No	No	Yes
Patient D: completed study due to death prior to 9-month visit and would have reached up to 18-Month visit had they not died	No	No	No	No	No
Patient E: CR up to 6-Month visit, discontinued study prior to 9-Month visit, would have reached up to 15-Month visit had they not discontinued	No	No	No	No	Yes
Patient F: CR up to 9-Month visit, indeterminate response at 12-Month visit, discontinued study prior to 15-Month visit, would have reached up to 15-Month visit had they not discontinued	No	No	No	No	Yes

Analysis 3

The above analysis will be repeated using denominator the number of patients who were CR at 3-month visit and reached the follow-up visit with conclusive disease evaluation (i.e., either CR or recurrence or progression) or would have reached the follow-up visit at the time of the data read-out had they not completed the study (due to recurrence, progression, or death).

This analysis makes no assumptions on the response status of patients who do not have a conclusive disease evaluation. Any patients ongoing in the study are excluded from timepoints that they have either not reached or have not had a conclusive disease evaluation (i.e. not confirmed by subsequent visits). Any patients who had CR at their last conclusive disease evaluation prior to completing/discontinuing the study are excluded from subsequent timepoints.

Illustrating cases of patients who were CR at 3-month visit and would be excluded from the denominator of Analysis 3 for computing proportion of patients who maintained CR at 6, 9, 12, 15, and 18 months are provided below. Highlighted in green are the cases where Analysis 3 and Analysis 2 differ.

	Exclusion from the timepoint

Case	6-Month	9-Month	12-Month	15-Month	18-Month
Patient A: ongoing, CR up to 15-Month visit, and has not reached 18-Month visit	No	No	No	No	Yes
Patient B: ongoing, CR up to 12-Month visit, indeterminate response at 15-Month visit and has not reached 18-Month visit	No	No	No	Yes	Yes
Patient C: completed study due to recurrence/progression at 9-Month visit, would have reached up to 15-Month visit had they not recurred/progressed	No	No	No	No	Yes
Patient D: completed study due to death prior to 9-month visit and would have reached up to 18-Month visit had they not died	No	No	No	No	No
Patient E: CR up to 6-Month visit, discontinued study prior to 9-Month visit, would have reached up to 15-Month visit had they not discontinued	No	Yes	Yes	Yes	Yes
Patient F: CR up to 9-Month visit, indeterminate response at 12-Month visit, discontinued study prior to 15-Month visit, would have reached up to 15-Month visit had they not discontinued	No	No	Yes	Yes	Yes

Sensitivity analyses may be performed based on the 3-month Imputed CR analysis set.

11.3.1.1. Imputation of Missing Visit Level Response

Some patients had disruptions to their visit schedules due to COVID-19 situation or due to the war situation in Europe, after completing their 3-Month Visit. As a result, scheduled visits beyond the 3-Month Visit were delayed (outside the protocol window) or missed. 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:

1. If any recorded response is CR, then all previous missing or indeterminate scheduled visits responses are CR (Patient A), using the method last observation carried backward (LOCB).
2. If any recorded response is recurrence disease (RD), then all previous missing or indeterminate scheduled visits responses are RD (Patient B; Patient C), using the

method last observation carried backward (LOCB); unless there is a previous CR response from an unscheduled visit closer in time (days) to the scheduled visit in which case the visit response will be CR (Patient D), using the method last observation carried forward (LOCF). In the case of ties, the scheduled visit response 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	Target Day/ Study Day	Visit Window	Patient A	Patient B	Patient C	Patient D	Patient E
3-Month	90 (± 1 week)	83-97	CR	CR	CR	CR	CR
6-Month	180 (± 2 week)	166-194	U/CR	U/ RD	U/RD	CR	CR
Unscheduled	190						CR
Unscheduled	220					CR	
9-Month	270 (± 2 week)	256-284	CR	RD	RD	CR	RD
Unscheduled	300				RD		
Unscheduled	350						RD
12-Month	360 (± 2 week)	346-374				RD	
…….							

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

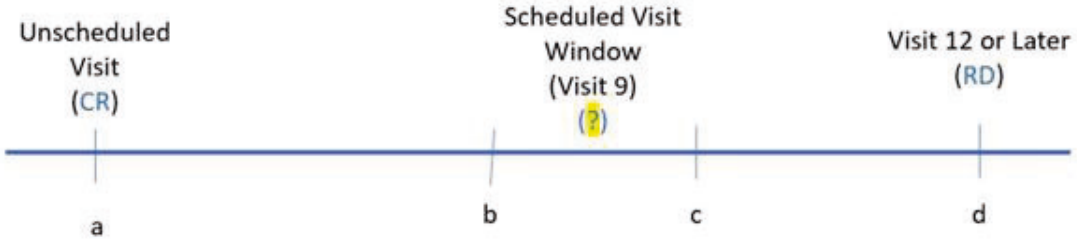
To further explain 'closer in time (days) to the scheduled visit,' the following variables are defined:

a = Study day of the CR unscheduled visit

b = Lower limit of next scheduled visit (days)

c = Upper limit of the next 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, 36 days before the beginning of the 9-Month Visit window. This patient also has a RD response on study day 346 (say),

62 days (346-284) after the end of the 9-Month Visit window. Since the CR response is closer to the 9-Month Visit, the response will be CR.

A patient listing will be generated to display visit level responses as recorded in the eCRF along with derived responses (indicating the method of derivation, LOCF or LOCB). Evaluation based on local laboratory results will be flagged.

If the 3-month EOR is 'Indeterminate', it will be imputed using the rules defined in Section 11.1.1.

11.3.2. DFS in Patients Who Were CR at the 3-Month Visit

The DFS in patients who were CR at the 3-month Visit, is defined as the time from the first instillation to the earliest date of recurrence or progression as determined using the date of cystoscopy, for cause biopsy, or cytology, or death due to any cause, whichever occurred first. If a patient has not had an event (recurrence or progression 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 DFS in months is calculated as (first event date / censored date – date of first instillation + 1) / 30.4375.

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. The median follow-up time along with 95% CI will be estimated using the inverse Kaplan-Meier method. A figure and listing of DFS time will also be provided. This analysis will be based on the 3-month$_{CR}$ analysis set.

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.

Sensitivity analyses similar to those described in Section 11.2.1 may be performed for DFS.

The rules for defining recurrence and censoring are described in Table 4. Censoring rules will follow those of the DOR analysis defined in Section 11.2.1.

An additional analysis will be performed on the ITT analysis set. For this analysis, the DFS$_{ITT}$ is defined the same as the DFS performed on the 3-month$_{CR}$ analysis set. Censoring rules will follow the table below. Sensitivity analyses similar to those described in Section 11.2.1 may be performed on the DFS$_{ITT}$.

Table 5: Assignments for Event and Censoring Dates for DFS$_{ITT}$ analysis

Scenario	Date of Event or Censoring	Outcome
No post-baseline adequate assessments[1]	Date of first instillation	Censored
Received NPT prior to 3-month disease assessment	Date of first instillation	Censored
Received NPT after 3-month disease assessment prior to documented recurrence or progression	Date of last 'adequate'[1,3] assessment on or prior to starting NPT[2]	Censored

Death prior to 3-month disease assessment	Date of death	Event
Progression at or prior to 3-month disease assessment	Date of progression[3]	Event
Death prior to documented disease recurrence or progression	Date of death	Event
No recurrence or progression or death	Date of last 'adequate' assessment[1,3]	Censored
No recurrence or progression and the last response evaluation is 'Indeterminate'	Date of last 'adequate' assessment[1,3]	Censored
Recurrence or progression or death after 2 or more missed visits	Date of last 'adequate'[1,3] assessment prior to missed assessments	Censored
Disease recurrence or progression documented at or between scheduled visits	Date of first recurrence or progression[3]	Event

[1] An adequate assessment is defined as an assessment where the visit level response is not missing and not indeterminate.

[2] If recurrence/progression and NPT occur on the same day assume the recurrence/progression was documented first.

[3] The earliest of (i) Date of cystoscopy; or (ii) Date of biopsy, or (iii) Date of cytology.

11.4. Exploratory Efficacy Analyses

11.4.1. TTR in Patients Who Were NCR at the 3-month Visit

The TTR in patients who were NCR (residual disease) at the 3-month Visit, is defined as the time from the date of first SoC treatment after NCR to the earliest date of recurrence or progression as determined using the date of cystoscopy, for cause biopsy, or cytology, whichever occurred first.

The TTR in months is calculated as (date of recurrence / progression / censoring – date of first SoC treatment after NCR + 1) / 30.4375.

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. The median follow-up time along with 95% CI will be estimated using the inverse Kaplan-Meier method. A figure and listing of TTR time will also be provided. This analysis will be performed on the 3-month$_{NCR}$ analysis set.

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.

Death will not be considered an event in this analysis. If a patient has not had an event (recurrence or progression), TTR 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 determining event and censoring are described in the following table.

Table 6: Assignments for Event and Censoring Dates for TTR analysis

Scenario	Date of Event or Censoring	Outcome
No adequate assessments[1]	Date of first SoC treatment after NCR	Censored
Received NPT after 3-month disease assessment prior to documented recurrence or progression	Date of last 'adequate'[1,3] assessment on or prior to starting NPT[2]	Censored
Death prior to documented disease recurrence or progression	Date of death	Censored
No recurrence or progression or death	Date of last 'adequate' assessment[1,3]	Censored
No recurrence or progression and the last response evaluation is 'Indeterminate'	Date of last 'adequate' assessment[1,3]	Censored
Recurrence or progression or death after 2 or more missed visits	Date of last 'adequate'[1,3] assessment prior to missed assessments	Censored
Disease recurrence or progression documented at or between scheduled visits	Date of first recurrence or progression[3]	Event

[1] An adequate assessment is defined as an assessment where the visit level response is not missing and not indeterminate.

[2] If recurrence/progression and NPT occur on the same day assume the recurrence/progression was documented first.

[3] The earliest of (i) Date of cystoscopy; or (ii) Date of biopsy, or (iii) Date of cytology.

An additional analysis will be performed on the ITT analysis set. For this analysis, the TTR_{ITT} is defined as the time from the date of the first instillation to the earliest date of recurrence or progression as determined using the date of cystoscopy, for cause biopsy, or cytology, whichever occurred first. Censoring rules for TTR_{ITT} analysis will follow those described in the table below.

The TTR_{ITT} in months is calculated as (date of recurrence / progression / censoring – date of first instillation + 1) / 30.4375.

Table 7: Assignments for Event and Censoring Dates for TTR_{ITT} analysis

Scenario	Date of Event or Censoring	Outcome
No post-baseline adequate assessments[1]	Date of first instillation	Censored
Received NPT prior to 3-month disease assessment	Date of first instillation	Censored
Received NPT after 3-month disease assessment prior to documented recurrence or progression	Date of last 'adequate'[1,3] assessment on or prior to starting NPT[2]	Censored
Death prior to 3-month disease assessment	Date of death	Censored

Progression at or prior to 3-month disease assessment	Date of progression[3]	Event
Death prior to documented disease recurrence or progression	Date of death	Censored
No recurrence or progression or death	Date of last 'adequate' assessment[1,3]	Censored
No recurrence or progression and the last response evaluation is 'Indeterminate'	Date of last 'adequate' assessment[1,3]	Censored
Recurrence or progression or death after 2 or more missed visits	Date of last 'adequate'[1,3] assessment prior to missed assessments	Censored
Disease recurrence or progression documented at or between scheduled visits	Date of first recurrence or progression[3]	Event

[1] An adequate assessment is defined as an assessment where the visit level response is not missing and not indeterminate.

[2] If recurrence/progression and NPT occur on the same day assume the recurrence/progression was documented first.

[3] The earliest of (i) Date of cystoscopy; or (ii) Date of biopsy, or (iii) Date of cytology.

11.4.2. Preliminary Responses Following SoC Treatment in 3-month CR Patients

Preliminary responses (approximately 3 months) following SoC treatment in 3-month CR patients who have disease recurrence or progression during the Follow-up Period are defined as the response outcomes evaluated at the first disease assessment visit after SoC. The number (%) of CR and NCR responders (as defined in EOS Evaluation of Response eCRF) will be summarized using descriptive statistics. A listing will also be provided. This analysis will be based on the FUP_{NCR} analysis set.

11.4.3. Impact of UGN-102 on Subsequent SoC treatment

The impact of UGN-102 on subsequent SoC treatment in patients who have NCR (residual disease or progression) at the 3-month Visit or disease recurrence/progression during the Follow-up Period will be analyzed to estimate the proportion of patients whose planned NMIBC treatment at baseline was downgraded following treatment with UGN-102 (e.g., from TURBT to biopsy and/or fulguration). The procedures for LG NMIBC during the study including patients who had downgraded procedures will be summarized using descriptive statistics. A listing will also be provided. This analysis will be performed on the ITT analysis set and the FUP_{NCR} analysis set.

12. PATIENT REPORTED OUTCOME (PRO) ANALYSES

The ITT analysis set will be used for analyzing PRO data unless specified differently. The PRO will be assessed using the EORTC QLQ-C30 and QLQ-NMIBC24 questionnaires

provided in Protocol Appendix 2 and Appendix 3, respectively. These questionnaires will be completed by patients at screening and subsequent visits following the SoA (Protocol Section 1.3). The baseline is defined as the last PRO assessment prior to the first instillation.

12.1. EORTC QLQ-C30

The EORTC QLQ-C30 assesses a variety of side effects that are common during cancer treatment, including fatigue, pain, nausea and vomiting, dyspnea, insomnia, appetite loss, constipation and diarrhea, mood, cognition, physical functioning, role functioning (daily activities), and a global health status scale. The recall period for the items is the past week.

The EORTC QLQ-C30 is composed of both multi-item scales and single-item measures. These include 5 functional scales, 3 symptom scales, a global health status / quality of life (QoL) scale, and 6 single items as detailed in the table below. No item occurs in more than one scale.

Table 8: EORTC QLQ-C30 Scoring

Functional Scales (15 Questions)	**Symptom Scales (7 Questions)**	**Symptom Single Items (6 Questions)**	**Global QoL (2 Questions)**
Physical (Items 1-5)	Fatigue (Items 10, 12, 18)	Constipation (Item 16)	Global QoL (Items 29, 30)
Role (Items 6, 7)	Pain (Items 9, 19)	Diarrhea (Item 17)	
Cognitive (Items 20, 25)	Nausea/Vomiting (Items 14, 15)	Insomnia (Item 11)	
Emotional (Items 21-24)		Dyspnea (Item 8)	
Social (Items 26, 27)		Appetite loss (Item 13)	
		Financial difficulties (Item 28)	

All of the scales and single-item measures range in score from 0 to 100. A high scale score represents a higher response level. Thus, a high score for a functional scale represents a high/healthy level of functioning; a high score for the global health status/QoL represents a high QoL, but a high score for a symptom scale/item represents a high level of symptomatology/problems.

Patients will be asked to complete the questionnaire as completely and accurately as possible. Missing data will be handled as recommended in the EORTC scoring manual:

Scales scores: if at least 50% of the items have been answered, the scale scores will be calculated according to the standard equations given on the manual and described below (any items with missing values will be ignored). In other cases, scores will be set to missing.

Single-item measures: the score will be set to missing.

For all scales, the Raw Score (RS), is the mean of the non-missing component items:

$$R = (I_1 + I_2 + \cdots + I_n)/n$$

For **functional scales**:

$$score = \left\{1 - \frac{(RS - 1)}{range}\right\} \times 100$$

For **symptom scales/items and global QoL**:

$$score = \{(RS - 1)/range\} \times 100$$

Range is defined as the difference between the maximum possible value of RS and the minimum possible value, e.g., most items are quoted from 1 to 4, and the range is equal to 3. The exceptions are the items contributing to the global QoL, which are 7-point questions and the range is equal to 6.

Summary statistics of the observed values will be provided for the functional/symptom scales, the global QoL, and the single items at baseline, each planned visit, and worst-case post-baseline (includes both scheduled and unscheduled assessments). The change from baseline will also be summarized with summary statistics at each above-mentioned timepoints. A listing of the above-mentioned score values will be produced for each patient.

12.2. EORTC QLQ-NMIBC24

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).

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

All scoring information specific to the QLQ-NMIBC24 is presented in the following table. 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.

For the functional scale and items, a high raw score represents a high/healthy level of functioning. Conversely, for the symptom scales and items, a high raw score represents a high level of symptomatology/problems. A high standardized score represents a worsened level of functioning and symptoms for all scales and items.

Table 9: Scoring 11 Domains of QLQ-NMIBC24

	Scale	Number of Items (n)	Item Range*	Item Numbers (I_1, I_2,…,I_n)
Symptom scales / items				
Urinary Symptoms	US	7	3	31- 37
Malaise	MAL	2	3	38, 39
Intravesical treatment issues	InV	1	3	40
Future worries	FW	4	3	41 - 44
Bloating and flatulence	BAF	2	3	45, 46
Male sexual problems [a]	SXmen	2	3	49, 50
Sexual intimacy [a]	SXI	1	3	51
Risk of contaminating partner [a]	SXCP	1	3	52
Female sexual problems [a]	SXfem	1	3	54
Functional scales / items				
Sexual function	SX	2	3	47, 48
Sexual enjoyment [a]	SXEN	1	3	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 49 to 54 are conditional questions and must only be scored if these are applicable to the patient.

For all scales, the Raw Score (RS), is the mean of the non-missing component items:

$$R = (I_1 + I_2 + \cdots + I_n)/n$$

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

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

Patients will be asked to complete the questionnaire as completely and accurately as possible. Missing data will be handled as recommended in the EORTC scoring manual:

Scales scores: if at least 50% of the items have been answered, the scale scores will be calculated according to the standard equations given on the manual (any items with missing values will be ignored). In other cases, scores will be set to missing.

Single-item measures: the score will be set to missing.

Summary statistics of the observed values will be provided for 11 domain scores at

baseline, each planned visit, and worst-case post-baseline (includes both scheduled and unscheduled assessments). The change from baseline for 11 domain scores will also be summarized with summary statistics at each above mentioned timepoints. A listing of the 11 domain scores values will be produced for each patient.

13. SAFETY ANALYSES

All safety analyses will be based on the safety analysis set. All tables and listings will display the analysis population as 'Safety Analysis Set'.

13.1. Extent of Exposure of UGN-102

Treatment exposure of UGN-102 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 Safety analysis set.

Descriptive statistics of UGN-102 dose (mg) instilled at week 1 through week 6 will be presented.

A data listing of treatment exposure will be presented.

13.2. Adverse Events

AEs will be coded according to the latest version of the MedDRA. AEs will be graded according to the Common Terminology Criteria for Adverse Events (CTCAE) Version 5.0.

All AEs (including local and systemic reactions not meeting the criteria for SAEs) and SAEs will be collected from the signing of the ICF until the 6-month Visit. After the 6-month Visit, all SAEs (regardless of causality) and non-serious AEs assessed as related to study treatment or study procedures should be collected until the EOS Visit at the time points specified in the Protocol SoA.

A treatment-emergent adverse event (TEAE) is defined as an AE that occurs on or after the day of the first instillation of UGN-102; or a pre-treatment AE that worsens during the study.

An overall summary with number and percentage of patients with all AEs, SAEs, TEAEs, different TEAE categories, like Grade $\geq$ 3, treatment-related, procedure related, serious, serious related to treatment or procedure, leading to treatment or study discontinuation, fatal, and AESIs will be provided.

Separate summaries will be provided for study treatment related TEAEs and procedure

related TEAEs. A study treatment-related AE is defined as an AE for which the investigator classifies the relationship to study treatment (UGN-102) as "Related". A study procedure-related AE is defined as an AE for which the investigator classifies the relationship to study procedure (e.g., cystoscopy, biopsy, MRI, pyelogram, CT urogram, catheter insertion, cytology) as "Related'. A worst-case scenario approach will be taken to handle missing relatedness data, i.e., the summary table will include events with the relationship to study treatment as 'Yes' if missing. The same approach for missing data will be applied to study procedure-related SAEs.

AEs will be summarized by number and percentage of patients having at least one AE, having at least one AE in each primary system organ class and for each PT using MedDRA coding. A patient with multiple occurrences of an AE will be counted only once in the respective AE category. A patient with multiple CTCAE grades for the same PT will be summarized under the maximum CTCAE grade recorded for the event.

In AE summaries, the primary system organ class will be presented by descending order of frequency (alphabetically for ties), then PTs will be sorted by descending order of frequency within each system organ class (alphabetically for ties). The sort order for the PT will be based on their frequency.

The overall observation period will be divided into three mutually exclusive time periods:

1. Pre-treatment period: from day of patient's informed consent to the day before administration of first treatment
2. Up to 3-months: from day of first treatment administration to 3-month visit
3. Post 3-months: after the 3-month visit until the end of the study.

Some key TEAE tables will be summarized by the following time periods: overall (anytime during the study from the treatment start date), up to 3-month and post 3-month. This will help to distinguish the AE profile of UGN-102 from SoC procedures as patients who are NCR at the 3-Month visit will undergo SoC. The calculation of percentages for the post 3-months period will be based on the number of patients who completed/discontinued or ongoing in the study after the 3-month visit.

Summaries based on the following subgroups will be provided for all TEAEs by PT:

- Age: < 65 years vs. ≥ 65 years
- Sex: male vs. female

All AEs will be listed.

13.3. Deaths and Serious Adverse Events

In the event that a patient has withdrawn consent, no data after the withdrawal of consent date from this patient including death is supposed to appear in the database, which should be part of the data cleaning process. All deaths occurring any time from the time of informed consent to the clinical cut-off date will be listed displaying the primary cause of death (Adverse event, related to cancer, or other).

All serious TEAEs will be tabulated based on the number and percentage of subjects who

experienced the event. Separate summaries will also be provided for study treatment related serious TEAEs and procedure related serious TEAEs. The summary tables will be displayed in descending order of incidence by system organ class (SOC) and PT.

A study treatment-related SAE is defined as an SAE for which the investigator classifies the relationship to study treatment (UGN-102) as "Related". A study procedure-related SAE is defined as an SAE for which the investigator classifies the relationship to study procedure (e.g., cystoscopy, biopsy, MRI, pyelogram, CT urogram, catheter insertion, cytology) as "Related'. A worst-case scenario approach will be taken to handle missing data, i.e., the summary table will include events with the relationship to study treatment as 'Yes' if missing. The same approach for missing data will be applied to study procedure-related SAEs.

A separate listing for SAEs will be generated.

13.4. Adverse Events Leading to Discontinuation of Study Treatment and/or Withdrawal from the Study

The following categories of TEAEs will be summarized separately in descending order of incidence by system organ class and PT and separate supportive listings will be generated with patient level details for those patients:

- TEAEs Leading to Treatment Discontinuation
- TEAEs Leading to Study Discontinuation

13.5. Adverse Events of Special Interest

An AESI is an AE of potential scientific and medical concern specific to UGN-102. AESIs may be identified through manual clinical/safety review of all AEs by system organ class and PT prior to the database lock.

The following categories of AE of special interest will be summarized as permitted by data:

- Allergic reactions
- Bone marrow suppression
- Genitourinary infections
- Lower urinary tract symptoms
- Voiding interruption due to urethral/penile edema (unrelated to prostatic hypertrophy)

Summaries of the number and percentage of patients with these events will be provided by category 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. These summaries will be provided for each type of AESI separately.

AESIs will be flagged in the general AE listing.

13.6. Clinical Laboratory Evaluations

Samples for hematology, serum chemistry and urinalysis will be collected at Screening and subsequent visits following the Protocol SoA (Section 1.3). Samples will be tested at the central laboratory. Additional tests may be part of a clinical site's local laboratory's standard panel and therefore reported along with these specified tests. Data from all sources (central and local laboratories) will be combined. When results from both central and local laboratories are available, central laboratory results will supersede the local laboratory results.

The estimated Glomerular Filtration Rate (eGFR) was calculated based on serum creatinine results from central labs using the MDRD 4-variable GFR equation:

GFR in mL/min per 1.73 m^2 = 175 x $SerumCr^{-1.154}$ x $age^{-0.203}$ x 1.212 (if patient is black) x 0.742 (if female)

In order to harmonize the results across central and local labs, the eGFR will be re-calculated for results from local labs using the formula above.

Hematology and serum chemistry results will be standardized to both conventional units and SI units and will be stored in the datasets. Summaries of laboratory parameters will include hematology and serum chemistry using SI units. Baseline will be defined as the latest non-missing assessment prior to the first instillation of UGN-102, i.e., prior to Day 1. Changes in hematology and serum chemistry variables between baseline and each subsequent assessment will be calculated. For values recorded with a leading greater than or less than ('>,' '<') symbol, the reported numeric value will be used for analysis and the value with the symbol will be included in the listings, unless otherwise specified. For example, a value of <0.01 will be analyzed as 0.01 and listed as <0.01.

The summary of observed values of the laboratory parameters will be presented at baseline, each planned visit and worst-case post-baseline (includes both scheduled and unscheduled assessments) as time points. Worst-case values 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 (Appendix 2). These summaries will display the number and percentage of patients 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. For laboratory tests that are graded for both low and high values, summaries will be done separately and labelled 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 patient has a decrease to low and an increase to high during the same time interval, then the patient is counted in both the "Decrease to Low" categories and the "Increase to High" categories.

Listing of patients with laboratory values outside the laboratory normal ranges will be generated. Complete listings of hematology and chemistry laboratory values will be generated. Scatter plots of baseline vs worst post-baseline values for ALT, AST, and TBL will also be generated.

No summary table of urinalysis will be produced. Urinalysis laboratory values with change from baseline will be listed.

13.6.1. Potentially Clinically Significant (PCS) 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 the following table.

Table 10: PCS Laboratory Criteria

Laboratory Parameter	SI Unit	Lower Limit	Upper Limit
Chemistry			
Creatinine	umol/L		>194
Sodium	mmol/L	≤130	>150
Potassium	mmol/L	<3.0	>5.5
Total bilirubin	umol/L		>1.5 × ULN
ALT	U/L		>3 × ULN
AST	U/L		>3 × ULN
GGT	U/L		>2.5 × ULN
Hematology			
Hemoglobin	g/L	<0.8 × LLN and >20% decrease from baseline	>1.3 × ULN and >30% increase from baseline
Leukocytes	$\times 10^9$/L	≤ 2.8	≥ 16.0
Lymphocytes	$\times 10^9$/L	<0.5	>20
Neutrophils	$\times 10^9$/L	<1.0	
Platelets	$\times 10^9$/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

13.6.2. Potentially Clinically Significant Abnormal Liver Function Tests

Patients meeting criteria for abnormal liver function tests (LFTs) by time point will be listed. Threshold values of interest for LFTs are provided in the following table.

Table 11: 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
Hy's Law met	ALT or AST >3×ULN and TBL >2×ULN and ALP <2×ULN

ALP = alkaline phosphatase; 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.

13.7. Vital Signs

Systolic blood pressure (SBP), diastolic blood pressure (DBP), pulse rate, respiration rate, and temperature, will be measured at Screening and subsequent visits following the Protocol SoA (Section 1.3).

The summary of observed values of the vital sign parameters will be presented at baseline, each planned visit, and worst-case post-baseline (includes both scheduled and unscheduled assessments) as time points. Worst-case values will be presented based on worst minimum and worst maximum values.

A listing of all vital sign assessments along with the change from baseline values will be produced.

13.7.1. Potentially Clinically Significant (PCS) 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 baseline, planned visit, and worst-case post-baseline (includes both scheduled and unscheduled assessments) will be provided. For computing percentages, the denominator will be the number of patients with a post-baseline value for the specific vital sign parameter and the respective time point. A listing of patients with PCS vital signs results will also be generated.

The PCS criteria are shown in the following table.

Table 12: 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

13.8. General Physical Examination

General physical examinations as described in Section 8.2.1 of the Protocol, will be performed at the screening visit only. Abnormal results (clinically significant or not clinically significant) will be summarized and listed by body system.

13.9. Urology-Oriented Examination

Urology-oriented examinations as described in Section 8.2.1 of the Protocol, will be performed at screening and subsequent visits following the Protocol SoA (Section 1.3).

Abnormal results (clinically significant or not clinically significant) will be summarized for each body system by planned visits, and worst-case post-baseline (includes both scheduled and unscheduled assessments) as time points.

A listing of patients with abnormal urology examination findings either clinically significant or not clinically significant will be provided.

14. PHARMACOKINETIC ANALYSES

Not applicable.

15. PHARMACODYNAMIC AND BIOMARKER ANALYS

Not applicable.

16. PHARMACOKINETIC/PHARMACODYNAMIC ANALYSES

Not applicable.

17. PHARMACOGENETIC DATA ANALYSES

Not applicable.

18. REFERENCES

Babjuk M, Burger M, Compérat EM, et al. European Association of Urology guidelines on nonmuscle-invasive bladder cancer (TaT1 and carcinoma in situ) - 2019 update. Eur Urol. 2019;76(5):639-657.

Brookmeyer R., Crowley J. A Confidence Interval for the Median Survival Time. *Biometrics.*Vol. 38, 1982:29-41.

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

19. APPENDICES

Appendix 1 Imputation Rules for Missing or Partial Dates

The partial date imputation will follow Analysis Data Model (ADaM) conventions. The ADaM approach is to populate the numeric date variables with the imputed date and add a flag variable to the dataset that indicates the level of imputation. The flag variable can contain the values: blank, 'D,' 'M,' 'Y.'

blank: indicates that no imputation was done

D='Day': indicates that the day portion of the date is imputed

M='Month': indicates that the month and day portions of the date are imputed

Y='Year': indicates that the entire date (year, month, and day) is imputed.

Details on imputing partial dates for specific datasets are outlined below.

Table 13: Imputation Rules of Adverse Event (AE) Start and End Dates

Dataset	Date	Missing Element	Rule
AE	Start Date	day, month, and year	• No Imputation for completely missing dates
		year	• No Imputation if year is missing
		day, month	• If year of AE start date = year of study treatment start date, then set AE start date = Study treatment start date • Else set AE start date = 01JANYYYY • If AE end date is not missing and imputed AE start date > AE end date, then imputed AE start date should be set to AE end date.
		day	• If month and year of start date = month and year of treatment start date, then set AE start date = study treatment start date. • Else set start date = 01MONYYYY. • If AE end date is not missing and imputed AE start date > AE end date, then imputed AE start date should be set to AE end date.
	End Date		• No imputation for partial end dates will be performed

Table 14: Imputation Rules of Prior or Concomitant Medication (CM) Start and End Dates

Dataset	Date	Missing Element	Rule
CM	Start Date	day, month,	• No Imputation for completely missing dates

		and year	
		year	• No Imputation if year is missing
		day, month	• If year of CM start date = year of study treatment start date, then set CM start date = Study treatment start date • Else set CM start date = 01JANYYYY • If imputed CM start date > CM end date (complete or imputed), then imputed CM start date should be set to CM end date.
		day	• If month and year of start date = month and year of treatment start date, then set CM start date = study treatment start date. • Else set start date = 01MONYYYY. • If imputed CM start date > CM end date (complete or imputed), then imputed CM start date should be set to CM end date.
	End Date	year	• No imputation for partial end dates will be performed
		day, month	• Min (Last visit date, 31DECYYYY, Date of Death) • If imputed CM start date > CM end date (complete or imputed), then imputed CM start date should be set to CM end date.
		day	• Min (Last visit date, last day of the month, death day) • If imputed CM start date > CM end date (complete or imputed), then imputed CM start date should be set to CM end date.

Table 15: Imputation Rules of Urothelial Carcinoma (UC) Related Medical History (MH) Start and End Dates

Dataset	Date	Missing Element	Rule (Only impute missing dates for MH related to UC)
MH	Start Date	day, month, and year	• No Imputation for completely missing dates
		year	• No Imputation if year is missing
		day, month	• If year of MH start date = year of start date of the current urothelial carcinoma diagnosis, then set MH start date = min (01JANYYYY, Start date of the current urothelial carcinoma diagnosis -1). • Else if the year of MH start date < year ofstart date of the current urothelial carcinoma diagnosis, then set MH start date = 01JANYYYY. • If imputed MH start date > MH end date (complete or imputed), then imputed MH start date should be set to MH end date.

		day	• If month and year of start date = month and year ofstart date of the current urothelial carcinoma diagnosis, then set MH start date = Start date of the current urothelial carcinoma diagnosis -1. • Else set start date = 01MONYYYY. • If imputed MH start date > MH end date (complete or imputed), then imputed MH start date should be set to MH end date.
	End Date	year	• No imputation for partial end dates will be performed
		day, month	• Min (Start date of the current urothelial carcinoma diagnosis, Last visit date, 31DECYYYY, Date of Death) • If imputed MH start date > MH end date (complete or imputed), then imputed MH start date should be set to MH end date.
		day	• Min (Start date of the current urothelial carcinoma diagnosis, Last visit date, last day of the month, death day) • If imputed MH start date > MH end date (complete or imputed), then imputed MH start date should be set to MH end date.

Table 16: Imputation Rules of Prior or Concomitant Surgical Procedures (PR) Start Date

Dataset	Date	Missing Element	Rule
PR	Start Date	day, month, and year	• No Imputation for completely missing dates
		year	• No Imputation if year is missing
		day, month	• If the surgical procedure was performed for a prior history of UC or MH, then do: o If the year of PR start date = year of inform consent date, then set PR start date = min (01JANYYYY, Inform consent date -1). o Else if the year of PR start date < year of inform consent date, then set PR start date = 01JANYYYY. • If the surgical procedure was NOT performed for a prior history of UC or MH, then do: o If year of PR start date = year of study treatment start date, then set PR start date = Study treatment start date -1. o Else set CM start date = 01JANYYYY
		day	• If month and year of PR start date = month and year of treatment start date, then set PR start date = study treatment start date -1. • Else set start date = 01MONYYYY.

Appendix 2 Laboratory CTCAE Grade Version 5.0 Criteria

Table 17: Laboratory CTCAE Grade Version 5.0 Criteria

Lab Parameter	SI Unit	CTCAE Grade v5.0				
		Grade 0	Grade 1	Grade 2	Grade 3	Grade 4
Hemoglobin decreased (Anemia)	g/L	≥LLN	100 – <LLN	80 – <100	< 80	
Hemoglobin increased	g/L	≤ULN	>ULN – 20 + ULN	>20 + ULN – 40 + ULN	>40 + ULN	
Albumin (Hypoalbuminemia)	g/L	≥LLN	30 - <LLN	20 - <30	<20	
Alkaline phosphatase increased	IU/L	≤ULN	BASE Normal[a]: >ULN – 2.5 × ULN BASE Abnormal: 2 × BASE – 2.5 × BASE	BASE Normal[a]: >2.5 × ULN – 5.0 × ULN BASE Abnormal: >2.5 × BASE – 5 × BASE	BASE Normal[a]: >5.0 × ULN – 20.0 × ULN BASE Abnormal: >5 × BASE – 20 × BASE	BASE Normal[a]: >20.0 × ULN BASE Abnormal: >20 × BASE
Alanine aminotransferase increased	U/L	≤ULN	BASE Normal[a]: >ULN – 3 × ULN BASE Abnormal: 1.5 × BASE – 3 × BASE	BASE Normal[a]: >3 × ULN – 5 × ULN BASE Abnormal: >3 × BASE – 5 × BASE	BASE Normal[a]: >5 × ULN – 20 × ULN BASE Abnormal: >5 × BASE – 20 × BASE	BASE Normal[a]: >20 × ULN BASE Abnormal: >20 × BASE
Aspartate aminotransferase increased	U/L	≤ULN	BASE Normal[a]: >ULN – 3 × ULN BASE Abnormal: 1.5 × BASE – 3 × BASE	BASE Normal[a]: >3 × ULN – 5 × ULN BASE Abnormal: >3 × BASE – 5 × BASE	BASE Normal[a]: >5 × ULN – 20 × ULN BASE Abnormal: >5 × BASE – 20 × BASE	BASE Normal[a]: >20 × ULN BASE Abnormal: >20 × BASE
Blood bilirubin increased	umol/L	≤ULN	BASE Normal[a]: >ULN – 1.5 × ULN BASE Abnormal: >BASE – 1.5 × BASE	BASE Normal[a]: >1.5 × ULN – 3 × ULN BASE Abnormal: >1.5 × BASE – 3 × BASE	BASE Normal[a]: >3 × ULN – 10 × ULN BASE Abnormal: >3 × BASE – 10 × BASE	BASE Normal[a]: >10 × ULN BASE Abnormal: >10 × BASE
Creatinine increased	umol/L	≤ULN	>ULN – 1.5 × ULN	>1.5 × ULN – 3 × ULN or >1.5 × BASE – 3 × BASE	>3 × ULN – 6 × ULN or >3 × BASE	>6 × ULN
Gamma-glutamyl transferase increased	U/L	≤ULN	BASE Normal[a]: >ULN – 2.5 × ULN BASE Abnormal: 2 × BASE – 2.5 × BASE	BASE Normal[a]: >2.5 × ULN – 5 × ULN BASE Abnormal: >2.5 × BASE – 5 × BASE	BASE Normal[a]: >5 × ULN – 20 × ULN BASE Abnormal: >5 × BASE – 20 × BASE	BASE Normal[a]: > 20 × ULN BASE Abnormal: >20 × BASE
Eosinophils increased (Eosinophilia)	$\times 10^9$/L	≤ULN	BASE Normal[a]: >ULN BASE Abnormal: >ULN and >BASE			
Lymphocyte count decreased	$\times 10^9$/L	≥LLN	0.8 – < LLN	0.5 – <0.8	0.2 – <0.5	<0.2
Lymphocyte count increased	$\times 10^9$/L	≤4		>4 - 20	>20	

Neutrophil count decreased	$\times 10^9$/L	≥LLN	1.5 – <LLN	1.0 – <1.5	0.5 – <1.0	<0.5
Platelet count decreased	$\times 10^9$/L	≥LLN	75 – <LLN	50 – <75	25 – <50	<25
White blood cell decreased	$\times 10^9$/L	≥LLN	3.0 – <LLN	2.0 – <3.0	1.0 – <2.0	<1.0
White blood cell increased (Leukocytosis)	$\times 10^9$/L	≤100			>100	
Hyperkalemia (Potassium increased)	mmol/L	≤ULN	>ULN – 5.5	>5.5 – 6.0	>6.0 – 7.0	>7.0
Hypokalemia (Potassium decreased)	mmol/L	≥LLN		3.0 – <LLN	2.5 – <3.0	<2.5
Hypernatremia (Sodium increased)	mmol/L	≤ULN	>ULN – 150	>150 – 155	>155 – 160	>160
Hyponatremia (Sodium decreased)	mmol/L	≥LLN	130 – <LLN	125 – <130	120 – <125	<120
Notes: BASE = baseline; LLN = lower limit of normal; ULN = upper limit of normal a. Also apply for baseline values, or when BASE is missing.						