Statistical Analysis Plan: BL006

Study Title:	A Randomized, Controlled, Open-label Study of the Efficacy, Durability, and Safety of UGN-102 With or Without TURBT in Patients with Low Grade Intermediate Risk Non-Muscle Invasive Bladder Cancer (LG IR-NMIBC) (ATLAS)
Study Number:	BL006
Study Phase:	3
Sponsor:	UroGen Pharma Ltd.
Version	Final 1.0
Date:	17 Feb 2023

Confidentiality Statement

SIGNATURE PAGE

PREPARED BY:

Author
Revathy Duvedi
Senior Statistical Consultant
UroGen Pharma, Ltd.

DocuSigned by:
Revathy Duvedi
52C5D85A88D54BC...

Signature Date: 2/17/2023

Reviewer
Thalis Tsaklas
Senior Biostatistician Consultant
ClinChoice, Inc.

DocuSigned by:
Thalis Tsaklas
3409CF600F09475...

Signature Date: 2/17/2023

APPROVED BY:

Approver
Soumi Lahiri, PhD
Senior Director, Biostatistics
UroGen Pharma, Ltd.

DocuSigned by:
Soumi Lahiri
83A07D4A7B0C4CE...

Signature Date: 2/17/2023

Approver
Sunil Raju, MD
Vice President, Clinical Development
UroGen Pharma, Ltd.

DocuSigned by:
Sunil Raju
F10B82FA75194A5...

Signature Date: 17/02/2023

TABLE OF CONTENTS

PAGE

ABBREVIATIONS 5

1. INTRODUCTION 7

2. STUDY OBJECTIVES AND ENDPOINTS 7

3. STUDY DESIGN 8
   - 3.1. Statistical Hypothesis 10
   - 3.2. Sample Size Justification 10
   - 3.3. Randomization and Blinding 10

4. PLANNED ANALYSES 11
   - 4.1. Interim Analyses 11
   - 4.2. Final Analysis 11

5. ANALYSIS POPULATIONS 11

6. TREATMENT COMPARISONS 12

7. CHANGES TO THE PLANNED ANALYSES 12

8. GENERAL CONSIDERATIONS FOR DATA ANALYSES 12
   - 8.1. Multicenter Studies 13
   - 8.2. Other Strata and Covariates 13
   - 8.3. Multiple Comparisons and Multiplicity 13

9. DATA HANDLING CONVENTIONS 13
   - 9.1. Premature Withdrawal and Missing Data 14
   - 9.2. Derived and Transformed Data 14
     - 9.2.1. Reference Date 14
     - 9.2.2. Study Day for Safety Measures 15
     - 9.2.3. Study Day for Efficacy 15
     - 9.2.4. Calculation of Durations 15
     - 9.2.5. Imputation of Partial Dates 15
     - 9.2.6. Baseline Definition 16
     - 9.2.7. Change from Baseline 16
     - 9.2.8. Analysis Visit Window 16
     - 9.2.9. Multiple Assessments 17
     - 9.2.10. Actual Treatment 18

10. PATIENT DISPOSITION, DEMOGRAPHICS AND OTHER BASELINE CHARACTERISTICS 18
    - 10.1. Disposition of Patients 18
    - 10.2. Protocol Deviations 19
    - 10.3. Demographic and Baseline Characteristics 19

10.4. General Medical History and Urothelial Carcinoma Related Medical History (UCMH) 20
10.5. Prior and Concomitant Medications 20
10.6. Prior and Concomitant Surgical Procedures 21

11. EFFICACY ANALYSES 21
11.1. Primary Efficacy Analyses 21
11.1.1. Disease-Free Survival (DFS) 21
11.2. Secondary Efficacy Analyses 25
11.2.1. Time to Recurrence (TTR) 25
11.2.2. Complete Response Rate (CRR) at 3-Month Visit 26
11.2.3. Visit Level Complete Response Rate (CRR) 27
11.2.4. Duration of Response (DOR) 29
11.2.5. Avoidance of Surgery (TURBT) for Treatment of LG IR-NMIBC 29

12. HEALTH RELATED QUALITY OF LIFE (HRQOL) ANALYSES 30

13. HEALTH RESOURCE UTILIZATION 31

14. SAFETY ANALYSES 31
14.1. Extent of Exposure of UGN-102 32
14.2. Adverse Events 32
14.3. Deaths and Serious Adverse Events 33
14.4. Adverse Events Leading to Discontinuation of Study Treatment and/or Withdrawal from the Study 34
14.5. Adverse Events of Special Interest 34
14.6. Clinical Laboratory Evaluations 34
14.6.1. Potentially Clinically Significant (PCS) Laboratory Values 35
14.6.2. Potentially Clinically Significant Abnormal Liver Function Tests 36
14.7. Vital Signs 37
14.7.1. Potentially Clinically Significant (PCS) Vital Sign Abnormalities 37
14.8. General Physical Examination 38
14.9. Urology-Oriented Examination 38

15. PHARMACOKINETIC ANALYSES 38

16. PHARMACODYNAMIC AND BIOMARKER ANALYS 38

17. PHARMACOKINETIC/PHARMACODYNAMIC ANALYSES 38

18. PHARMACOGENETIC DATA ANALYSES 38

19. REFERENCES 38

20. APPENDICES 39
Appendix 1 Imputation Rules for Missing or Partial Dates 39
Appendix 2 Laboratory CTCAE Grade Version 5.0 Criteria 42
Appendix 3 Data Display Specifications 43

ABBREVIATIONS

AE	adverse event
AESI	adverse event of special interest
ALT	alanine aminotransferase
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
DBP	diastolic blood pressure
DFS	disease-free survival
DOR	duration of response
eCRF	electronic case report form
EOS	end of study
HR	hazard ratio
HRQoL	health related quality of life
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
QLQ-NMIBC24	24-item quality of life questionnaire for patients with NMIBC
RD	Recurrence disease
SAE	serious adverse event
SAP	statistical analysis plan
SAS	Statistical analysis system
SBP	systolic blood pressure
SOA	schedule of activities
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
ULN	upper limit of the normal range
WHO	World health organization

1. INTRODUCTION

This statistical analysis plan (SAP) describes all planned analyses for the Clinical Study Report (CSR) of study BL006 (ATLAS), a phase 3, randomized, controlled, open-label study of UGN-102 with or without transurethral resection of bladder tumors (TURBT) in patients with Low Grade Intermediate Risk Non-Muscle Invasive Bladder Cancer (LG IR-NMIBC).

This SAP is written with due consideration of the recommendations outlined in the most recent International Council on Harmonisation (ICH) E9 Guideline entitled Statistical Principles for Clinical Trials and the most recent ICH E3 Guideline entitled Structure and Content of Clinical Study Reports (CSR). The SAP is based on the following study documents:

- Protocol, Version 2.0 dated 20 December 2021
- Case report form (CRF), Version 6.1 dated 30 June 2022

All decisions regarding final analysis, as defined in the SAP document, have been made prior to database lock and formal unblinding of the study data.

2. STUDY OBJECTIVES AND ENDPOINTS

Table 1 lists the study objectives and corresponding endpoints which are under the scope of this SAP.

Table 1: Study Objectives and Endpoints

OBJECTIVE	ENDPOINT
Primary	
To evaluate the efficacy of UGN-102 with or without TURBT versus TURBT alone with respect to disease-free survival (DFS) in patients with Low Grade Intermediate Risk Non-Muscle Invasive Bladder Cancer (LG IR-NMIBC).	Disease-free survival (DFS) is defined as the time from randomization until the earliest date of any of the following events: • Failure to be rendered free of local disease at the 3-Month assessment after the TURBT procedure. • Recurrence of low-grade disease after the 3-Month assessment (i.e., during the follow-up period). • Progression to high-grade disease. • Death due to any cause.
Secondary	
1. To evaluate the efficacy of UGN-102 with or without TURBT versus TURBT alone with respect to: a) Time to Recurrence (TTR) b) Complete Response Rate (CRR) at 3-Month disease assessment	1. The following efficacy endpoints will be evaluated: a) Time to recurrence (TTR) is defined as the time from randomization until the earliest date of recurrence of low-grade disease or progression to high-grade disease. b) Complete response rate (CRR), defined as the proportion of patients who achieved CR at the 3-Month disease assessment.

OBJECTIVE	ENDPOINT
c) Duration of Response (DOR) d) Avoidance of surgery (TURBT) for treatment of LG IR-NMIBC 2. To evaluate the safety profile of UGN-102 with or without TURBT versus TURBT alone. 3. To assess the effect of UGN-102 with or without TURBT versus TURBT alone on Patient Reported Outcomes (PROs) including disease related symptoms, functioning, and health-related quality of life (HRQoL). 4. To evaluate visit level Complete Response Rate (CRR)	c) Duration of response (DOR), defined as the time from first documented CR until the earliest date of any of the following events: • Recurrence of low-grade disease. • Progression to high-grade disease. • Death due to any cause. d) Proportion of patients requiring TURBT in each arm and average number of TURBT interventions per patient in each arm 2. The safety profile of UGN-102 and TURBT will be evaluated as assessed through standard clinical and laboratory tests (hematology and chemistry, urinalysis, physical examination, vital sign measurements, diagnostic tests, etc.) and through the collection of reports of adverse events (AEs) and serious adverse events (SAEs) including AEs of special interest. 3. Changes from baseline in HRQoL measures assessed by European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire for Non-muscle Invasive Bladder Cancer patients (EORTC-QLQ-NMIBC24). 4. Observed CRR at scheduled disease assessment timepoints, defined as the proportion of patients who had CR at 3-Month disease assessment and maintained CR up to that particular follow-up disease assessment.
Exploratory	
To explore potential differences in health resource utilization with UGN-102 vs TURBT in patients with LG NMIBC	Number of patients hospitalized for non-elective reasons, total number and length of non-elective hospitalizations

3. STUDY DESIGN

This study is a global, randomized, controlled, open-label Phase 3 study designed to assess the long-term efficacy and safety of UGN-102 (mitomycin) for intravesical solution (i.e., intravesical instillation) with or without TURBT (treatment arm) versus TURBT alone (control arm) in the treatment of patients with LG IR-NMIBC defined as 1 or 2 of the following: multiple tumors, solitary tumor > 3 cm, or recurrence (≥ 1 occurrence of LG NMIBC within 1 year of the current diagnosis at the initial Screening Visit) (Kamat, 2014). Eligible patients were treated with 6 weekly instillations of UGN-102 followed by TURBT only if needed, or with TURBT alone.

Eligible patients were randomized in a 1:1 ratio to UGN-102 with or without TURBT or TURBT alone. Randomization was stratified by the presence of previous LG NMIBC episodes within 1 year of the current diagnosis at the initial Screening Visit (yes versus no). Starting at Day 1, patients randomized to the UGN-102 group were scheduled to receive

6 weekly intravesical instillations of UGN-102 followed by TURBT only if needed and patients randomized to the TURBT alone group were scheduled to undergo TURBT followed by repeat TURBT if needed.

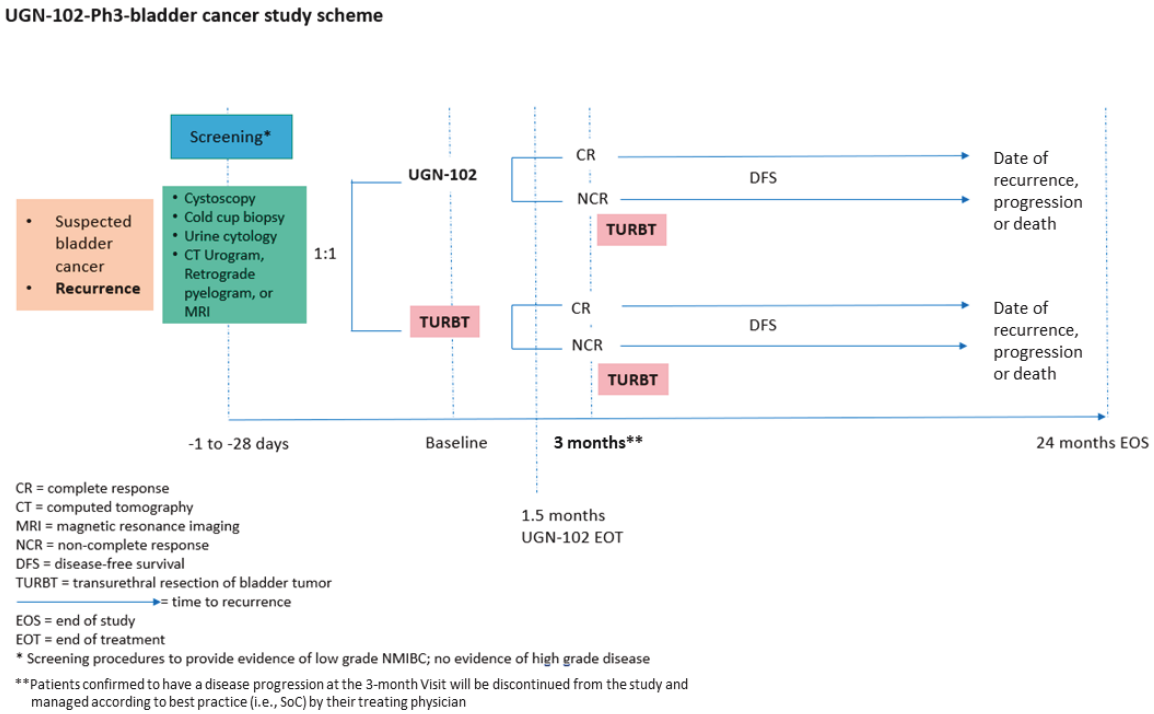
All patients (i.e., both treatment arms) were scheduled to return to the clinic at approximately 3 months from the initiation of treatment (i.e., 7 weeks ±1 week after the last weekly instillation for the UGN-102 treatment arm; and 12 weeks ±1 week after TURBT for the TURBT alone treatment arm) for their first evaluation of response. Following this initial evaluation, all patients who had a CR received no further treatment and entered the follow-up period of the study. Patients who had a non-complete response (NCR), residual disease, in either treatment arm had TURBT of any remaining lesions and then entered the follow-up period of the study. Note: If the Investigator determined that a formal TURBT in the operating room was unnecessary following cystoscopy and biopsy of remaining lesions (e.g., residual disease can be addressed by in office biopsy and/or fulguration), such an alternative procedure was permitted based on the Investigator's medical judgment.

Response was determined based on visual evaluation by cystoscopy (white light) (appearance, number, and size of any remaining lesions), histopathology of any remaining lesions, and voiding urine cytology. Details of the evaluation methods are available in Appendix 1 (Guidance on Evaluation of Response) of the protocol.

The original plan was to follow patients up to 24 months or until disease recurrence, disease progression, or death was documented, whichever occurred first. But the study was closed early for business reasons after all patients completed their 15-Month visit (i.e. 12 months follow-up after primary disease evaluation at 3-Month) or until disease recurrence, disease progression, or death was documented, whichever occurred first.

The study design is depicted in Figure 1 below. Schedules of activities from Screening to end of study (EOS) are available in protocol Section 1.3 and Section 1.4 for patients in the UGN-102 ± TURBT and TURBT alone treatment arms, respectively.

Figure 1: Study Schema



3.1. Statistical Hypothesis

UroGen Pharma Ltd. (Sponsor) closed enrollment in Study BL006 on 10 Nov 2021 in order to pursue an alternative development strategy for UGN-102 in patients with LG IR NMIBC. Although patients who were already randomized were permitted to continue their study participation until the EOS, the study is not powered to perform any hypothesis testing. Therefore, all analyses will be descriptive in nature.

3.2. Sample Size Justification

As of 10 Nov 2021, 282 patients were randomized in this study.

3.3. Randomization and Blinding

Patients were randomly assigned in a 1:1 ratio to UGN-102 with or without TURBT or TURBT alone by using Interactive Response Technology (IRT) system. Randomization was stratified by the presence of previous LG NMIBC episodes within 1 year of the current diagnosis at the initial Screening Visit (yes versus no). The randomization schedule was generated by the ClinChoice randomization statistician, using SAS version 9.4 that incorporates a standard procedure for generating randomization numbers. One randomization list, using block randomization method was generated, and all centers are using the same list to minimize any imbalance in the number of patients assigned to each treatment arm.

Open-label studies, particularly randomized ones, require a rigorous approach to protect trial and data integrity. Per ICH E9 and E10 Principles, blinding or masking is intended to

limit the occurrence of conscious and unconscious bias in the conduct and interpretation of a clinical trial arising from the influence that the knowledge of treatment may have on the recruitment and allocation of patients, their subsequent care, attitudes of patients to the treatments, assessment of endpoints, handling of withdrawals, exclusion of data from analysis and so on. A Trial Integrity Plan (TIP) was created to define data restrictions and data access for the randomized open-label Study BL006, sponsored by UroGen Pharma Ltd (UroGen) without compromising patient safety. Details of study blinding plan are provided in the Trial Integrity Plan version 1.1.

4. PLANNED ANALYSES

4.1. Interim Analyses

No interim analysis is planned for this study.

4.2. Final Analysis

The final analysis of this study will be performed after all patients complete their 15-Month visit (i.e., 12 months follow-up after primary disease evaluation at 3-Month) or until disease recurrence, disease progression, or death is documented, whichever comes first.

5. ANALYSIS POPULATIONS

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

Analysis Set	Description
Intent-to-Treat Analysis Set (ITT)	The ITT comprises of all randomized patients regardless of whether the treatment was administered. Patients will be analyzed according to the treatment and strata they have been assigned to during the randomization procedure. Any patient who receives a randomization number will be considered to have been randomized. Efficacy analyses will be conducted using the ITT population.
Safety Set	The Safety Set includes all patients who received any dose of UGN-102 (treatment arm) or at least one TURBT intervention (control arm). Patients will be analyzed according to the study treatment they actually received. All safety analyses will be conducted using the safety population.
Per Protocol Set (PPS)	The Per Protocol set (PPS) will include the subset of the patients in the Safety analysis set without major protocol deviations that would confound efficacy evaluation. All protocol deviations or conditions leading to exclusion from the PPS will be detailed in the protocol deviation plan. Sensitivity analyses of the primary endpoint of DFS may be performed using data from PPS if the PPS and ITT differ sufficiently.
3-Month CR Set	The 3-Month CR Analysis Set will consist of all patients from ITT set who achieved CR at 3-Month disease assessment (Study Visit 3).

	Efficacy analyses of the secondary endpoints DOR and visit level CRR will be conducted using the 3-Month CR Set.

Patients considered NCR at the 3-Month Visit could receive an alternative procedure (e.g., biopsy and/or fulguration) if a formal TURBT was unnecessary based on the Investigator's medical judgment. All NCR patients at the 3-Month Visit, irrespective of their mode of treatment, will be included in the analysis.

6. TREATMENT COMPARISONS

The study is not powered to perform any hypothesis testing as described in Section 3.1. Therefore, all analyses will be descriptive in nature. Hazard ratios (HR) will be generated for the primary endpoint to understand the trend of any clinical benefit in the treatment arm versus the control arm. No statistical conclusions will be made.

The following treatment descriptors will be used on all applicable displays:

Treatment Group		Data Display	Order of Treatment Groups
Code	Description		
1	UGN-102 ± TURBT	UGN-102 ± TURBT	1
2	TURBT Alone	TURBT Alone	2

7. CHANGES TO THE PLANNED ANALYSES

1. Definition of ITT analysis set is modified to include all randomized patient regardless of whether the treatment was administered.
2. Full Analysis Set is not needed anymore and hence has been deleted.
3. Definition for PPS is modified to be a subset of Safety analysis set.
4. All analyses under this SAP are descriptive in nature. Mixed effect model for repeated measures (MMRM) will not be performed to evaluate the two treatment arms with respect to changes in the QLQ-NMIBC24 measures.
5. Due to the change in the ITT definition, all health outcomes related analyses will be performed under Safety analysis set.
6. The overall safety observation periods have been updated to:
    - Pre-treatment period: from day of patient's informed consent to the day before administration of first treatment or TURBT,
    - Up to 3-months: from the First Dose date of UGN-102/ date of Initial TURBT Procedure to 3-Month Visit, and
    - Post 3-months: after 3-Month Visit until the end of the study.

8. GENERAL CONSIDERATIONS FOR DATA ANALYSES

Data will be listed and summarized according to Clinical Data Interchange Standards Consortium (CDISC) standards and ICH E9 Guideline. SAS software will be used to

perform all data analyses, generate tables, figures, and listings.

Unless otherwise stated, all listings will be sorted by treatment group, patient identification number (ID), 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.

Generally, summaries will be produced by treatment arms. All laboratory data will be presented using the International System of Units (SI) Units.

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

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

Prior to randomization, eligible patients are stratified by previous LG NMIBC episodes within 1 year (Yes versus No).

For analysis purposes randomized strata from IRT will be used. Strata will also be captured on the electronic case report form (eCRF) and will be listed if this differs from the classification reported at the time of randomization.

A table will be provided to summarize the number of patients within each category of the stratification variables for the ITT population.

All statistical analyses, unless otherwise noted, will be stratified by the randomization stratification.

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), 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.

All available time-to-event data will be analyzed using suitable statistical methods. For endpoints which determine the percentage of complete responders (CR), indeterminate responses will be imputed following the imputation rules described in Section 11.2.2 and Section 11.2.3.1. Patients who terminated the study early, prior to the first disease assessment at 3-Month, will be assumed to be NCRs and will be included in the denominator when calculating the percentages.

Missing data occurs when any requested data is 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 should be displayed as such.

Patients with the designation of treatment relationship for adverse events (AE) and serious adverse events (SAEs) missing will have the worst-case assumed to impute the relationship: if relationship 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 Imputation Rules for Missing or Partial Dates.

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 three reference dates:

- The reference date for age is the date of Informed Consent as age is an eligibility requirement.

- The safety reference date is the date of first instillation of UGN-102 in the UGN-102 ± TURBT arm or date of first TURBT procedure for patients in the TURBT Alone arm and will be used to calculate study day for safety measures.

- The efficacy reference date is the date of randomization and will be used to calculate study day for efficacy measures and baseline characteristics (such as time since initial diagnosis), as well as efficacy durations.

9.2.2. Study Day for Safety Measures

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

9.2.3. Study Day for Efficacy

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

9.2.4. Calculation of Durations

Durations (e.g., duration of adverse event, duration of response, 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.

9.2.5. Imputation of Partial Dates

In general, imputed partial dates will not be used to derive study day or duration (e.g., duration of adverse events). 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 to study time periods or 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.

9.2.6. Baseline Definition

For Safety analyses, baseline will be defined as the most recent non-missing value prior to the first instillation of UGN-102 for patients in UGN-102 ± TURBT arm or prior to the first TURBT for patients in TURBT alone arm, i.e., prior to Day 1.

The baseline is defined as the last HRQoL assessment on or prior to randomization for the HRQoL analysis.

9.2.7. Change from Baseline

Change from baseline will be presented for HRQoL and safety data as described in Section 12 and Section 12 respectively.

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.

9.2.8. 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:

- Health related Quality of Life (HRQoL)
- 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.

Due to the different schedule of assessments per treatment arm, the visit windows are

defined separately in the following tables.

Table 2: Analysis windows for UGN-102 ± TURBT arm

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

Table 3: Analysis windows for TURBT Alone arm

Analysis Visit	Target Day	Analysis Window (Study Day)
Baseline	< 1	< 1
Day 1	1	1
Month 3	90	2 - 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

9.2.9. Multiple Assessments

All data will be reported according to the nominal visit date for which it was reported. Analysis windows will be applied during dataset creation for the summaries of clinical hematology, chemistry, urinalysis, vital signs, urological examination, and the HRQoL 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.

9.2.10. Actual Treatment

If a patient's actual treatment is the same as the randomized treatment, then actual treatment is the randomized treatment. If a patient receives a study treatment that is different from the randomized treatment, then actual treatment is the different treatment (the treatment actually received). A listing of patients with actual treatment deviating from the randomized treatment will be generated.

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 all summaries and data listings will use treatment labels as specified in Section 6.

10.1. Disposition of Patients

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 by treatment arm using the ITT analysis set. A listing of patients excluded from analysis set along with the reasons will also be provided.

A summary of treatment status of 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 listing of UGN-102 treatment discontinuation will be generated.

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 or withdrew from the study, and primary reasons for study completion/withdrawal. Reasons for study completion or withdrawal will

be presented in the order they are displayed in the eCRF. The corresponding listing will also 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 (PDMP) for details of protocol deviations.

Patients with major protocol deviations that would confound efficacy evaluation, would be excluded from the PPS. All protocol deviations or conditions leading to exclusion from the PPS will be detailed in the protocol deviation plan and will be finalized prior to the database lock (DBL). Sensitivity analyses of the primary endpoint DFS may be performed using data from PPS if the PPS and ITT differ sufficiently.

10.3. Demographic and Baseline Characteristics

The demographic characteristics (e.g., age, sex, race, ethnicity, height, weight and BMI) will be summarized and listed. Age, height, weight and BMI 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.
The count and percentage will be computed for sex, race and ethnicity.

The following baseline prognostic factors used in the subgroup analyses will be provided by count and percentage:

- Age categories (<65, ≥ 65; < 75, ≥ 75)
- BMI category (<30 kg/m^2, ≥ 30 kg/m^2)
- Disease history (number of previous LG NMIBC episodes, previous LG NMIBC episodes within 1 year of current diagnosis)
- Smoking history (smoker, non-smoker)
- Disease burden (tumor size: ≤ 3 cm, >3 cm; tumor count: single, multiple)
- Prior TURBT to treat at least 1 prior episode of LG NMIBC (Yes, No)

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)
- Status (Smoker vs Non-Smoker)
- Duration (<10, 10-20, 21-30, >30 years)

A summary of prior LG NMIBC along with its standard of care (TURBT) will be provided. Number of prior LG NMIBC, time to recurrence between NMIBC episodes, 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 group (0,

1, 2, ≥ 3) and time since last TURBT procedure at the time of current diagnosis (≤ 365 days, >365 days). A listing of prior NMIBC diagnosis and the 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 SOC and PT using the most recent version of MedDRA Version 23.1 or later. Number and percent of patients reporting any medical history by SOC and PT for the ITT analysis set will be provided for each treatment group. A patient with multiple medical conditions will be counted once per SOC and PT. For computing percentages, the denominator will be the number of patients in the ITT analysis Set. The summary table will be sorted by descending order of frequency of SOC in the UGN-102 ± TURBT column (then alphabetically for ties), then by descending order of frequency of PT within each SOC (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 month between the date of initial diagnosis and baseline
- Number of UCMH episodes
- Staging of the UCMH episode (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, Other)
- Grade of the UCMH episode (Papillary urothelial carcinoma - low grade, Papillary urothelial neoplasm of low malignant potential, Urothelial carcinoma in situ, Papillary urothelial carcinoma - high grade, Other)
- UCMH episode treated (Yes, No)
- Treatment given for the prior episode 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 data. Analysis of prior and concomitant medications use will be performed in the following manner:

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

UGN-102 ± TURBT arm or following the first TURBT procedure for patients in TURBT alone arm. The number and percentage of patients reporting the use of concomitant medications by ATC Level 2, ATC Level 3 and preferred terms 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 for patients in UGN-102 ± TURBT arm or prior to the first TURBT for patients in TURBT alone arm.

Concomitant surgical procedures are surgical procedures that were done following the first instillation of UGN-102 for patients in UGN-102 ± TURBT arm or the first TURBT for patients in TURBT alone arm through end of study.

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

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

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 treatment arm.

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

11.1. Primary Efficacy Analyses

11.1.1. Disease-Free Survival (DFS)

DFS is the primary endpoint of this study, and it is defined as the time from randomization until the earliest date of failure to be rendered free of local disease at the 3-Month assessment after the TURBT procedure, recurrence of low-grade disease after the 3-Month assessment (i.e., during the follow-up period), progression to high-grade disease, or death due to any cause.

NCR (i.e., residual disease) at the 3-Month disease assessment in the UGN-102 ± TURBT treatment arm is not considered as a DFS event because treatment failure is not an event in a neoadjuvant setting. However, having residual disease at the 3-month disease assessment

after the TURBT procedure (i.e., in the TURBT arm) will be considered as a DFS event. Progression to HG disease any time during the study (even at the 3-month disease assessment) will be considered as a DFS event in both arms.

The date of documented recurrence or progression will be determined using the date of cystoscopy, biopsy, or cytology, whichever occurs first. The date of death (due to any cause) should be taken from the Record of Death page.

Patients without any adequate post baseline disease assessments will be censored at the date of randomization.

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). An adequate assessment is defined as an assessment where the visit level response is not missing, and not Indeterminate. A patient will be considered to have an extended lost to follow-up (i.e., missed two or more consecutive adequate assessments), if the patient did not have an adequate assessment during the time period 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 DFS status of this patient will be censored, and the censoring date will be the date of 6-month visit.

Non-Protocol Therapy (NPT)

Non-Protocol Therapy (NPT) are prohibited concomitant anti-cancer medication or surgical procedures as outlined in Section 6.8 of the protocol. These medications will be identified via clinical review of the CONMEDS 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 DFS, TTR and DOR.

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 Table 11 and Table 13 of 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 DFS 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 DFS will be censored at the date of randomization.
- If a patient does not have any post-baseline adequate assessment, DFS will be censored at the date of randomization.
- If a patient has neither recurrence or progression nor died nor started NPT, then DFS will be censored at the date of the last adequate assessment.

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

Table 4: Assignments for Event and Censoring Dates for DFS Analysis

Situation	Date of Event or Censoring	Outcome
No post-baseline adequate assessments[1]	Randomization date	Censored
Received NPT prior to 3-Month disease assessment	Randomization date	Censored
Received NPT after 3-Month disease assessment prior to documented recurrence or progression	Date of last 'adequate' assessment[1,3] 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
Residual disease at 3-Month disease assessment in TURBT Alone arm	Date of NCR[3]	Event
Death between adequate assessment visits	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 more than two 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

[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.

DFS will be analyzed based on the data observed in the ITT population, according to the treatment arm patients were randomized and previous LG NMIBC episodes within 1 year (Yes vs. No) as the strata assigned at randomization. The distribution of DFS will be

estimated using the Kaplan-Meier method. The estimated median, first, and third quartiles of DFS along with 95% CIs (Brookmeyer and Crowley, 1982) will be presented by treatment arm. A figure and a listing of DFS time will be provided.

A stratified Cox regression model will be used to estimate the HR of DFS, along with 95% CIs (using the same strata information as above). No p-value will be generated. The HR will be generated to understand the trend of any clinical benefit in the treatment arm versus the control arm. No statistical conclusion will be made.

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

11.1.1.1. Supportive Analyses for DFS

11.1.1.1.1. *Sensitivity Analysis*

The primary analysis for DFS may be repeated with data based on PPS, as described in Section 5, if the PPS and ITT differ sufficiently.

The proportionality of hazard assumption of the DFS distribution functions will be evaluated by appropriate methods.

Additionally, a sensitivity analysis for DFS will be conducted, where the recurrence or progression or death after more than two or more missed visits (i.e., after extended loss to follow-up) will be considered as event in the analysis.

11.1.1.1.2. *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. Additionally, a forest plot of estimated HRs with the 95% CI by subgroups will be presented.

Subgroups will include (but not limited to):

- Baseline age group (years): (< 75, ≥ 75)
- BMI category (kg/m^2): (< 30, ≥ 30)
- Sex: (Male, Female)
- Tumor size (cm): (≤ 3, > 3)
- Tumor count: (Single, Multiple)
- Previous LG NMIBC Episodes within 1 year: (Yes, No)
- Previous LG NMIBC Episode (Yes, No)
- Number of previous LG NMIBC episodes: (≤ 2, > 2)
- Prior TURBT (Yes, No)
- Smoking History: (Smoker, Non-smoker)

Prior TURBT value ‘Yes’ indicates at least 1 prior episode of LG NMIBC was treated with TURBT.

Smoker category includes both former and current smokers; Non-smoker category includes patients who "Never" smoked.

The HRs will be generated to understand the trend of any clinical benefit in the treatment arm versus the control arm in each of the subgroups. No statistical conclusions will be made.

11.2. Secondary Efficacy Analyses

11.2.1. Time to Recurrence (TTR)

TTR analysis will be based on the ITT analysis set. TTR is defined as the time from randomization until the earliest date of recurrence of low-grade disease or progression to high-grade disease. Death will not be considered as 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 defining event and censoring are described in the following table.

Table 5: Assignments for Event and Censoring Dates for TTR Analysis

Situation	Date of Event or Censoring	Outcome
No post-baseline adequate assessments[1]	Randomization date	Censored
Received NPT prior to 3-Month disease assessment	Randomization date	Censored
Received NPT after 3-Month disease assessment prior to documented recurrence or progression	Date of last 'adequate' assessment[1,3] on or prior to starting NPT[2]	Censored
Death prior to 3-Month disease assessment	Date of death	Censored
Progression at or prior 3-Month disease assessment	Date of progression[3]	Event
Residual disease at 3-Month disease assessment in TURBT Alone arm	Date of NCR	Event
Death between adequate assessment visits	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 more than two 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

[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.

The distribution function of TTR will be estimated using the Kaplan-Meier method. The estimated median, first, and third quartiles times to the first documented disease recurrence or progression along with 95% CIs (Brookmeyer and Crowley, 1982) will be presented by treatment arm. A listing of TTR endpoint will also be provided.

A stratified Cox regression model will be used to estimate the HR of TTR, along with 95% CIs. No p-value will be generated. The HR will be generated to understand the trend of any clinical benefit in the treatment arm versus the control arm. No statistical conclusion will be made.

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

11.2.2. Complete Response Rate (CRR) at 3-Month Visit

Complete response rate (CRR) is defined as the proportion of patients in the ITT population who achieved CR at the 3-Month disease assessment. 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 Evaluation of Response CRF page.

If the response at the 3-Month Visit is Indeterminate, then it will be imputed as below:

- If the patient required a TURBT or other alternative procedures based on the investigator's judgment (e.g., intraoperative laser resection, office fulguration, office cold cup biopsy, etc.), to treat the residual disease observed at the 3-Month evaluation, then the 3-Month response will be imputed as 'NCR.'
- Else if the patient did not have any disease assessment post 3-Month visit, then the 3-Month response will be imputed as 'NCR'.
- Otherwise, it will be imputed as a CR.

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

CRR will be presented by treatment arm along with the exact 95% CIs. A test for binomial proportions (SAS PROC FREQ with binomial option) will be used to derive the exact two-sided 95% CI for the CRR and using the Clopper-Pearson method (Clopper and Pearson, 1934). In addition, reasons for non-complete response (NCR) comprised of 'Residual Disease,' 'Progression to HG Disease,' and 'Indeterminate' will be tabulated using the number and percentages of patients.

This endpoint will also be analyzed for a subgroup of patients who had at least 1 prior episode of LG NMIBC which was treated with TURBT.

11.2.3. Visit Level Complete Response Rate (CRR)

Observed CRR 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 (i.e., 3-MONTH$_{CR}$ analysis set) by treatment arm. Patient response will be evaluated by the investigator according to the EOR in protocol Appendix 1 and will be recorded on the follow-up Evaluation of Response CRF page. Missing visit level responses will be imputed as described in Section 11.2.3.1.

Evaluation of response will be based on central laboratory results when 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.

If a patient receives any NPT during the follow-up period, the response status at all subsequent follow-up visits after the NPT start date will be considered as NCR. Patients who terminated the study early or patients who have indeterminate response (which are not confirmed by subsequent visits) will also be considered as NCR.

Observed CRR will be presented with nominal 95% exact CI (Clopper-Pearson) for each treatment arm at each scheduled timepoint (i.e., 6-, 9-, 12- and 15-months visits). The denominator to calculate the proportion will include all patients who were CR at 3-month visit.

This endpoint will also be analyzed for a subgroup of patients who had at least 1 prior episode of LG NMIBC which was treated with TURBT.

11.2.3.1. 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 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).

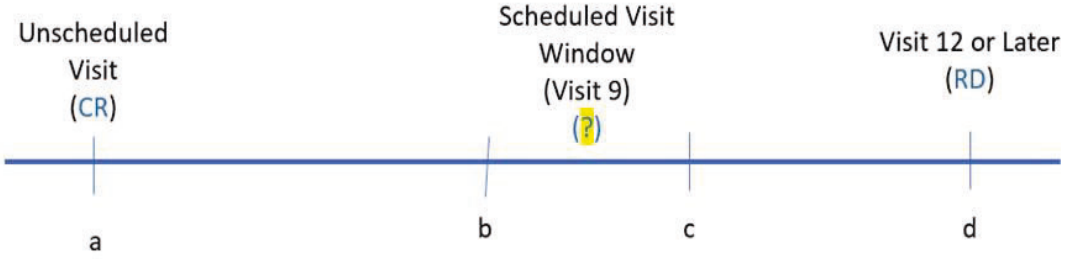
3. Otherwise, responses will not be carried forward to scheduled visits that have not occurred (Patient C; Patient E)

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	288						RD
Unscheduled	300				RD		
12-Month	360 (± 2 week)	346-374				RD	
…….							

RD=Recurrence disease; U=Unknown/Indeterminate; Red implicates 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 Visit 9 window. This patient also has a RD response on study day 346 (say), 62 days (346-284) after the end of the visit 9 window. Since the CR response is closer to the visit 9, the response will be CR.

A patient listing will be generated to display visit level responses as recorded in the CRF

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 11.2.2.

11.2.4. Duration of Response (DOR)

DOR will be summarized only for CR patients at the 3-Month visit (i.e., 3-MONTH$_{CR}$ analysis set) by treatment arm. DOR is defined as the time from first documented CR until the earliest date of recurrence of low-grade disease, progression to high-grade disease or death due to any cause. If a patient has not had an event (recurrence, progression, or death), DOR will be censored at the date of the last adequate disease assessment (i.e., assessment where visit level response is CR). The rules for defining recurrence and censoring are described in Table 4. Censoring rules will follow those of the primary DFS analysis defined in Section 11.1.1.

The distribution function of DOR will be estimated using the Kaplan-Meier method. The estimated median, first, and third quartiles of DOR along with 95% CIs (Brookmeyer and Crowley, 1982) will be presented by treatment arm. A figure and listing of DOR endpoint will also be provided.

A stratified Cox regression model will be used to estimate the HR of DOR, along with 95% CIs. No p-value will be generated. The HR will be generated to understand the trend of any clinical benefit in the treatment arm versus the control arm. No statistical conclusion will be made.

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.

Additionally, a sensitivity analysis for DOR will be conducted, where the recurrence or progression or death after more than two or more missed visits (i.e., after extended loss to follow-up) will be considered as event in the analysis.

This endpoint will also be analyzed for a subgroup of patients who had at least 1 prior episode of LG NMIBC which was treated with TURBT.

11.2.5. Avoidance of Surgery (TURBT) for Treatment of LG IR-NMIBC

This analysis will be performed using the Safety analysis set. Patients confirmed to have had a non-complete response (NCR) in either treatment arm will undergo TURBT of any remaining lesions and will then enter the follow-up period of the study. If the Investigator determines that a formal TURBT in the operating room is unnecessary following cystoscopy and biopsy of remaining lesions, residual disease can be addressed by the following in office procedures based on the Investigator's medical judgment:

- Intraoperative Laser Resection

- Office Fulguration (Ablation, Cauterization)
- Office Cold Cup Biopsy
- Other

Proportion of patients requiring TURBT (total and after 3-Month disease evaluation) in each arm will be summarized along with other in office alternative procedures. Number of TURBT intervention per patient will also be summarized using appropriate descriptive statistics.

12. HEALTH RELATED QUALITY OF LIFE (HRQoL) ANALYSES

Health Related Quality of Life (HRQoL) is a secondary endpoint of this study. The safety analysis set will be used for analyzing QoL data unless specified differently.

HRQoL will be assessed using The EORTC QLQ-NMIBC24 Questionnaire provided in Appendix 2 of the protocol and summarized by treatment arm. These questionnaires will be completed by patients at screening and subsequent visits following the Schedule of Activities tables (protocol Section 1.3 and Section 1.4).

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.

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

All scoring information specific to the QLQ-NMIBC24 is presented in Table 6. The scoring method of this questionnaire will consist of a calculation of a raw score, followed by a linear transformation to standardize the raw score so that the scores range from 0 to 100. Raw scores from the EORTC QLQ-NMIBC24 questionnaires will be transformed based on the scoring manual.

Table 6: Scoring 11 Domains of QLQ-NMIBC24

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

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

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

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

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

Summary statistics 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) by treatment arm. The change from baseline for 11 domain scores will also be summarized with summary statistics at each above mentioned timepoints by treatment group. A listing of the 11 domain scores values will be produced for each patient.

The baseline is defined as the last HRQoL assessment on or prior to randomization.

13. HEALTH RESOURCE UTILIZATION

Potential differences in health resource utilization with UGN-102 ± TURBT versus TURBT alone, is an exploratory objective of this study. Number of patients hospitalized due to non-elective reasons along with the duration of hospitalization will be summarized using appropriate descriptive statistics and safety analysis set by treatment arms.

14. SAFETY ANALYSES

All safety analyses will be based on the safety analysis set.

14.1. 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.

14.2. Adverse Events

Adverse Events (AEs) will be coded according to the latest version of the Medical Dictionary for Regulatory Affairs (MedDRA). AEs will be graded according to the CTCAE Version 5.0.

AEs shall be recorded starting at signing of ICF until the end of the follow-up or early termination. A treatment-emergent adverse events (TEAE) is defined as an AE that occurs on or after the day of the first instillation of UGN-102 for patients in the UGN-102 ± TURBT arm or the day of the initial TURBT for patients in TURBT alone arm; or a pre-treatment AE that worsens during the study.

An overall summary with number and percentage of patients with all AEs, all serious AEs (SAE), TEAEs, different TEAE categories, like treatment-related, procedure related, serious, serious related to treatment or procedure, leading to treatment or study discontinuation, fatal, and AE of special interest (AESI) 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 as “Yes.” 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’ or missing.

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 (SOC) and for each preferred term (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 preferred term will be summarized under the maximum CTCAE grade recorded for the event.

In AE summaries, the primary SOC will be presented by descending order of frequency in UGN-102 ± TURBT column (alphabetically for ties), then PTs will be sorted by descending

order of frequency within each SOC (alphabetically for ties). The sort order for the PT will be based on their frequency in the treatment arm (UGN-102 ± TURBT).

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 or TURBT
2. Up to 3-month**s:** from the First Dose date of UGN-102/ date of Initial TURBT Procedure to 3-Month Visit and
3. Post 3-months: after 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 TURBT (or other alternative) procedures as patients with residual disease at the 3-Month visit will undergo TURBT (or alternative procedures).

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

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

All AEs will be listed.

14.3. 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 in the UGN-102 ± TURBT arm by PT and Time Period.

A study treatment-related SAE is defined as an SAE for which the investigator classifies the relationship to study treatment as "Yes." 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' or missing.

A separate listing for SAEs will be generated.

14.4. 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 in combination arm by PT only and separate supportive listings will be generated with patient level details for those patients:

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

14.5. Adverse Events of Special Interest

An adverse event of special interest is a grouping of adverse events that are of scientific and medical concern specific to UGN-102. A comprehensive list of MedDRA terms based on clinical review will be used to identify each type of events. For some events several AE preferred terms may be 'collapsed.'

The following categories of AE of special interest will be summarized by treatment arm 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)

Additional AE of special interest may be identified through manual review of all AEs by SOC and PT prior to the database lock.

Summaries of the number and percentage of patients with these events will be provided by category and PT and by treatment arm. 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.

AESI will be flagged in the general AE listing.

14.6. Clinical Laboratory Evaluations

Samples for hematology, serum chemistry and urinalysis will be collected at Screening and

subsequent visits following the SoA tables (protocol Section 1.3 and Section 1.4). 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.

Summaries of laboratory parameters will include hematology and serum chemistry using standard units. Baseline will be defined as the latest non-missing assessment prior to the first instillation of UGN-102 for patients in the treatment arm or prior to the first TURBT for patients in the control arm. 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. Missing baseline grade will be assumed as grade 0. For laboratory tests that are graded for both low and high values, summaries will be done separately and 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.

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

14.6.1. 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 7: PCS Laboratory Criteria

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

14.6.2. Potentially Clinically Significant Abnormal Liver Function Tests

There were no patients who met these criteria in the Phase 2 study (BL005). 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 8: 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

Parameter	Criterion
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.

14.7. 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 SoA tables (protocol Section 1.3 and Section 1.4).

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 by treatment arm.

14.7.1. 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 by treatment arm. 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 by treatment arm.

The PCS criteria are shown in Table 9.

Table 9: 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

Parameter	PCS Criterion
	≥105 mmHg ≥105 mmHg and increase of ≥15 mmHg from Baseline

14.8. General Physical Examination

General physical examinations as described in the Section 8.3.1 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 by treatment arm.

14.9. Urology-Oriented Examination

Urology-oriented examinations as described in the Section 8.3.1 of the protocol, will be performed at screening and subsequent visits following the protocol SoA table (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 by treatment arm.

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

15. PHARMACOKINETIC ANALYSES

Not applicable.

16. PHARMACODYNAMIC AND BIOMARKER ANALYS

Not applicable.

17. PHARMACOKINETIC/PHARMACODYNAMIC ANALYSES

Not applicable.

18. PHARMACOGENETIC DATA ANALYSES

Not applicable.

19. REFERENCES

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.

20. 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 10: 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 11: Imputation Rules of Prior or Concomitant Medication (CM) Start and End Dates

Dataset	Date	Missing Element	Rule
CM	Start Date	day, month, and year	• No Imputation for completely missing dates
		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 12: 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 inform consent date, then set MH start date = min (01JANYYYY, Inform consent date -1). • Else if the year of MH start date < year of inform consent date, 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 of treatment start date, then set CM start date = Inform consent date -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 (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 (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 13: 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 14: Laboratory CTCAE Grade Version 5.0 Criteria

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

Lab Parameter	Conventional Unit	CTCAE Grade v5.0				
		Grade 0	Grade 1	Grade 2	Grade 3	Grade 4
Neutrophil count decreased	$\times 10^3/\mu L$	≥LLN	1.5 – <LLN	1.0 – <1.5	0.5 – <1.0	<0.5
Platelet count decreased	$\times 10^3/\mu L$	≥LLN	75 – <LLN	50 – <75	25 – <50	<25
White blood cell decreased	$\times 10^3/\mu L$	≥LLN	3.0 – <LLN	2.0 – <3.0	1.0 – <2.0	<1.0
White blood cell increased (leukocytosis)	$\times 10^3/\mu L$	≤100			>100	
Hyperkalemia (Potassium increased)	mEq/L	≤ULN	>ULN – 5.5	>5.5 – 6.0	>6.0 – 7.0	>7.0
Hypokalemia (Potassium decreased)	mEq/L	≥LLN	3.0 – <LLN	Symptomatic with 3.0 – <LLN	2.5 – <3.0	<2.5
Hypernatremia (Sodium increased)	mEq/L	≤ULN	>ULN – 150	>150 – 155	>155 – 160	>160
Hyponatremia (Sodium decreased)	mEq/L	≥LLN	130 – <LLN	125 – <130	120 – <125	<120
Notes: ULN = upper limit of normal; LLN = lower limit of normal						

Appendix 3 Data Display Specifications

The data display specifications are contained in a separate document. The display mock-shells are provided as a guideline - the format and layout may be revised due to potential limitations of the programs and tools used to produce the displays. The following data displays will be provided for the End of Study analysis.

Table 15: Table of Contents for Data Display Specifications

Type	Section	Number	Title
Table	Patient Info	14.1.1	Screened and Randomized Patients
Table	Patient Info	14.1.2	Summary of Analysis Sets
Table	Patient Info	14.1.3	Summary of Patient Enrollment by Country and Site
Table	Patient Info	14.1.4.1	Summary of Patient Disposition through the End of Treatment Period for UGN-102

Type	Section	Number	Title
Table	Patient Info	14.1.4.2	Summary of Patient Disposition through the End of Study
Table	Patient Info	14.1.5	Summary of Protocol Deviations
Table	Patient Info	14.1.6.1	Summary of Demographics and Baseline Characteristics
Table	Patient Info	14.1.6.2	Summary of Baseline Prognostic Factors
Table	Patient Info	14.1.6.3	Summary of Baseline Smoking History
Table	Patient Info	14.1.7.1	Summary of General Medical History
Table	Patient Info	14.1.7.2	Summary of Prior and Concurrent Urothelial Carcinoma Medical Condition
Table	Patient Info	14.1.8.1	Summary of Urothelial Carcinoma Related Medical History at Screening
Table	Patient Info	14.1.8.2	Summary of Prior NMIBC Episodes and Prior TURBT
Table	Patient Info	14.1.9.1	Summary of Prior Medications
Table	Patient Info	14.1.9.2	Summary of Concomitant Medications
Table	Patient Info	14.1.10.1	Summary of Prior Surgical Procedures
Table	Patient Info	14.1.10.2	Summary of Concomitant Surgical Procedures
Table	Patient Info	14.1.11.1	Summary of Duration of Follow-up
Table	Efficacy	14.2.1.1	Summary of Disease-free Survival
Table	Efficacy	14.2.1.2	Summary of Kaplan-Meier Estimates of Disease-free Survival
Table	Efficacy	14.2.1.3	Sensitivity Analysis: Summary of Disease-free Survival Ignoring Extended Loss to Follow-up
Table	Efficacy	14.2.1.4	Sensitivity Analysis: Summary of Disease-free Survival Per Protocol Analysis Set
Table	Efficacy	14.2.1.5	Subgroup Analysis: Summary of Disease-free Survival

Type	Section	Number	Title
Table	Efficacy	14.2.2.1.1	Summary of Time to Recurrence
Table	Efficacy	14.2.2.1.2	Summary of Kaplan-Meier Estimates of Time to Recurrence
Table	Efficacy	14.2.2.2.1a	Summary of Response Rate at 3-Month Disease Assessment
Table	Efficacy	14.2.2.2.1b	Summary of Response Rate at 3-Month Disease Assessment for Patients with Prior TURBT
Table	Efficacy	14.2.2.2.2a	Summary of Response Rates at Scheduled Disease Assessments
Table	Efficacy	14.2.2.2.2b	Summary of Response Rates at Scheduled Disease Assessments for Patients with Prior TURBT
Table	Efficacy	14.2.2.3.1a	Summary of Duration of Response
Table	Efficacy	14.2.2.3.1b	Summary of Duration of Response for Patients with Prior TURBT
Table	Efficacy	14.2.2.3.1a	Summary of Duration of Response Ignoring Extended Loss to Follow-up
Table	Efficacy	14.2.2.3.2a	Summary of Kaplan-Meier Estimates of Duration of Response
Table	Efficacy	14.2.2.3.2b	Summary of Kaplan-Meier Estimates of Duration of Response for Patients with Prior TURBT
Table	Efficacy	14.2.2.4	Summary of Procedures for NMIBC During Study
Table	Health Outcomes	14.2.2.5	Summary of Change from Baseline in HRQoL
Table	Other	14.2.3.1	Summary of Patients Hospitalized for Non-Elective Reasons
Table	Safety	14.3.1.1	Overall Summary of Adverse Events by Time Period
Table	Safety	14.3.1.2.1	Incidence of Treatment-Emergent Adverse Events by System Organ Class, Preferred Term and Time Period
Table	Safety	14.3.1.2.2	Incidence of Treatment-Emergent Adverse Events by Preferred Term and Time Period
Table	Safety	14.3.1.2.3	Incidence of Treatment-Emergent Adverse Events by Preferred Term, Age and Time Period
Table	Safety	14.3.1.2.4	Incidence of Treatment-Emergent Adverse Events by Preferred Term, Sex and Time Period
Table	Safety	14.3.1.3.1	Incidence of Treatment-Emergent Adverse Events by Preferred Term and Worst Severity
Table	Safety	14.3.1.3.2	Incidence of Treatment-Emergent Adverse Events by Preferred Term and Worst Outcome

Type	Section	Number	Title
Table	Safety	14.3.1.4.1	Incidence of Treatment Related Treatment-Emergent Adverse Events by System Organ Class, Preferred Term and Time Period
Table	Safety	14.3.1.4.2	Incidence of Treatment Related Treatment-Emergent Adverse Events by Preferred Term and Time Period
Table	Safety	14.3.1.4.3	Incidence of Procedure Related Treatment-Emergent Adverse Events by Preferred Term
Table	Safety	14.3.2.1.1	Incidence of Serious Treatment-Emergent Adverse Events by Preferred Term and Time Period
Table	Safety	14.3.2.1.2	Incidence of Treatment Related Serious Treatment-Emergent Adverse Events by Preferred Term and Time Period
Table	Safety	14.3.2.1.3	Incidence of Procedure Related Serious Treatment-Emergent Adverse Events by Preferred Term and Time Period
Table	Safety	14.3.2.2.1	Incidence of Treatment-Emergent Adverse Events Leading to Treatment Discontinuation by Preferred Term
Table	Safety	14.3.2.2.2	Incidence of Treatment-Emergent Adverse Events Leading to Study Discontinuation by Preferred Term
Table	Safety	14.3.2.3.1	Incidence of Treatment-Emergent Adverse Events of Special Interest by Category and Preferred Term
Table	Safety	14.3.2.3.2	Summary of Characteristics of Treatment-Emergent Adverse Event of Special Interest by Category
Table	Safety	14.3.2.3.3	Summary of Onset and Duration of the First Occurrences of Treatment-Emergent Adverse Event of Special Interest
Table	Safety	14.3.4.1.1	Summary of Hematology Results: Observed Values
Table	Safety	14.3.4.1.3	Summary of Hematology Results: Change from Baseline with Respect to the Normal Range
Table	Safety	14.3.4.2.1	Summary of Chemistry Results: Observed Values
Table	Safety	14.3.4.2.2	Summary of Chemistry CTCAE Grade Changes from Baseline Grade
Table	Safety	14.3.4.2.3	Summary of Chemistry Results: Change from Baseline with Respect to the Normal Range
Table	Safety	14.3.4.3	Summary of Potentially Clinically Significant Laboratory Values by Time Point
Table	Safety	14.3.5.1	Summary of Vital Signs Results: Observed Values
Table	Safety	14.3.5.2	Summary of Potentially Clinically Significant Vital Signs by Time Point
Table	Safety	14.3.6.1	Abnormal General Physical Examination Results at Screening

Type	Section	Number	Title
Table	Safety	14.3.6.2	Abnormal Urology-Oriented Examination Results
Table	Safety	14.3.7	Summary of Treatment Exposure of UGN-102
Table	Safety	14.3.1.1	Summary of UGN-102 Dose (mg) by Visit
Listing	Patient Info	14.3.1.2.1	Randomized and Actual Treatment
Listing	Patient Info	16.2.1.1	Reason for Study Discontinuation
Listing	Patient Info	16.2.1.2	Patients Discontinued from UGN-102 Treatment
Listing	Patient Info	16.2.2.1	Protocol Deviations
Listing	Patient Info	16.2.2.2	Patients with Inclusion/Exclusion Criteria Deviation
Listing	Patient Info	16.2.3	Reason for Excluding Patients from Analysis Sets Full Analysis Set
Listing	Patient Info	16.2.4.1	Demographics and Baseline Characteristics
Listing	Patient Info	16.2.4.2	Baseline Prognostic Factors
Listing	Patient Info	16.2.4.3	Smoking History
Listing	Patient Info	16.2.4.4	Medical History
Listing	Patient Info	16.2.4.5	Prior and Concurrent Urothelial Carcinoma Medical Condition
Listing	Patient Info	16.2.4.6	Prior NMIBC Diagnosis and TURBT History
Listing	Patient Info	16.2.4.7	Prior and Concomitant Medications
Listing	Patient Info	16.2.4.8	Prior and Concomitant Surgical Procedures
Listing	Safety	16.2.5	UGN-102 Administration
Listing	Efficacy	16.2.6.1	Visit Level Disease Assessment Based on Cystoscopy, Histopathology, and Cytology

Type	Section	Number	Title
Listing	Efficacy	16.2.6.2	Evaluation of Visit Level Response
Listing	Efficacy	16.2.6.3	Disease-free Survival
Listing	Efficacy	16.2.6.4	Time to Recurrence
Listing	Efficacy	16.2.6.5	Duration of Response (DOR)
Listing	Health Outcomes	16.2.6.6	QLQ-NMIBC24
Listing	Safety	16.2.7.1	All Adverse Events
Listing	Safety	16.2.7.2	Serious Adverse Events
Listing	Safety	16.2.7.3	Treatment-emergent Adverse Events Leading to Treatment or Study Discontinuation
Listing	Safety	16.2.7.4	Deaths
Listing	Safety	16.2.8.1	Hematology Laboratory Values Lying Outside the Laboratory Reference Ranges
Listing	Safety	16.2.8.2	Chemistry Laboratory Values Lying Outside the Laboratory Reference Ranges
Listing	Safety	16.2.8.3	Urinalysis Laboratory Values
Listing	Safety	16.2.8.4	Potentially Clinically Significant (PCS) Laboratory Values and Abnormal LFT
Listing	Safety	16.2.9.1	Vital Signs Results
Listing	Safety	16.2.9.2	Potentially Clinically Significant Vital Sign Results
Listing	Safety	16.2.10.1	General Physical Examinations with Abnormal Results
Listing	Safety	16.2.10.2	Urology-Oriented Physical Examinations with Abnormal Results
Figure	Efficacy	14.2.1.1a	Kaplan-Meier Plot of Disease-free Survival
Figure	Efficacy	14.2.1.1b	Kaplan-Meier Plot of Disease-free Survival Ignoring Extended Loss to Follow-up
Figure	Efficacy	14.2.1.2	Hazard Ratios and 95% Confidence Intervals for Disease-free Survival by Subgroups
Figure	Efficacy	14.2.2a	Kaplan-Meier Plot of Duration of Response
Figure	Efficacy	14.2.2b	Kaplan-Meier Plot of Duration of Response Ignoring Extended Loss to Follow-up