Janssen Research & Development

Statistical Analysis Plan

Phase 2b Clinical Study Evaluating Efficacy and Safety of TAR-200 in Combination with Cetrelimab, TAR-200 Alone, or Cetrelimab Alone in Participants with High-Risk Non-Muscle Invasive Bladder Cancer (NMIBC) Unresponsive to Intravesical Bacillus Calmette-Guerin (BCG) who are Ineligible for or Elected Not to Undergo Radical Cystectomy

Protocol 17000139BLC2001; Phase 2b

JNJ-17000139 (TAR-200)
JNJ-63723283 (Cetrelimab)

Status:	Approved
Date:	04 October 2024
Prepared by:	Janssen Research & Development, LLC
Document No.:	CCI [redacted]

Compliance: The study described in this report was performed according to the principles of Good Clinical Practice (GCP).

TABLE OF CONTENTS

VERSION HISTORY

Table 1: – SAP Version History Summary

SAP Version	Approval Date	Change	Rationale
1.0		Not Applicable	Initial release
2.0		• Instances of cetrelimab specified as IV cetrelimab. • Change wording of secondary endpoint to DOR • Changes made to study design section based on changes in protocol • Evaluable analysis set updated to include subjects who have disease assessment at any timepoint • Analyses simplified in sections 5.6.2.1-5.6.2.3 • Subsequent systemic therapy removed from subgroup table • Redundancy of papillary tumor removed in baseline characteristics and subgroup tables • Flexibility added to interim analysis • Expand statistical hypothesis to include all cohorts	• Updated interim analysis plan and statistical hypothesis • Protocol Amendment 2/3 and other minor changes
3.0		• Add details on Cohort 4 in design and analysis • Update time to event endpoints (OS and Time to Cystectomy) to start from first treatment. • Clarify populations (analysis sets) for analysis • Update DOR to DoR • Updates to analysis sections based on clinical review and FDA feedback • Add analysis of subsequent therapies	• Protocol Amendment 4 • Clarification of Analysis • Updates based on FDA feedback
4.0		• Add BCG strain subgroup • Update DoR to DOR • Add language regarding 60-day efficacy update • Clarify disease assessments in Table 4	• Updates based on FDA feedback

1. INTRODUCTION

This document presents the Statistical Analysis Plan (SAP) for protocol 17000139BLC2001. This SAP contains definitions of analyses sets, derived variables, and statistical methods for the planned analyses for this study.

This SAP follows guidelines provided in the International Conference on Harmonization (ICH) Topic E9 Statistical Principles for Clinical Trials. In the event of future amendments to the protocol, this SAP may be modified as necessary to account for changes relevant to the statistical analysis.

1.1. Objectives

Primary Objective (Cohort 1, 2, and 3 only)

To evaluate the overall complete response (CR) rate in participants treated with TAR-200 in combination with intravenous (IV) cetrelimab (Cohort 1), or TAR-200 alone (Cohort 2), or IV cetrelimab alone (Cohort 3) with Carcinoma in Situ (CIS), with or without concomitant high-grade Ta or T1 papillary disease.

Primary Objective (Cohort 4 only)

To evaluate disease-free survival (DFS) in participants treated with TAR-200 alone with papillary only disease (Cohort 4).

Secondary Objectives (Cohort 1, 2, and 3 only)

The secondary objectives are:

- To evaluate the duration of response (DOR) in participants treated with TAR-200 in combination with IV cetrelimab (Cohort 1), or TAR-200 alone (Cohort 2), or IV cetrelimab alone (Cohort 3) with CIS (with or without concomitant Ta or T1 papillary disease) who achieve a CR.
- To determine the overall survival (OS) in all participants
- To evaluate the pharmacokinetics (PK) of gemcitabine and major metabolite 2',2'-difluorodeoxyuridine (dFdU) in urine and plasma in TAR-200 in combination with IV cetrelimab (Cohort 1), or TAR-200 alone (Cohort 2 & Cohort 4).
- To evaluate the PK and immunogenicity of IV cetrelimab in serum in combination with TAR-200 (Cohort 1), or IV cetrelimab alone (Cohort 3).
- To evaluate health-related quality of life (HRQoL) in all participants.
- To assess the safety and tolerability of participants receiving TAR-200 in combination with IV cetrelimab (Cohort 1), or TAR-200 alone (Cohort 2 & Cohort 4), or IV cetrelimab alone (Cohort 3).

Exploratory Objectives (Cohort 1, 2, and 3 only)

The exploratory objectives are:

- To determine the incidence of and time to cystectomy in the study in all participants.
- To identify biomarkers (DNA, RNA, protein) in blood, urine and tissue samples that elucidate mechanisms of action and predict response to TAR-200 in combination with IV cetrelimab (Cohort 1), or TAR-200 alone (Cohort 2), or IV cetrelimab alone (Cohort 3); and its' association with RFS, in participants with Ta/T1/CIS BCG-unresponsive HR-NMIBC.
- To explore gemcitabine and dFdU concentration in biopsied tissue in TAR-200 in combination with IV cetrelimab (Cohort 1), or TAR-200 alone (Cohort 2 & Cohort 4).
- To characterize the impact of medical resource utilization in all participants.
- To explore the relationships between PK, Pharmacodynamic, adverse event (AE) profile, and antitumor activity in all participants.

1.2. Study Design

This is an open-label, parallel-group, multi-center study evaluating the efficacy and safety of intravesical TAR-200 in combination with IV cetrelimab, TAR-200 alone, or IV cetrelimab alone in participants with HR-NMIBC unresponsive to prior intravesical BCG therapy who are either ineligible for or have elected not to undergo radical cystectomy (RC) after making an informed decision. Ineligibility for or reason for refusal of cystectomy will be documented in the electronic case report form (eCRF). All enrolled participants must have received adequate BCG and confirmed CIS [with or without papillary disease] or confirmed papillary only disease (high-grade Ta or any T1, without CIS), within 12 months of completion of the last dose of BCG therapy. All visible papillary disease must be fully resected (absent) prior to randomization (residual CIS acceptable for participants eligible for Cohort 1, 2, and 3 only) and documented in the eCRF based on Screening cystoscopy. Adequate BCG therapy is defined as a minimum of 5 of 6 full doses of an induction course (adequate induction) plus 2 of 3 doses of a maintenance course, or at least 2 of 6 doses of a second induction course (Food and Drug Administration 2018[1]). Single-dose perioperative intravesical chemotherapy prior to study treatment is allowed in accordance with institutional guidelines.

There are 4 cohorts in this study: Cohort 1 TAR-200 in combination with IV cetrelimab; Cohort 2 TAR-200 alone; Cohort 3 IV cetrelimab alone; and, added with Amendment 4, Cohort 4 TAR-200 alone. Per previous protocol versions, participants with CIS [with or without papillary disease] were randomly assigned (2:1:1) to receive intravesical TAR-200 in combination with IV cetrelimab (Cohort 1) or TAR-200 alone (Cohort 2) or IV cetrelimab alone (Cohort 3). Following review of the futility data, the Sponsor elected to close further enrollment into Cohorts 1 and 3. Additionally, in accordance with Amendment 3, the Sponsor decided to expand enrollment into Cohort 2 to approximately 80 participants. Participants already enrolled in Cohorts 1 and 3 who are in the Treatment Phase and benefiting from treatment will continue to receive the assigned treatment until the protocol-defined criteria for treatment discontinuation are met (eg, recurrence/progression) (refer to Protocol Section 7.1 for discontinuation criteria). All participants, including those in Cohort 1 and 3, who discontinue study treatment should not be considered

withdrawn from the study and will continue to be followed for assessments in the Follow-up (FU) Phase as described in the Protocol SOA (Section 1.3).

As of Protocol Amendment 4, the study design includes Cohort 4 of approximately 50 participants with BCG-unresponsive papillary disease only (high-grade Ta or any T1). Enrollment into Cohort 4 may also be discontinued at Sponsor discretion (eg, if Cohort 2 enrollment is completed).

For the purposes of this study, both TAR-200 and IV cetrelimab, alone or in combination, will be referenced as study drug. For Cohort 1, TAR-200 will be dosed every 3 weeks (Q3W) up to the first 24 weeks. Dosing of TAR-200 will then be every 12 weeks through Week 99 (Year 2). Cetrelimab will be dosed at 360 mg IV every 3 weeks (Q3W) through Week 78 (18 months).

For Cohort 2 and Cohort 4, TAR-200 alone will be dosed Q3W for up to the first 24 weeks on study. Dosing of TAR-200 will then be every 12 weeks through Week 99 (Year 2).

For Cohort 3, cetrelimab will be dosed at 360 mg IV Q3W through Week 78 (18 months).

There is only 1 dose and schedule option for TAR-200. TAR-200 is available in one dose only (containing 225-mg free base equivalents of gemcitabine) and will remain indwelling for up to 3 weeks (approximately 21 days) per each dosing cycle. Stopping criteria is allowed based on TAR-200 related adverse events. Dose modifications are not permitted due to the nature of TAR-200, except for early removal, delayed placement (i.e., a missed dosing cycle), or discontinuation of TAR-200. There is only 1 dose and schedule option for IV cetrelimab. Cetrelimab is dosed and scheduled for 360 mg IV Q3W. Dose delays and/or skipped doses are allowed. Guidelines for management of immune-related toxicities are provided.

Disease assessments will be performed at Screening, then every CCI weeks for the first 2 years and every CCI weeks thereafter through Year 3, based on an integrated evaluation of local cystoscopy and centrally read urine cytology. In patients with CR, a disease assessment will be conducted at CCI from date of initial CR. Chest and abdominopelvic imaging will be performed at codified study timepoints per the Protocol Schedule of Activities (SOA). Transurethral Resection of Bladder Tumor (TURBT)/biopsies will be performed per the schedule of activities, and as clinically indicated. Biopsy assessment from TURBT, performed at screening, to determine the participant's eligibility may be utilized from the time point in which the participant was deemed to have a recurrence of CIS after adequate BCG therapy. Biopsy samples will be evaluated by the local histopathology laboratory for clinical management at the discretion of the Investigator; however, any tissue biopsy sampling will be submitted to the sponsor for a central confirmation of participant's disease status. Response status will be programmatically determined based on local cystoscopy results, centrally read urine cytology results, centrally evaluated biopsy results, and local imaging, if available. See section 5.3.1 for definition of complete response (CR).

Cohort 2 participants with a CR must have a CCI disease assessment visit at CCI from the date of onset of CR. The CCI disease assessment from onset of CR may occur in lieu of CCI disease assessment, if the two visits are within 6 weeks of each other. The CCI disease assessment post initial CR date will include assessment cystoscopy, urine cytology, TURBT bladder biopsy (if clinically indicated) and CT/MRI with contrast (if applicable).

Assessment cystoscopy during Screening will still be required to determine if bladder accessibility is adequate for insertion of TAR-200 and that any papillary disease has been completely excised. Participants are required to discontinue from study treatment if there is confirmed high-risk (HR) disease persistence, recurrence or progression based on central urine cytology and/or central biopsy assessment or local biopsy/local imaging after consultation with the Sponsor.

Disease status will be assessed through the end of study in accordance with the institutional standard of care and will be collected at least at 6-month intervals through the end of study. Survival status will also be determined and collected in the electronic case report form (eCRF) until death, withdrawal of consent, or the end of the study, whichever occurs first.

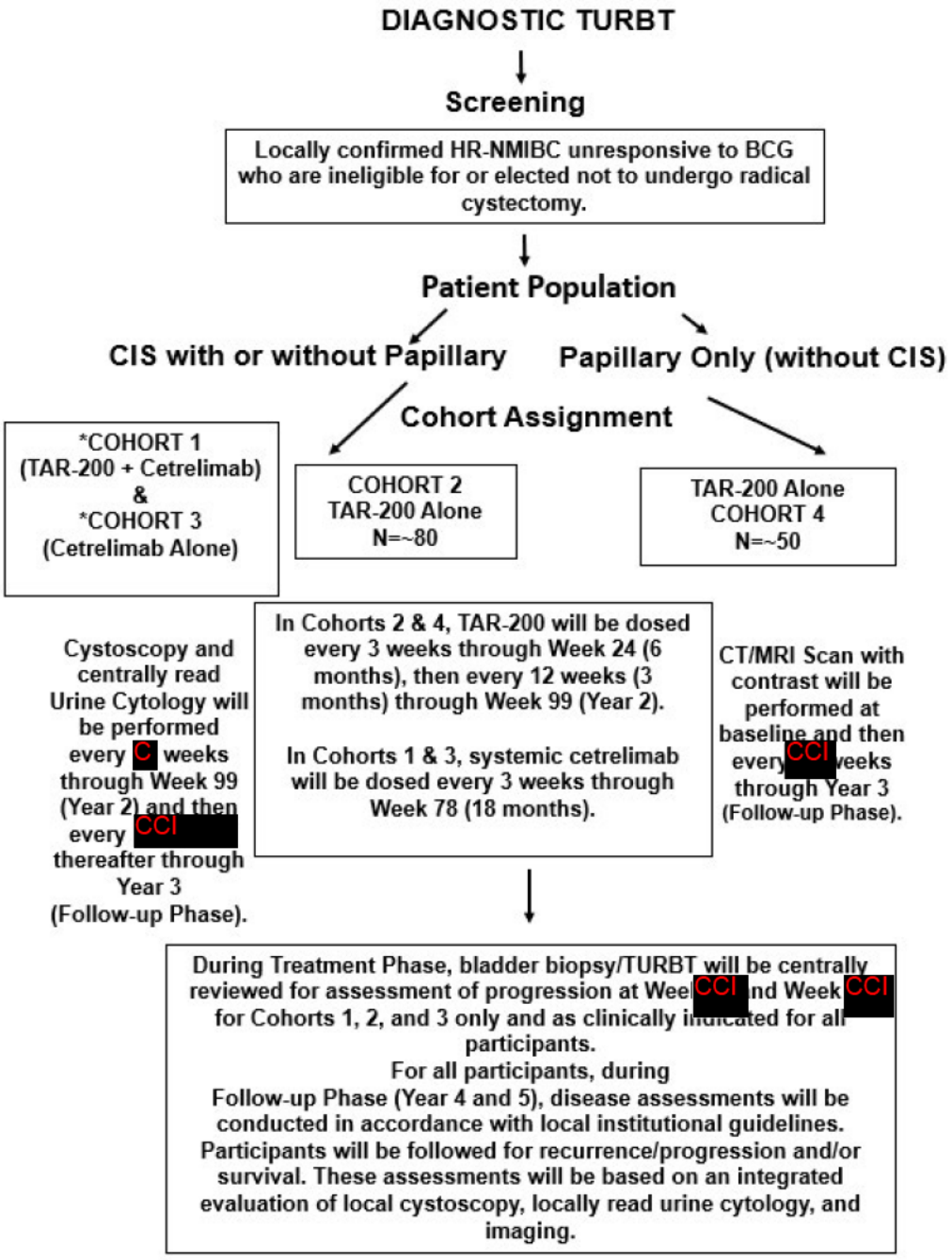
Safety oversight will be performed by the medical monitor and Independent Data Monitoring Committee (IDMC), which will be composed of individuals with the appropriate expertise. The IDMC will operate under the rules of an approved charter and will meet as outlined in the IDMC charter.

Study Treatment After Futility Analysis

The IDMC will meet at pre-specified analysis and as specified in the IDMC charter to review the data and to make a recommendation.

A futility analysis was performed after there were 75 evaluable participants from Cohorts 1-3 who had adequate disease assessment post-baseline (See Section 4 for the definition of evaluable analysis set). Any cohort could be closed for futility if its overall CR rate is less than 20%. Additionally, the IDMC reviewed and evaluated the data from individual cohorts and the totality of the data to provide its recommendation to continue the study. Following Amendment 3 and the Futility Analysis, Cohorts 1 and 3 were closed for further enrollment and the sample size for Cohort 2 was expanded to approximately 80 participants.

A diagram of the study design is provided.



Abbreviations: BCG=Bacillus Calmette-Guérin; CT=computed tomography; MRI=magnetic resonance imaging

Abbreviations: BCG=Bacillus Calmette-Guérin, CIS=carcinoma in situ, CT=computed tomography, HR-NMIBC=high-risk non-muscle invasive bladder cancer; IDMC=Independent Data Monitoring Committee, MRI=magnetic resonance imaging; TURBT=transurethral resection of bladder tumor.

*For Cohorts 1 and 3 only, enrollment is closed. Participants randomized to Cohorts 1 or 3 will continue on study as outlined in the Treatment and Follow-up Phase of the Schedule of Activities (Refer to Section 1.3).

2. STATISTICAL HYPOTHESES

The statistical hypothesis for Cohorts 1, 2 and 3 is that the combination of TAR-200 and IV cetrelimab, TAR-200 alone, or IV cetrelimab alone will lead to a higher overall CR rate than observed in historical studies with chemotherapy (20% based on Valrubicin data[2] described in Protocol Section 2.1, i.e, H0 $\leqslant$ 20% vs. Ha > 20%) in the population with CIS unresponsive to intravesical BCG therapy who are ineligible for or have elected not to undergo radical cystectomy.

The statistical hypothesis for Cohort 4 is that TAR-200 alone will lead to a 1-year DFS higher than 40% in the population with papillary disease only (high grade Ta or any T1) unresponsive to intravesical BCG therapy who are ineligible for or have elected not to undergo radical cystectomy.

3. SAMPLE SIZE DETERMINATION

Per previous versions of the protocol, the sample size of 100 participants would provide over 90% power to reject the null hypothesis assuming an overall CR rate of 55% (Cohort 1) and the sample size of 50 participants will provide about 90% power assuming an overall CR rate of 40% (Cohorts 2 and 3), using 2-sided Z test at alpha = 0.05. After the futility analysis, the adjusted sample size of 80 participants for Cohort 2 will provide over 90% power to reject the null hypothesis at 2-sided alpha of 0.05. Table 2 describes the point estimate and its 95% exact confidence interval (CI) (two-sided) at selected number of participants with CR based on 80 participants in Cohort 2.

Table 2: Point Estimate and the 95% Exact Confidence Interval based on 80 Participants

Number of participants with CR	Observed CR rate	95% exact CI (2-sided)
40	50%	(39%, 61%)
44	55%	(43%, 66%)
48	60%	(48%, 71%)
52	65%	(54%, 75%)
56	70%	(59%, 80%)
60	75%	(64%, 84%)

CR=complete response; CI=confidence interval

The sample size of 50 participants in Cohort 4 will provide 93% probability to have the lower boundary of 95% CI of DFS at 1 year exceeding 35% (or about 78% probability exceeding 40% 1-year DFS), if we assume the 1-year DFS rate is 60%, 5% annual dropout rate, 7-month enrollment and additional 12-month follow-up.

4. POPULATIONS (ANALYSIS SETS) FOR ANALYSIS

For purposes of analysis, participants will be classified into the following analysis sets (Table 3).

Table 3: Analysis Sets

Population	**Description**
Enrolled	Participants who were not screen failures and were assigned to an intervention group
Full	All enrolled participants who received at least 1 dose of study treatment
Safety	All enrolled participants who received at least 1 dose of study treatment

5. STATISTICAL ANALYSES

5.1. General Considerations

5.1.1. Study Phases

Screening Phase

The screening phase is only defined for enrolled participants who were not screen failures and were assigned to an intervention cohort. The reference start date of the screening phase is the day when participants signed the informed consent form. The reference end date is the day before the reference start date of the treatment phase.

Treatment Phase

The treatment phase is only defined for participants who received at least 1 dose of study intervention. The treatment phase is defined to be between the date of first dose of study intervention and the date of the end of treatment visit. If the date of the end of treatment visit is not available, the date of last dose of study medication + 21 days will be used. The assessments performed during the 'End-of-Treatment Visit' will be included in this phase.

Follow-up Phase

The follow-up phase is only defined for participants who entered the follow-up phase. The reference start date of the follow-up phase is the day after the reference end date of the treatment phase. The reference end date of follow-up phase is the date of death, withdrawal of consent or the end of the study, whichever occurs first.

5.1.2. Baseline Measurement and Study Day Definition

Baseline is defined as the last observation on or before receiving the first dose of any study treatment.

The first dose of treatment is study day 1 and is the reference start date in the study. If a participant is not treated the date of randomization is used as the reference start date.

5.2. Participant Dispositions

The number of participants in the following disposition categories will be summarized using tables throughout the study by cohort:

- Participants enrolled
- Participants who received study drug
- Participants who discontinued study drug
- Reasons for discontinuation of study drug
- Participants who completed study drug
- Participants who completed the study (see definition below)

- Participants who discontinued study
- Reasons for discontinuing study

Listings of participant's dispositions will be provided for the following categories:

- Participants who discontinued study drug
- Participants discontinued study
- Participants who were enrolled yet did not receive study drug.
- Listing of overall study start and end dates

The following definitions will be used in the determination of participant disposition:

Participant Completion of Treatment Definition

A participant will have completed treatment for Cohort 1, Cohort 2, and Cohort 4, when they have completed treatment through Week 99 (removal of the TAR-200); and for Cohort 3, when they have completed treatment through Week 78. Participants who prematurely discontinue study treatment for any reason before completion of the treatment phase will not be considered to have completed the study treatment.

Participant Completion of Study Definition

A participant will be considered to have completed the study if they have completed the treatment and Follow-up Phases or died before the end of the treatment or Follow-up Phases. Participants who have been lost to follow-up or have withdrawn consent for study participation before the end of the Follow-up Phase will not be considered to have completed the study.

5.3. Primary Endpoint(s) Analysis

5.3.1. Definition of Endpoint(s)

For Cohort 1, 2, and 3 the primary endpoint is the CR rate, which is defined as the proportion of participants who met at least one of the following: negative cystoscopy and negative (including atypical) centrally read urine cytology, or positive cystoscopy with biopsy-proven benign or low-grade NMIBC and negative (including atypical) centrally read cytology at any time point.

Patients with suspicious cytology will be imputed based on the subsequent cytology. If the subsequent cytology is positive, the previous cytology will be analyzed as positive, and the patient will be considered to be a non-CR at the previous visit. If the subsequent cytology is negative, the previous cytology will be analyzed as negative, and the response status will be determined together with other assessments at the previous visit.

Disease assessment schedule and other details are described in Table 4. Disease assessments occurring within +/-6 weeks of the scheduled visit date are considered as in the same analysis visit window to determine the response status.

Table 4: Disease assessments for deriving a response.

Disease Assessments	Scheduled Visits	Central Confirmation
Cystoscopy	Weeks 12, 24, 36, 48, 60, 72, 84, 96, 120, 144 and ~365 days post CR, if applicable	Local assessment
Urine Cytology	Weeks 12, 24, 36, 48, 60, 72, 84, 96, 120, 144 and ~365 days post CR, if applicable	Central review
TURBT (Biopsy)	Weeks 24, 48 and at any time point if clinically indicated (to prove benign/low-grade NMBIC when there is a positive cystoscopy)	Local assessment and Central review
Imaging (CT/MRI)	Week 24, 48, 72, 96, 120, 144, and when clinically indicated	Local assessment

The primary endpoint for Cohort 4 is the disease-free survival (DFS) for BCG-unresponsive papillary only disease.

DFS is defined as the time from first dose to the time of one of the following events, whichever occurs first:

1. The first recurrence of high-risk disease (high-grade Ta, any T1 or CIS).
2. Progression to MIBC (≥T2) or to lymph node (N+) or to distant disease (M+), whichever occurs first
3. Death due to any cause

A DFS event is defined as a positive urine cytology based on central assessment, high risk disease (high grade Ta, Tis, T1, ≥T2) based on central pathology assessment, or lymph node (N+) or distant metastasis (M+) disease from local imaging assessment or central pathology report, or death due to any cause. Censoring rules can be found in Table 5.

Patients with suspicious cytology will be considered to have DFS event if the subsequent cytology is positive, i.e., the previous suspicious cytology will be imputed based on the subsequent cytology. If the subsequent cytology is negative, this will not be considered a DFS event.

Table 5: Censoring Rules for DFS

Situation	Date of Disease Event or Censoring	Censoring Indicator
Disease recurrence/progression/death	Date of: - The first recurrence of high-risk disease (high-grade Ta, any T1 or CIS).	No

	- progression to MIBC (≥T2) or to lymph node (N+) okay or to distant disease (M+), whichever occurs first - Death due to any cause	
No DFS event at data cutoff	Date of last disease assessment	Yes
DFS event after subsequent therapy	Date of last disease assessment before the start of subsequent therapy	Yes
Study discontinuation without a DFS event	Date of last disease assessment	Yes
No post-baseline disease assessment	Date of enrollment	Yes
DFS event after missing 2 or more consecutive disease assessments	Date of last disease assessment before the consecutive missing assessments	Yes

5.3.2. Estimand

Primary Estimand for Cohorts 1, 2, and 3

The primary estimand, the main clinical quantity of interest to be estimated in the study, is defined by the following five components:

- Population: participants with HR-NMIBC unresponsive to intravesical BCG therapy who are ineligible for or have elected not to undergo radical cystectomy.
- Variable: overall CR.
- Study Drug: TAR-200 in combination with IV cetrelimab, or TAR-200 alone, or IV cetrelimab alone.
- Intercurrent event: subsequent therapy. Hypothetical Strategy: Response after this intercurrent event will not be considered as a response.
- Summary: overall CR rate.

The primary efficacy endpoint, overall CR rate, is defined as the percentage of participants achieving a CR at any time post-treatment.

Primary Estimand for Cohort 4

- Population: participants with HR-NMIBC papillary disease only unresponsive to intravesical BCG therapy who are ineligible for or have elected not to undergo RC.
- Variable: disease-free survival (DFS)
- Study Drug: TAR-200 alone
- Intercurrent event: Subsequent anticancer therapy. While on Treatment Strategy: censor DFS event occurred after subsequent therapy, as if the participants had continued treatment as planned and had not switched to subsequent anticancer therapies.
- Summary: DFS rate at 1 year

The primary efficacy endpoint for Cohort 4, DFS, is defined as the time from treatment to recurrence, progression, or death due to any reason.

5.3.3. Analysis Methods

All enrolled participants in Cohorts 1, 2, and 3 will have confirmed CIS at enrollment and will be stratified based upon the presence or absence of concomitant papillary disease. The cohorts will be analyzed separately. There will be no formal comparison between the 4 cohorts.

For Cohorts 1, 2, and 3, a Z test with normal approximation will be used to compare the CR rate with 20%. The overall CR rate and its 95% exact confidence interval will also be calculated for each cohort based on binomial distribution. They type 1 error rate will be controlled at 5% (2 sided). The CR rate at specified timepoints (i.e. 3, 6, 9, 12, 15 months) will be provided. Participants who are not evaluable (i.e. due to trial discontinuation or missing adequate post-baseline disease assessment) will be considered a non-responder in the analysis.

For Cohort 4, the distribution (median and Kaplan-Meier curves) of DFS will be provided using Kaplan Meier estimate. DFS rate at 12 months, 18 months and 24 months will be calculated based on the Kaplan-Meier estimate.

Primary Analysis Timepoints:

The primary analysis of Cohorts 1-3 will be performed on the full analysis set based on assigned treatment cohort, approximately 6 months after treatment initiation of the last participant in Cohorts 1, 2, or 3.

Two pre-specified analyses will be performed will be performed at approximately 9 and 15 months (primary analysis of Cohort 4) after treatment initiation of the last participant in Cohort 4.

Analyses Related to Primary Endpoint:

Cohorts 1-3:

- Concordance of central and local cytology and pathology samples
- Concordance of CIS presence at screening
- Summary of worst disease stage for non-responders and responders with disease recurrence/progression
- Summary of worst TNM stage and corresponding listing
- A summary of worst CT/MRI result and corresponding listing
- Listing of disease evaluation results
- Summary of cystoscopy modality used at baseline vs post-baseline and corresponding listing

Similar analyses will be performed on Cohort 4 at the time of the primary analysis in Cohort 4.

5.3.4. Sensitivity Analyses

A sensitivity analysis of overall CR rate in Cohorts 1-3 on subjects who had at least one evaluable post-baseline disease assessment.

A sensitivity analysis of overall CR rate in Cohorts 1-3 where CR is determined based on local investigator assessment (i.e. urine cytology, biopsy/TURBT, and CT/MRI is locally assessed) will be performed.

A sensitivity analysis of overall CR rate and DOR in Cohorts 1-3 on only subjects who meet the FDA's definition of BCG unresponsive will be performed.

A sensitivity analysis of overall CR rate and DOR in Cohorts 1-3 on only subjects who meet the FDA's definition of BCG unresponsive and used consistent cystoscopy technology or white light at baseline and enhanced cystoscopy modality at follow-up.

A sensitivity analysis in Cohort 4 where DFS is based on local investigator's assessment of a DFS event.

Other sensitivity analyses may be performed ad hoc, as needed.

5.4. Secondary Endpoint(s) Analysis

- Duration of Response (DOR)
- Overall survival
- To evaluate the pharmacokinetics (PK) of gemcitabine and the PK and immunogenicity of TAR-200 with or without IV cetrelimab
- Health related quality of life (HRQoL) in all participants
- To assess the safety of TAR-200 with IV cetrelimab, TAR-200 alone and cetrelimab alone

5.4.1.1. Definition of Endpoint(s)

Duration of response (DOR) is defined from the date of first CR achieved to first evidence of recurrence, progression or death due to any cause (whichever is earlier) for participants who achieve a CR. Censoring rules can be found in Table 6 below. Response or recurrence/progression will be assessed by cystoscopy, and centrally read urine cytology. Centrally read pathology and local imaging results will also be used, as applicable.

Table 6: Censoring Rules for DOR

Situation	Date of DOR event or censoring	Censoring indicator
Disease recurrence/progression or death due to any cause	Date of disease progression/recurrence or death, whichever occurs first	No
Initiated subsequent therapy without prior recurrence/progression	Date of last disease assessment before the start of subsequent therapy	Yes

Discontinued from study without disease recurrence/progression or death due to any cause	Date of last disease assessment	Yes
No documented disease progression/recurrence or death at the end of study (or at data cutoff)	Date of last disease assessment	Yes
DOR event after missing 2 or more consecutive disease assessments	Date of last disease assessment before the consecutive missing assessments	Yes

Overall survival (OS) is measured from the date of first dose to the date of the participant's death. If the participant is alive or the vital status is unknown, then the participant's data will be censored at the date the participant was last known to be alive.

5.4.1.2. Analysis Methods

The Kaplan-Meier method will be used to estimate the distributions of DOR and OS along with corresponding 95% confidence interval for their median and selected timepoints. DOR will be based on participants with complete response only while OS analysis will be based on the full analysis set. The DOR rate at 1 and 2-years post-CR will also be provided based on Kaplan-Meier method. A Kaplan-Meier plot and swim plot of DOR will be provided. Observed DOR will be summarized descriptively (N, mean, SD, median, and range [minimum, maximum]). The number and percentage of subjects with DOR >= 6 months, 9 months, 12 months, through 36 months will be provided.

Observed 1-year DOR

An updated analysis of DOR, after the last participant has 365 days follow-up from initial CR, or has died or withdrawn from study, whichever occurs first, will be provided at the 60-day efficacy follow-up to the NDA.

5.5. Exploratory Endpoints/Analysis

- Incidence of cystectomy for all participants in the full analysis set. Measured by determining the proportion of participants who underwent a cystectomy. The incidence rate and its 95% exact confidence interval (CI) will be calculated for each cohort based on binomial distribution.
- Time to cystectomy for all participants in the full analysis set. Measured from the date of first dose to the date of the cystectomy. For participants who have not received cystectomy, data will be censored at the last study assessment. The Kaplan-Meier method will be used to estimate the distributions of time to cystectomy with corresponding 95% confidence interval for their median and selected timepoints.
- To identify biomarkers (DNA, RNA, protein) in blood, urine and tissue samples that elucidate mechanisms of action and predict response to TAR-200 in combination with IV cetrelimab (Cohort 1), or TAR-200 alone (Cohort 2), or IV cetrelimab alone (Cohort 3); and its' association with RFS, in participants with Ta/T1/CIS BCG-unresponsive HR-NMIBC.

- To characterize the impact of medical resource utilization (MRU) in all participants.
- Relationships between PK, PD, adverse event profile, and antitumor activity in all participants

5.6. Safety Analyses

All safety analyses will be based on the safety analysis set based on actual intervention received, unless otherwise specified. Participants assigned to receive both TAR-200 and cetrelimab (Cohort 1), but only received one study drug are summarized under their assigned intervention group.

For all continuous safety variables, descriptive statistics by intervention group will include the N, mean, standard deviation, median, minimum, and maximum. Categorical variables will be summarized by intervention group using frequency counts and percentages.

5.6.1. Extent of Exposure

The number and percentage of participants who receive each study drug (TAR-200+cerelimab, TAR-200 alone, cetrelimab alone) within a study intervention will be summarized.

Total duration of intervention is defined as (date of last dose of study intervention – date of first dose of study intervention) +1. Descriptive statistics (N, mean, SD, median, and range (minimum, maximum)) of total duration of intervention will be summarized for each study drug within a study intervention. Total treatment (Participant-years) of intervention is calculated as the sum of each participant's days of intervention/365.25 and will be summarized by intervention group. Duration of intervention will also be summarized in the following duration categories: [<12 weeks, 12-<24 weeks, 24-<36 weeks…] for each study drug within a study intervention. Cumulative duration of intervention [>=12 weeks, >=24 weeks, …] will be summarized.

Total number of administrations received is defined as insertions that were successfully performed. Descriptive statistics (N, mean, SD, median, and range [minimum, maximum]) of total number of administrations received will be summarized. Total number of administrations received will also be summarized by category (1-4, 5-8, 9-14, etc.).

Total dose of study treatment is defined as the sum of all doses received by a participant during the treatment phase. Total dose of study treatment will be summarized descriptively statistics (N, mean, SD, median, and range [minimum, maximum]) across all participants.

For TAR-200, a participant's average indwelling time is defined as the date of TAR-200 removal – date of TAR-200 insertion+1 averaged across the number of administrations of TAR-200 that participant received. Average indwelling time will be summarized with descriptive statistics (N, mean, SD, median, and range [minimum, maximum]) across all participants.

Study intervention compliance will be summarized descriptively. See Appendix 7 for further details.

A summary of dose modifications will be provided on:

- The total number of skipped insertions early removals, and late removals

- The number (%) of participants with skipped insertions. Skipped insertions include those that were not attempted or were attempted and unsuccessful
- The number (%) of participants with TAR-200 removed prior to or later than protocol schedule
- Reasons for skipped insertion (insertion not attempted or attempted but unsuccessful insertion) and reason for early or late removal
- Dose modification prior to cetrelimab infusion (rate decreased, infusion skipped, infusion delayed) or during infusion (infusion aborted, infusion interrupted, infusion rate decreased)

A listing of dose modifications will be provided.

A skipped dose is defined as a TAR-200 insertion that was not attempted or attempted but was unsuccessful or a cetrelimab infusion that is not administered within the study window and is not due to study discontinuation. For cetrelimab only, a dose interruption is when the infusion is temporarily stopped before the entire dose has been administered.

Total number (N) of insertions attempted, successful insertions, insertions attempted but not successful, insertions not attempted, early removals, and late removals will be totaled across all participants by cohort. Insertion success rate (%), defined as the proportion of successful insertions out of all insertion attempts, will be provided.

A summary of prophylactic antibiotic use at each insertion will be provided.

5.6.2. Adverse Events

The verbatim terms used in the CRF by Investigators to identify AEs will be coded using the Medical Dictionary for Regulatory Activities (MedDRA). Any AE with onset date and time after the first dose through 100 days after the last study drug administration or prior to the start of subsequent anticancer therapy, whichever is earlier, or the follow-up AE (linked to an existing treatment-emergent adverse event [TEAE]) with onset date and time beyond 100 days after the last dose of study intervention but prior to the start of subsequent therapy is considered to be treatment-emergent. If the event occurs on the day of the initial administration of study intervention, and either event time or time of administration are missing, then the event will be assumed to be treatment emergent. If the event date is recorded as partial or completely missing, then the event will be considered to be treatment emergent unless it is known to be prior to the first administration of study intervention or later than 100 days after last study drug administration based on partial onset date or resolution date. If the event is considered drug-related regardless of the start date of the event, or the event that is present at baseline but worsens in toxicity grade or is subsequently considered drug-related by the investigator, then this event will be assumed to be treatment emergent. All reported TEAEs will be included in the analysis. Immune-related AEs will be determined by the investigator only. For each AE, the number and percentage of participants who experience at least 1 occurrence of the given event will be summarized overall and by cohort.

Summary tables will be provided for TEAEs:

- AEs
- Serious AEs (SAEs)
- AEs leading to discontinuation of each study drug within a study intervention
- AEs based on NCI-CTCAE toxicity grade
- AEs and SAEs by relationship to each of the following:
  - In Cohorts 1, 2, and 4: TAR-200, TAR-200 insertion procedure, TAR-200 removal procedure, Urinary Placement Catheter (UPC)
  - In Cohorts 1 and 3: cetrelimab,
- AEs leading to dose interruption of each study agent within a study intervention
- AEs with fatal outcome
- Other safety observations (e.g. immune related AEs by preferred term (PT) and toxicity grade)

In addition to the summary tables, listings will be provided for participants who had:

- TEAEs
- SAEs
- TEAEs and related TEAEs leading to discontinuation of study intervention
- TEAEs leading to dose interruption
- TEAEs with fatal outcome
- Immune related TEAEs

Duration of TEAEs and lower urinary tract TEAEs will be summarized by cohort.

A listing of participants who died will be provided.

5.6.3. Clinical Laboratory Tests

Clinical laboratory tests will be displayed for the participants included in the safety analysis set.

Descriptive statistics will be presented for select chemistry and hematology laboratory tests at scheduled time points. Change from baseline to each scheduled time point will be summarized for select chemistry and hematology tests by cohort.

A listing of participants with abnormal chemistry and hematology values will be provided.

Applicable laboratory results will be graded according to NCI-CTCAE version 5.0. The worst toxicity grade will be summarized by cohort. Selected lab tests for chemistry to be analyzed include Alkaline Phosphatase (ALP), Alanine Aminotransferase (ALT), Aspartate Aminotransferase (AST), Bilirubin, Creatinine, C Reactive Protein, lactate dehydrogenase, Lipase, Sodium, Potassium, Glucose, Thyroid Stimulating Hormone (TSH).

Selected lab tests for hematology to be summarized include Hemoglobin, Platelets, White Blood Cell Count, Prothrombin international normalized ratio.

On treatment urine culture results and shift from baseline to post baseline will be summarized at scheduled timepoints by cohort. A listing will be provided.

5.6.4. Electrocardiogram

Electrocardiogram (ECG) measurements are taken at baseline and end of treatment visits. ECG interpretations (clinically significant, not clinically significant) will be summarized at baseline and end of treatment visits by cohort.

Listings will be produced for abnormal or clinically significant ECG data including unscheduled visit data.

5.6.5. Vital Signs

Vital sign parameters including heart rate, weight and blood pressure (systolic and diastolic) will be summarized descriptively at each assessment timepoint by cohort. Descriptive statistics (mean, standard deviation, median, minimum and maximum) of vital signs and the change from baseline will be presented by cohort.

Abnormality criteria (based on criteria defined in Table 7) will be applied to baseline and postbaseline values. For baseline values, increase or decrease criteria are not applied.

Postbaseline values will be considered treatment emergent (TE) if they meet both value and change criteria in the table below. For criteria that do not include an increase or decrease from baseline:

- TE will be concluded if the postbaseline value is above the upper limit and the baseline value is below the upper limit (e.g., Normal or Low). The same applies to the postbaseline value being below the lower limit with the baseline value being above the lower limit (e.g., Normal or High).
- If the baseline value is missing, a postbaseline abnormality will always be considered as TE.

Incidence of markedly abnormal vital signs during intervention, as defined in Table 7, will be summarized for participants who had a baseline assessment and at least 1 postbaseline assessment for that vital sign. A listing of participants with TE abnormal vital signs will be presented.

Table 7: Markedly Abnormal Vital Signs

Vital Sign	Criteria
Pulse	>110 bpm and with >30 bpm increase from baseline
	<50 bpm and with >20 bpm decrease from baseline
Systolic blood pressure	>180 mm Hg and with >40 mm Hg increase from baseline
	<90 mm Hg and with >30 mm Hg decrease from baseline
Diastolic blood pressure	>105 mm Hg and with >30 mm Hg increase from baseline
	<50 mm Hg and with >20 mm Hg decrease from baseline
Weight	Increase 10% kg from baseline

Vital Sign	Criteria
	Decrease 10% kg from baseline

5.7. Other Analyses

Biomarker analyses will be performed if feasible and may be provided in a separate document.

5.7.1. Pharmacokinetics and Immunogenicity

Pharmacokinetic Analyses

TAR-200:

Urine and plasma concentration of gemcitabine and dFdU, and urine amount of gemcitabine and dFdU will be listed (for both sparse and serial urine PK and plasma PK) and summarized (for serial urine PK only) at each timepoint using descriptive statistics (arithmetic mean, standard deviation [SD], coefficient of variation [CV], geometric mean, geometric CV, median, minimum, and maximum). All concentrations below the lowest quantifiable concentration or missing data will be labeled as such in the listing. For serial urine PK, PK parameters will be listed and summarized using descriptive statistics. Details of the analyses will be described in the Clinical Pharmacology Analysis Plan.

Cetrelimab:

Cetrelimab concentrations will be listed and summarized at each timepoint using descriptive statistics (arithmetic mean, SD, CV, geometric mean, geometric CV, median, minimum, and maximum) and figures. All concentrations below the lowest quantifiable concentration or missing data will be labeled as such in the concentration data presentation. Concentrations below the lowest quantifiable concentration will be treated as zero in the summary statistics and for the calculation of PK parameters. All participants and samples excluded from the analysis will be clearly documented in the study report.

5.7.1.1. Immunogenicity of Cetrelimab

The incidence of anti-cetrelimab antibodies will be summarized for all participants who receive at least 1 dose of IV cetrelimab and have appropriate samples for detection of antibodies to cetrelimab (ie, participants with at least 1 sample obtained after their first dose of cetrelimab). A listing of participants who are positive for antibodies to cetrelimab will be provided. The maximum titers of antibodies to cetrelimab will be summarized for participants who are positive for antibodies to cetrelimab. Other immunogenicity analyses may be performed to further characterize the immune responses that are generated.

5.7.1.2. Population PK analysis

If feasible, population PK analysis of serum concentration-time data of IV cetrelimab or/and urine exposure-time data of TAR-200 may be performed using nonlinear mixed-effects modeling. Data may be combined with those of other selected studies to support a relevant structural model or to compare across different patient populations. Available baseline participant characteristics (demographics, laboratory variables, genotypes, race, etc.) may be tested as potential covariates affecting PK parameters. Details will be given in a population PK analysis plan and the results of the population PK analysis will be presented in a separate report.

5.7.1.3. Pharmacokinetic/Pharmacodynamic and Exposure-Response Analyses

Pharmacokinetic/pharmacodynamic models of serum cetrelimab or/and urine exposure-time data of TAR-200 may be explored to understand and characterize the exposure-response relationship for key efficacy, safety, and pharmacodynamic/biomarker data parameters, to detect the influence of covariates, and to identify inter-individual variability in response. The details will be provided in a separate analysis plan and the results of the analyses will be summarized in a separate report.

5.7.2. Medical Resource Utilization and Health Economics

Medical resource utilization (MRU) data, associated with medical encounters, will be collected on an ongoing basis whenever an encounter occurs, including information from regarding utilization of healthcare services (including the type of services). Protocol-mandated procedures, tests, and encounters are excluded.

A descriptive summary of MRU data will be provided, including number and duration of medical encounters, duration of hospitalization and frequency of outpatient medial encounters by treatment cohort.

5.7.3. Patient Reported Outcomes (PRO)

Patient-reported outcomes will be measured by two instruments: the European Organization for Research and Treatment of Cancer (EORTC) Quality of Life Questionnaire (QLQ)-C30 and EORTC-QLQ Non-Muscle-Invasive Bladder Cancer-24 (NMIBC24). The term “each PRO endpoint” refers collectively to the EORTC-QLQ-C30 global health status (GHS) score, functional scale scores, and symptom scale/single item scores and EORTC-QLQ-NMIBC24 scale scores/single item scores.

The EORTC-QLQ-C30 comprises 30 items that form a global health status, 5 functional scales (physical functioning role functioning, emotional functioning, cognitive functioning, and social functioning), 3 symptom scales (fatigue, nausea and vomiting, and pain), and a set of 6 single items (dyspnea, insomnia, appetite loss, constipation, diarrhea, and financial impact). The recall period is the past week (except for the first 5 items in the assessment).

The EORTC-QLQ-NMIBC24 comprises 24 items that cover the domains of urinary symptoms, malaise, future worries, bloating and flatulence, sexual function, and male sexual problems. Additionally, 5 single items are also included for intravesical treatment issues, sexual intimacy, risk of contaminating a partner, sexual enjoyment, and female sexual problems. The recall period is either the past week or past 4 weeks depending on the item.

Each instrument’s user manual is included in Appendix 11.

PRO data will be collected as outlined in the Schedule of Activities in the protocol. During the Treatment Phase, PROs will be collected at Baseline (Week 0/Day 0), Weeks 6, 12, 24, 36, 48, 60, 72, 84 (except Cohort 3), 96 (except Cohort 3), and the end-of-treatment (EOT) visit. PRO measures will also be collected at the 30-day and 100-day safety follow-up visit.

Analyses of each PRO endpoint will be conducted on the full analysis set, separately for each cohort. The analyses of PRO endpoints are considered exploratory and as such, any p-values reported should be considered as nominal. No adjustments will be made for multiplicity. There will be no imputation for missing PRO data.

Each PRO endpoint will be analyzed with descriptive summaries (e.g. mean, SD, minimum, maximum, change from baseline) at each assessment time point and graphical summaries will be provided. Subjects who have a baseline measurement and at least one post-baseline value will be included in the analysis. Where variables are summarized in terms of change from baseline, the baseline mean score of those subjects with non-missing change scores will also be included. A mixed-effects model using repeated measures (MMRM) based on observed data will be performed to obtain the least squares (LS) mean estimates and 95% CI at each assessment time point. Time to symptom deterioration will be analyzed descriptively and estimated using Kaplan-Meier methods. Time to symptom deterioration is defined as the time from enrollment until the date of the first clinically meaningful deterioration in PRO score, or death. A clinically meaningful change in EORTC-QLQ-C30 scores is defined as a decrease in GHS/QoL or functional scale scores, or an increase for symptom scale/single item scores from baseline of ≥10 points[3]. While there are no published clinically meaningful change thresholds for the EORTC QLQ-NMIBC24, a threshold of ≥10 points will also be applied (i.e., decrease of at least 10 for sexual function and sexual enjoyment, or an increase of at least 10 for all other scales/single items). Participants who have not shown a deterioration or have not died at the time of analysis or before subsequent therapy will be censored at their last PRO assessment date.

In addition to PRO endpoint analyses, participant disposition by cohort for each scheduled treatment visit that PROs are completed will be summarized using the following categories:

- PRO assessment expected (participant on therapy)
- PRO assessment not expected due to recurrence or progression
- PRO assessment not expected due to death
- PRO assessment not expected due to other reasons

Compliance rates for the completion of each PRO instrument will be generated by cohort based on the actual number assessments received over the number of expected assessments for each scheduled visit of PRO collection. Special consideration will be made to account for the conditional items in the EORTC-QLQ-NMIBC24.

PRO completion rates among patients who are expected to have PRO assessments will be summarized for each PRO instrument by cohort using the following categories:

- All questions completed
- At least meets minimum requirements for scoring of PRO

5.7.4. Definition of Subgroups

Subgroup analysis will be performed for the selected potential prognostic variables (as listed in Table 8 below) to assess the consistency and robustness of the treatment benefit. The subgroup variables and the cut-off values are subject to change, if warranted, to better represent the data. Subgroup analyses will be performed on the primary endpoint and secondary endpoint of DOR in Cohorts 1-3.

Subgroup analysis will be presented graphically in a forest plot.

Table 8: Subgroup Analyses

Subgroup	Definition
Region	• America: Canada, USA • Asia Pacific: Australia, Japan, South Korea • EMEA: Belgium, France, Germany, Greece, Italy, Netherlands, Portugal, Russia, Spain
Age Group	• <65 • 65 to <75 • >=75
Nicotine Use History	• Current • Former • Never
Tumor Stage at Screening	• CIS (Tis) • CIS + Papillary Disease
Baseline ECOG status	• 0 • >=1
PD-L1 status 1	• CPS >=1 (Positive) • CPS <1 (Negative)
PD-L1 status 2	• CPS >=10 (Positive) • CPS <10 (Negative)
Gender	• Female • Male
Race	• American Indian or Alaska Native • Asian • Black or African American • Native Hawaiian or other Pacific Islander • White
Reason for not receiving RC at screening	• Ineligible • Elected not to undergo RC
Creatinine Clearance	• <60 • >=60
BCG Doses	• 7-9 • 10-14 • >14
BCG Strain	• BCG TICE • BCG Non-TICE

5.8. Interim Analyses

A futility analysis was performed after 40 participants in Cohort 1 had been on study for at least 12 weeks. Any cohort could be terminated for futility if its overall CR rate is less than 20%.

CCI, the IDMC reviewed and evaluated the data from the pre-specified futility analysis and recommended to continue the study with none of the cohorts meeting the futility threshold. Please refer to the Independent Data Monitoring Committee (IDMC) Charter and IDMC SAP for details as to specific outputs that were reviewed by the IDMC at the futility analysis.

5.8.1. Independent Data Monitoring Committee (IDMC) or Other Review Board

An IDMC of at least 2 medical experts in the relevant therapeutic area and at least 1 statistician will be established to monitor data on an ongoing basis to ensure the safety of the subjects enrolled in this study. The safety review will focus on deaths, treatment discontinuations, serious adverse events, Grade ≥3 events, and emerging safety data. Based on the results from these scheduled safety review meetings, the IDMC chair may request additional safety monitoring measures. Throughout the conduct of the study, all individual patient deaths, treatment discontinuations and serious adverse events will be reviewed by the Sponsor's medical monitor and the safety physician on an ongoing basis to identify specific patient safety issues, and the IDMC will be informed of any new potential signals of concern. The plan for monitoring subject safety and evaluating efficacy, and the roles and responsibilities of the IDMC, are detailed in the IDMC Charter.

6. SUPPORTING DOCUMENTATION

6.1. Appendix 1: List of Abbreviations

AE	Adverse event
ATC	Anatomical Therapeutic Chemical
BCG	Bacillus Calmette-Guérin
BMI	Body Mass Index
BPM	Beats per minute
CI	Confidence Interval
CIS	Carcinoma in situ
CR	Complete response
CT/MRI	Computed Tomography/ Magnetic Resonance Imaging
CV	Coefficient of Variation
dFdU	2',2'-difluorodeoxyuridine
DFS	Disease Free Survival
DOR	Duration of response
DNA	Deoxyribonucleic acid
ECG	electrocardiogram
ECOG	Eastern Cooperative Oncology Group
eCRF	electronic case report form
EORTC-QLQ-C30	European Organization for Research and Treatment of Cancer-Quality of Life Questionnaire-Core
EORTC-QLQ-NMIBC-24	European Organization for Research and Treatment of Cancer-Quality of Life Questionnaire-Non-Muscle-Invasive Bladder Cancer-24
F/U	Follow-up
GHS	Global Health Score
HR	High risk
HRQoL	Health related quality of life
ICH	International conference on harmonization
IDMC	Independent Data Monitoring Committee
IV	Intravenous
LS	Least-squares
MedDRA	Medical dictionary for regulatory activities
MMRM	Mixed-effects model using repeated measures
MRU	Medical resource utilization
N	Sample size
NCI-CTCAE	National Cancer Institute - common terminology criteria for adverse events
NMIBC	Non-muscle invasive bladder cancer
OS	Overall survival
PD	Pharmacodynamic
PD	Protocol Deviation
PD-L1	Programmed death-ligand 1
PK	Pharmacokinetic
PRO	Patient-reported outcome
PT	Preferred term
QTc	Corrected QT interval
QTcF	Corrected QT interval using Frederica's formula
RC	Radical Cystectomy
RFS	Recurrence free survival
RNA	Ribonucleic Acid
SAE	Severe adverse event
SAP	Statistical analysis plan
SD	Standard deviation
SOA	Schedule of activities
TE	Treatment emergent
TNM	T Stage, N Stage, M Stage
TURBT	Transurethral resection of bladder tumor

WHO-DD World Health Organization Drug Dictionary

6.2. Appendix 2: Changes to Protocol-Planned Analyses

A side-by-side descriptive summary of efficacy will not be provided as each cohort will be summarized separately.

Time to onset of response will be assessed in Cohorts 1-3 and is defined as the time from first dose to the first determination CR. It will be summarized descriptively (intervention (N, mean, SD, median, and range [minimum, maximum])).

6.3. Appendix 3: Demographics and Baseline Characteristics

The number of participants in each analysis set will be summarized and listed by cohort and overall. In addition, the distribution of participants by region, country, and site ID will be presented.

Tables 9 and Table 10 presents a list of the demographic variables and baseline characteristics that will be summarized by cohort and overall, for the enrolled analysis set.

Table 9: Demographic Variables

Continuous Variables:	**Summary Type**
Age (years)	Descriptive statistics (N, mean, standard deviation [SD], median and range [minimum and maximum], and IQ range).
BMI	
Categorical Variables	
Age (<65 years, 65-74 years, 75-84 years, and >=85 years)	Frequency distribution with the number and percentage of participants in each category.
Age (≥65, ≥75, ≥85)	
Sex (male, female)	
Region[a] (America, Asia Pacific, EMEA)	
Race[b] (American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or other Pacific Islander, White, Multiple)	
Ethnicity (Hispanic or Latino, not Hispanic or Latino)	
BMI ([underweight <18.5 kg/m2, normal 18.5-<25 kg/m2, overweight 25-<30 kg/m2, obese >=30 kg/m2])	
Nicotine Use History (Current, former, never)	

[a]America includes Canada, USA; Asia Pacific includes Australia, Japan, South Korea; Europe/Middle East/Africa (EMEA) includes Belgium, France, Germany, Greece, Italy, Netherlands, Portugal, Russia, Spain
[b]If multiple race categories are indicated, the Race is recorded as 'Multiple'

Table 10: Baseline Characteristics

Continuous Variables	**Summary Type**
Time from last BCG dose to diagnosis (months)	Descriptive statistics (N, mean, standard deviation [SD], median and range [minimum and maximum], and IQ range).
Categorical Variables	

Baseline ECOG (0,1,2)	Frequency distribution with the number and percentage of participants in each category.
Tumor Stage (CIS, CIS+papillary disease [CIS+Ta, CIS+T1])	
PD-L1 status 1 (CPS<1 (Negative), CPS>=1 (Positive), indeterminate)	
PD-L1 status 2 (CPS<10 (Negative), CPS>=10 (Positive), indeterminate)	
Reason for not receiving RC (ineligible, elected not to undergo RC)	

6.4. Appendix 4: Protocol Deviations

In general, the following list of major protocol deviations (PD) may have the potential to impact participants' rights, safety or well-being, or the integrity and/or the primary endpoint of the clinical study. Participants whose major protocol deviations are determined to meet the above criteria (have the potential to impact participants' rights, safety or well-being, or the integrity and/or the primary endpoint of the clinical study) will be identified by medical and statistical review prior to database lock and will be summarized by the following categories:

- Developed withdrawal criteria but not withdrawn
- Entered but did not satisfy eligibility criteria
- Received a disallowed concomitant treatment
- Received wrong treatment or incorrect dose
- Other medically important PDs identified during medical and statistical PD review

A Protocol Deviation Specification (PDS) has been developed to provide more information about the major protocol deviations. Periodic meetings are required to investigate each potential protocol deviation.

6.5. Appendix 5: Prior, Concomitant Medications and Subsequent Therapies

Prior and Concomitant medications will be coded using the World Health Organization Drug Dictionary (WHO-DD). Prior medications are defined as any therapy used before the day of first dose (partial or complete) of study intervention. Concomitant medications are defined as any therapy used on or after the same day as the first dose of study drug, including those that started before and continue on after the first dose of study drug.

Summaries of concomitant medications will be presented by ATC term, cohort, and intervention phase. The proportion of participants who receive each concomitant medication will be summarized as well as the proportion of participants who receive at least 1 concomitant medication.

Prior medications will be summarized by cohort and ATC term.

Prior therapies for bladder cancer will be summarized by cohort.

These summaries (n and %) may include:

- Total number of participants with any prior therapies for bladder cancer
  - Prior systemic therapy
  - Prior BCG therapy
  - Prior intravesical therapy (other than BCG)
  - Prior NMIBC cancer-related surgery/procedure
- Number of prior intravesical BCG doses and induction and second induction/maintenance doses of BCG

Percentages will be calculated with the number of participants in each cohort as the denominator.

The total number of participants who received subsequent anticancer therapy will be reported for each cohort. A summary of first subsequent anticancer therapy or radiotherapy and selected subsequent surgeries/procedures (i.e. radical cystectomy) will be presented by chemical class, pharmacologic class and drug name, coded using the WHO-DD for each cohort. A listing will be provided.

6.6. Appendix 6: Medical History

A listing for participant medical history will be provided.

6.7. Appendix 7: Intervention Compliance

Compliance will be summarized descriptively for each study drug within a study intervention.

Compliance will be calculated as follows:

For TAR-200: Compliance (%) = number of TAR-200 insertions completed/the number of TAR-200 insertions expected through CCO x100, during the treatment period

For Cetrelimab: Compliance (%) = number of cetrelimab infusions completed/the number of cetrelimab infusions expected through CCO x100, during the treatment period

The maximum number of TAR-200 insertions and cetrelimab infusions are 14 and 27, respectively.

6.8. Appendix 8: Adverse Events of Special Interest

There are no AEs of special interest.

6.9. Appendix 9: Medications of Special Interest

Concomitant medications of special interest include steroids and immunosuppressive medications used to treat immune-related adverse events related to cetrelimab. A listing will be provided.

6.10. Appendix 10: Laboratory Toxicity Grading

The grading scale use for lab assessments is based on 'Common Terminology Criteria for Adverse Events (CTCAE) v5.0'.

If a laboratory value falls within the grading as specified below but also within the local laboratory normal limits, the value is considered to be normal and will be reset to grade 0.

Pre-baseline measurements will use the same grading ranges as applied to baseline measurements. In case a test has two sets of ranges – one for baseline normal and one for baseline abnormal, the one for baseline normal will be applied for all measurements taken pre-baseline and on baseline.

Text in gray italic in the table is present in the grading scale, but is not applied by JNJ when grading lab data.

CTCAE Term	Grade 1	Grade 2	Grade 3	Grade 4	Janssen implementation notes
Blood and lymphatic system disorders					
Anemia	Hemoglobin (Hgb) <LLN - 10.0 g/dL; <LLN - 6.2 mmol/L; <LLN - 100 g/L	Hemoglobin (Hgb) <10.0 - 8.0 g/dL; <6.2 - 4.9 mmol/L; <100 - 80g/L	Hemoglobin (Hgb) <8.0 g/dL; <4.9 mmol/L; <80 g/L; *transfusion indicated*	*Life-threatening consequences; urgent intervention indicated*	Clinical signs and symptoms are not taken into consideration for grading.
Leukocytosis	-	-	>100,000/mm3; >100 x 10e9 /L	*Clinical manifestations of leucostasis; urgent intervention indicated*	Clinical signs and symptoms are not taken into consideration for grading; Added ranges in SI unit (x 10e9 /L)
Investigations					
Activated partial thromboplastin time prolonged	>ULN - 1.5 x ULN	>1.5 - 2.5 x ULN	>2.5 x ULN; *bleeding*	-	Clinical signs and symptoms are not taken into consideration for grading.
Alanine aminotransferase increased	>ULN - 3.0 x ULN if baseline was normal; 1.5 - 3.0 x baseline if baseline was abnormal	>3.0 - 5.0 x ULN if baseline was normal; >3.0 - 5.0 x baseline if baseline was abnormal	>5.0 - 20.0 x ULN if baseline was normal; >5.0 - 20.0 x baseline if baseline was abnormal	>20.0 x ULN if baseline was normal; >20.0 x baseline if baseline was abnormal	Ranges defined for "abnormal baseline" are applied only if baseline > ULN. If baseline < LLN, then ranges for "normal baseline" are applied.
Alkaline phosphatase increased	>ULN - 2.5 x ULN if baseline was normal; 2.0 - 2.5 x baseline if baseline was abnormal	>2.5 - 5.0 x ULN if baseline was normal; >2.5 - 5.0 x baseline if baseline was abnormal	>5.0 - 20.0 x ULN if baseline was normal; >5.0 - 20.0 x baseline if baseline was abnormal	>20.0 x ULN if baseline was normal; >20.0 x baseline if baseline was abnormal	Ranges defined for "abnormal baseline" are applied only if baseline > ULN. If baseline < LLN, then ranges for "normal baseline" are applied.
Aspartate aminotransferase increased	>ULN - 3.0 x ULN if baseline was normal; 1.5 - 3.0 x baseline if baseline was abnormal	>3.0 - 5.0 x ULN if baseline was normal; >3.0 - 5.0 x baseline if baseline was abnormal	>5.0 - 20.0 x ULN if baseline was normal; >5.0 - 20.0 x baseline if baseline was abnormal	>20.0 x ULN if baseline was normal; >20.0 x baseline if baseline was abnormal	Ranges defined for "abnormal baseline" are applied only if baseline > ULN. If baseline < LLN, then ranges for "normal baseline" are applied.
Blood bilirubin increased	>ULN - 1.5 x ULN if baseline was normal;	>1.5 - 3.0 x ULN if baseline was normal;	>3.0 - 10.0 x ULN if baseline was normal;	>10.0 x ULN if baseline was normal;	Ranges defined for "abnormal baseline" are

CTCAE Term	Grade 1	Grade 2	Grade 3	Grade 4	Janssen implementation notes
	> 1.0 - 1.5 x baseline if baseline was abnormal	>1.5 - 3.0 x baseline if baseline was abnormal	>3.0 - 10.0 x baseline if baseline was abnormal	>10.0 x baseline if baseline was abnormal	applied only if baseline > ULN. If baseline < LLN, then ranges for "normal baseline" are applied.
CD4 lymphocytes decreased	<LLN - 500/mm3; <LLN - 0.5 x 10e9 /L	<500 - 200/mm3; <0.5 - 0.2 x 10e9 /L	<200 - 50/mm3; <0.2 x 0.05 - 10e9 /L	<50/mm3; <0.05 x 10e9 /L	
Cholesterol high	>ULN - 300 mg/dL; >ULN - 7.75 mmol/L	>300 - 400 mg/dL; >7.75 - 10.34 mmol/L	>400 - 500 mg/dL; >10.34 - 12.92 mmol/L	>500 mg/dL; >12.92 mmol/L	
CPK increased	>ULN - 2.5 x ULN	>2.5 x ULN - 5 x ULN	>5 x ULN - 10 x ULN	>10 x ULN	
Creatinine increased	Creatine Kinase >ULN - 1.5 x ULN	Creatine Kinase >1.5 - 3.0 x baseline; >1.5 - 3.0 x ULN	Creatine Kinase >3.0 x baseline; >3.0 - 6.0 x ULN	Creatine Kinase >6.0 x ULN	
Fibrinogen decreased	<1.0 - 0.75 x LLN; if abnormal, <25% decrease from baseline	<0.75 - 0.5 x LLN; if abnormal, 25 - <50% decrease from baseline	<0.5 - 0.25 x LLN; if abnormal, 50 - <75% decrease from baseline	<0.25 x LLN; if abnormal, 75% decrease from baseline; absolute value <50 mg/dL	Ranges defined for "abnormal" are applied only on values < LLN. Grade 0 will be assigned to values > ULN.
GGT increased	>ULN - 2.5 x ULN if baseline was normal; 2.0 - 2.5 x baseline if baseline was abnormal	>2.5 - 5.0 x ULN if baseline was normal; >2.5 - 5.0 x baseline if baseline was abnormal	>5.0 - 20.0 x ULN if baseline was normal; >5.0 - 20.0 x baseline if baseline was abnormal	>20.0 x ULN if baseline was normal; >20.0 x baseline if baseline was abnormal	Ranges defined for "abnormal baseline" are applied only if baseline > ULN. If baseline < LLN, then ranges for "normal baseline" are applied.
Haptoglobin decreased	<LLN	-	-	-	
Hemoglobin increased	Increase in >0 - 2 g/dL; Increase in >0 - 20 g/L	Increase in >2 - 4 g/dL; Increase in >20 - 40 g/L	Increase in >4 g/dL; Increase in >40 g/L	-	The increase indicates the level of increase above normal (above ULN). Applied as, e.g. grade 1 (g/dL): >ULN – ULN+2 g/dL; Added ranges in SI unit (g/L).
INR increased	>1.2 - 1.5; *>1 - 1.5 x baseline if on anticoagulation; monitoring only indicated*	>1.5 - 2.5; *>1.5 - 2.5 x baseline if on anticoagulation; dose adjustment indicated*	>2.5; *>2.5 x baseline if on anticoagulation; bleeding*	-	Concomitant therapy or clinical signs and symptoms are not taken into consideration for grading.

CTCAE Term	Grade 1	Grade 2	Grade 3	Grade 4	Janssen implementation notes
Lipase increased	>ULN - 1.5 x ULN	>1.5 - 2.0 x ULN; *>2.0 - 5.0 x ULN and asymptomatic*	>2.0 - 5.0 x ULN with signs or symptoms; *>5.0 x ULN and asymptomatic*	>5.0 x ULN and with signs or symptoms	"Asymptomatic" ranges are not taken into consideration for grading, i.e. worst case grading is applied.
Lymphocyte count decreased	<LLN - 800/mm3; <LLN - 0.8 x 10e9/L	<800 - 500/mm3; <0.8 - 0.5 x 10e9 /L	<500 - 200/mm3; <0.5 - 0.2 x 10e9 /L	<200/mm3; <0.2 x 10e9 /L	
Lymphocyte count increased	-	>4000/mm3 - 20,000/mm3; >4 - 20 x 10e9 /L	>20,000/mm3; >20 x 10e9 /L	-	Added ranges in SI unit (x 10e9 /L).
Neutrophil count decreased	<LLN - 1500/mm3; <LLN - 1.5 x 10e9 /L	<1500 - 1000/mm3; <1.5 - 1.0 x 10e9 /L	<1000 - 500/mm3; <1.0 - 0.5 x 10e9 /L	<500/mm3; <0.5 x 10e9 /L	Both Neutrophils and segmented neutrophils are graded using these criteria.
Platelet count decreased	<LLN - 75,000/mm3; <LLN - 75.0 x 10e9 /L	<75,000 - 50,000/mm3; <75.0 - 50.0 x 10e9 /L	<50,000 - 25,000/mm3; <50.0 - 25.0 x 10e9 /L	<25,000/mm3; <25.0 x 10e9 /L	
Serum amylase increased	>ULN - 1.5 x ULN	>1.5 - 2.0 x ULN; *>2.0 - 5.0 x ULN and asymptomatic*	>2.0 - 5.0 x ULN *with signs or symptoms; >5.0 x ULN and asymptomatic*	>5.0 x ULN *and with signs or symptoms*	"Asymptomatic" ranges are not taken into consideration for grading, i.e. worst case grading is applied.
White blood cell decreased	<LLN - 3000/mm3; <LLN - 3.0 x 10e9 /L	<3000 - 2000/mm3; <3.0 - 2.0 x 10e9 /L	<2000 - 1000/mm3; <2.0 - 1.0 x 10e9 /L	<1000/mm3; <1.0 x 10e9 /L	
Metabolism and nutrition disorders					
Acidosis	pH <normal, but >=7.3	-	pH <7.3	*Life-threatening consequences*	pH <normal is implemented as pH <LLN. Clinical signs and symptoms are not taken into consideration for grading.
Alkalosis	pH >normal, but <=7.5	-	pH >7.5	*Life-threatening consequences*	pH >normal is implemented as pH >ULN. Clinical signs and symptoms are not taken into consideration for grading.

CTCAE Term	Grade 1	Grade 2	Grade 3	Grade 4	Janssen implementation notes
Hypercalcemia	Corrected serum calcium of >ULN - 11.5 mg/dL; >ULN - 2.9 mmol/L; Ionized calcium >ULN - 1.5 mmol/L	Corrected serum calcium of >11.5 - 12.5 mg/dL; >2.9 - 3.1 mmol/L; Ionized calcium >1.5 - 1.6 mmol/L; *symptomatic*	Corrected serum calcium of >12.5 - 13.5 mg/dL; >3.1 - 3.4 mmol/L; Ionized calcium >1.6 - 1.8 mmol/L; *hospitalization indicated*	Corrected serum calcium of >13.5 mg/dL; >3.4 mmol/L; Ionized calcium >1.8 mmol/L; *life-threatening consequences*	Clinical signs and symptoms are not taken into consideration for grading.
Hyperkalemia	Potassium >ULN - 5.5 mmol/L	Potassium >5.5 - 6.0 mmol/L; *intervention initiated*	Potassium >6.0 - 7.0 mmol/L; *hospitalization indicated*	Potassium >7.0 mmol/L; *life-threatening consequences*	Clinical signs and symptoms are not taken into consideration for grading.
Hypermagnesemia	Magnesium >ULN - 3.0 mg/dL; >ULN - 1.23 mmol/L	-	Magnesium >3.0 - 8.0 mg/dL; >1.23 - 3.30 mmol/L	Magnesium >8.0 mg/dL; >3.30 mmol/L; *life-threatening consequences*	Clinical signs and symptoms are not taken into consideration for grading.
Hypernatremia	Sodium >ULN - 150 mmol/L	Sodium >150 - 155 mmol/L; *intervention initiated*	Sodium >155 - 160 mmol/L; *hospitalization indicated*	Sodium >160 mmol/L; *life-threatening consequences*	Clinical signs and symptoms are not taken into consideration for grading.
Hypertriglyceridemia	Triglycerides 150 mg/dL - 300 mg/dL; 1.71 mmol/L - 3.42 mmol/L	Triglycerides >300 mg/dL - 500 mg/dL; >3.42 mmol/L - 5.7 mmol/L	Triglycerides >500 mg/dL - 1000 mg/dL; >5.7 mmol/L - 11.4 mmol/L	Triglycerides >1000 mg/dL; >11.4 mmol/L; *life-threatening consequences*	Clinical signs and symptoms are not taken into consideration for grading.
Hypoalbuminemia	Albumin <LLN - 3 g/dL; <LLN - 30 g/L	Albumin <3 - 2 g/dL; <30 - 20 g/L	Albumin <2 g/dL; <20 g/L	*Life-threatening consequences; urgent intervention indicated*	Clinical signs and symptoms are not taken into consideration for grading.
Hypocalcemia	Corrected serum calcium of <LLN - 8.0 mg/dL; <LLN - 2.0 mmol/L; Ionized calcium <LLN - 1.0 mmol/L	Corrected serum calcium of <8.0 - 7.0 mg/dL; <2.0 - 1.75 mmol/L; Ionized calcium <1.0 - 0.9 mmol/L;	Corrected serum calcium of <7.0 - 6.0 mg/dL; <1.75 - 1.5 mmol/L; Ionized calcium <0.9 - 0.8 mmol/L;	Corrected serum calcium of <6.0 mg/dL; <1.5 mmol/L; Ionized calcium <0.8 mmol/L;	Clinical signs and symptoms are not taken into consideration for grading.

CTCAE Term	Grade 1	Grade 2	Grade 3	Grade 4	Janssen implementation notes
		symptomatic	*hospitalization indicated*	*life-threatening consequences*	
Hypoglycemia	Glucose <LLN - 55 mg/dL; <LLN - 3.0 mmol/L	Glucose <55 - 40 mg/dL; <3.0 - 2.2 mmol/L	Glucose <40 - 30 mg/dL; <2.2 - 1.7 mmol/L	Glucose <30 mg/dL; <1.7 mmol/L; *life-threatening consequences; seizures*	Clinical signs and symptoms are not taken into consideration for grading. Urine glucose is not graded.
Hypokalemia	*Potassium <LLN - 3.0 mmol/L*	*Symptomatic with* Potassium <LLN - 3.0 mmol/L; intervention indicated	Potassium <3.0 - 2.5 mmol/L; *hospitalization indicated*	Potassium <2.5 mmol/L; *life-threatening consequences*	"Symptomatic" ranges are applied for grade 2, grade 1 not assigned, i.e. worst case applied. Clinical signs and symptoms are not taken into consideration for grading of grade 3 and 4.
Hypomagnesemia	Magnesium <LLN - 1.2 mg/dL; <LLN - 0.5 mmol/L	Magnesium <1.2 - 0.9 mg/dL; <0.5 - 0.4 mmol/L	Magnesium <0.9 - 0.7 mg/dL; <0.4 - 0.3 mmol/L	Magnesium <0.7 mg/dL; <0.3 mmol/L; *life-threatening consequences*	Clinical signs and symptoms are not taken into consideration for grading.
Hyponatremia	Sodium <LLN - 130 mmol/L	*Sodium 125-129 mmol/L and asymptomatic*	*Sodium 125-129 mmol/L symptomatic; 120-124 mmol/L regardless of symptoms* Sodium <130-120 mmol/L	Sodium <120 mmol/L; *life-threatening consequences*	Clinical signs and symptoms are not taken into consideration for grading. Worst case ("<130-120 mmol/L" for grade 3 added by Janssen) is applied across grade 2/3 ranges: 120-129 mol/L assigned to grade 3, grade 2 not used.
Renal and urinary disorders					
Proteinuria	1+ proteinuria; urinary protein ≥ULN - <1.0 g/24 hrs; urinary protein ≥ULN - <1000 mg/day	**Adult:** 2+ and 3+ proteinuria; urinary protein 1.0 - <3.5 g/24 hrs;	**Adult:** 4+ proteinuria; urinary protein >=3.5 g/24 hrs;	-	In case both 24-h urine collection and dipstick are collected, then worst case is taken, as opposed to having 24-h urine

CTCAE Term	Grade 1	Grade 2	Grade 3	Grade 4	Janssen implementation notes
		urinary protein 1000 - <3500 mg/day **Pediatric:** Urine P/C (Protein/Creatinine) ratio 0.5 - 1.9; Urine P/C (Protein/Creatinine) 56.5 – 214.7 g/mol	urinary protein >=3500 mg/day; **Pediatric:** Urine P/C (Protein/Creatinine) ratio >1.9; Urine P/C (Protein/Creatinine) >214.7 g/mol		collection take precedence over dipstick. Added ranges in SI unit for urinary protein (mg/day) and for urine P/C (g/mol). Pediatric grading is applied to age range [0-18]. Adult grading is applied for ages [>18].

* Grade 0 is assigned to a lab assessment when the lab test is described in the table, but the lab value is not assigned a grade 1 or higher.

6.11. Appendix 11: Patient-Reported Outcomes User Manuals

EORTC QLQ-C30 Scoring Manual

The EORTC QLQ-C30

Introduction

The EORTC quality of life questionnaire (QLQ) is an integrated system for assessing the health-related quality of life (QoL) of cancer patients participating in international clinical trials. The core questionnaire, the QLQ-C30, is the product of more than a decade of collaborative research. Following its general release in 1993, the QLQ-C30 has been used in a wide range of cancer clinical trials, by a large number of research groups; it has additionally been used in various other, non-trial studies.

This manual contains scoring procedures for the QLQ-C30 versions 1.0, (+3), 2.0 and 3.0; it also contains summary information about supplementary modules.

All publications relating to the QLQ should use the scoring procedures described in this manual.

This manual will be updated at regular intervals, to reflect future changes to the QLQ and to incorporate new supplementary modules.

Background

The EORTC

The European Organisation for Research and Treatment of Cancer (EORTC) was founded in 1962, as an international non-profit organisation. The aims of the EORTC are to conduct, develop, co-ordinate and stimulate cancer research in Europe by multidisciplinary groups of oncologists and basic scientists. Research is accomplished mainly through the execution of large, prospective, randomised, multicentre, cancer clinical trials.

The EORTC Central Office Data Center, created in 1974, is concerned with all aspects of phase II and phase III cancer clinical trials, from their design to the publication of the final results. Since its inception, over 80,000 patients have been entered in trials handled by the EORTC Data Center.

In 1980, the EORTC created the Quality of Life Group, which in 1986 initiated a research programme to develop an integrated, modular approach for evaluating the QoL of patients participating in cancer clinical trials. This led to the development of the EORTC QLQ-C30, a quality of life instrument for cancer patients. To date, more than 2200 studies using the QLQ-C30 have been registered.

EORTC QLQ-C36

A first generation core questionnaire, the EORTC QLQ-C36, was developed in 1987. This 36-item questionnaire was designed to be (1) cancer specific, (2) multidimensional in structure, (3) appropriate for self-administration (i.e. brief and easy to complete), and (4) applicable across a range of cultural settings. Detailed results of the international field-testing of the EORTC QLQ-C36 have been reported (Aaronson *et al.*, 1991). While the overall psychometric results were promising, they also pointed to some areas in which the questionnaire could benefit from further development. Most of the revision involved only minor changes in the wording of items. A few items were found to be non-informative, and were discarded. The only scale requiring substantial revision, because of inadequate reliability, was the eight-item emotional functioning scale. In the next generation of the instrument, this scale was substituted by a four-item emotional functioning scale that had been used previously in EORTC clinical trials.

EORTC QLQ-C30 version 1.0

A second generation core questionnaire, the first version of the 30-item EORTC QLQ-C30 (Appendix 1a), was subsequently developed. The content areas covered by the questionnaire reflect the multi-dimensionality of the QoL construct. This questionnaire was field tested in a cross-cultural sample of lung cancer patients in 13 countries to confirm the hypothesised scale structure, to establish reliability and to evaluate validity (Aaronson *et al.*, 1993).

The QLQ-C30 version 1.0 (QLQ-C30(V1)) incorporates five functional scales (physical, role, cognitive, emotional, and social), three symptom scales (fatigue, pain, and nausea and vomiting), a global health status / QoL scale, and a number of single items assessing additional symptoms commonly reported by cancer patients (dyspnoea, loss of appetite, insomnia, constipation and diarrhoea) and perceived financial impact of the disease.

EORTC QLQ-C30 (+3)

The third generation core questionnaire, the 33-item EORTC QLQ-C30(+3) (Appendix 1b), arose following international testing of the QLQ-C30(V1), when refinement of the questionnaire by adding three new test items was recommended. Two of these test items (QLQ-C30(+3)/Q_{26}, and QLQ-C30(+3)/Q_{27}) were introduced as possible alternatives to the two-item role functioning scale (QLQ-C30(V1)/Q_6, QLQ-C30(V1)/Q_7), which was found to have sub-optimal internal consistency in previous studies. The third new test item, overall health (QLQ-C30(+3)/Q_{32}), was evaluated as a possible replacement for the overall physical condition item (QLQ-C30(V1)/Q_{29}) in the global health status / QoL scale, and employed the same 7-point response scale as the other two questions in that scale.

EORTC QLQ-C30 version 2.0

The QLQ-C30(+3) was an interim version, which retained all the original questions of the QLQ-C30 version 1.0 while evaluating the additional three items. There was a marked improvement in the internal consistency of the new role functioning scale. The new overall health item places less emphasis upon physical functioning, and did not alter the internal consistency. Having formally validated these new items, the older questions were replaced by the new ones (Osoba *et al.*, 1997). The result was the 30-item version 2.0 of the QLQ, the QLQ-C30(V2) (Appendix 1c).

EORTC QLQ-C30 version 3.0

Version 3.0 of the QLQ-C30 differs from version 2.0 in that it has four-point scales for the first five items (QLQ-C30(V3), Appendix 1d). These are coded with the same response categories as items 6 to 28, namely "Not at all", "A little", "Quite a bit" and "Very much." To allow for these categories, question 4 has been re-worded as "Do you have to stay in a bed or a chair during the day?" Version 3.0 has been tested in EORTC field studies (Bjordal *et al.*, 2000).

Version 3.0 is currently the standard version of the QLQ-C30, and should be used for all new studies unless investigators wish to maintain compatibility with previous studies, which used an earlier version of the QLQ-C30.

Latest information about development of the QLQ-C30 and its modules may be found on the EORTC Quality of Life web pages, at:
http://www.eortc.be/home/qol/

Citation and Availability

Citation in published reports

Any publications which describe the use of the EORTC QLQ-C30 or its modules, or which describe analyses of data arising from application of these questionnaires, should explicitly cite the following reference:

> Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ, Filiberti A, Flechtner H, Fleishman SB, de Haes JCJM, Kaasa S, Klee MC, Osoba D, Razavi D, Rofe PB, Schraub S, Sneeuw KCA, Sullivan M, Takeda F.
> The European Organisation for Research and Treatment of Cancer QLQ-C30: A quality-of-life instrument for use in international clinical trials in oncology.
> *Journal of the National Cancer Institute* 1993; **85:** 365-376.

For details of the scoring procedure, a suggested format of citation for this manual is:

> Fayers PM, Aaronson NK, Bjordal K, Groenvold M, Curran D, Bottomley A, on behalf of the EORTC Quality of Life Group.
> *The EORTC QLQ-C30 Scoring Manual (3rd Edition).*
> Published by: European Organisation for Research and Treatment of Cancer, Brussels 2001.

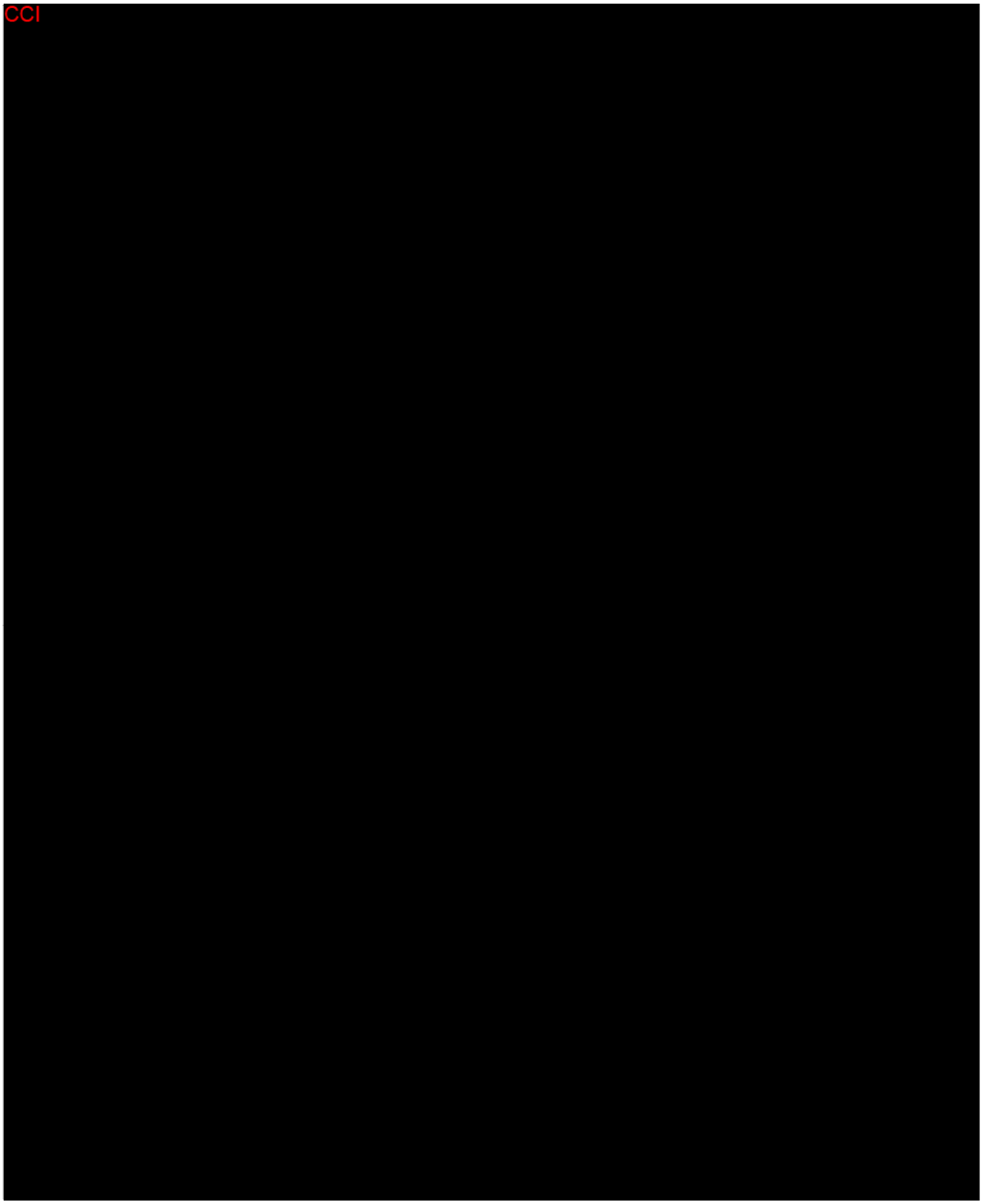
Contact address

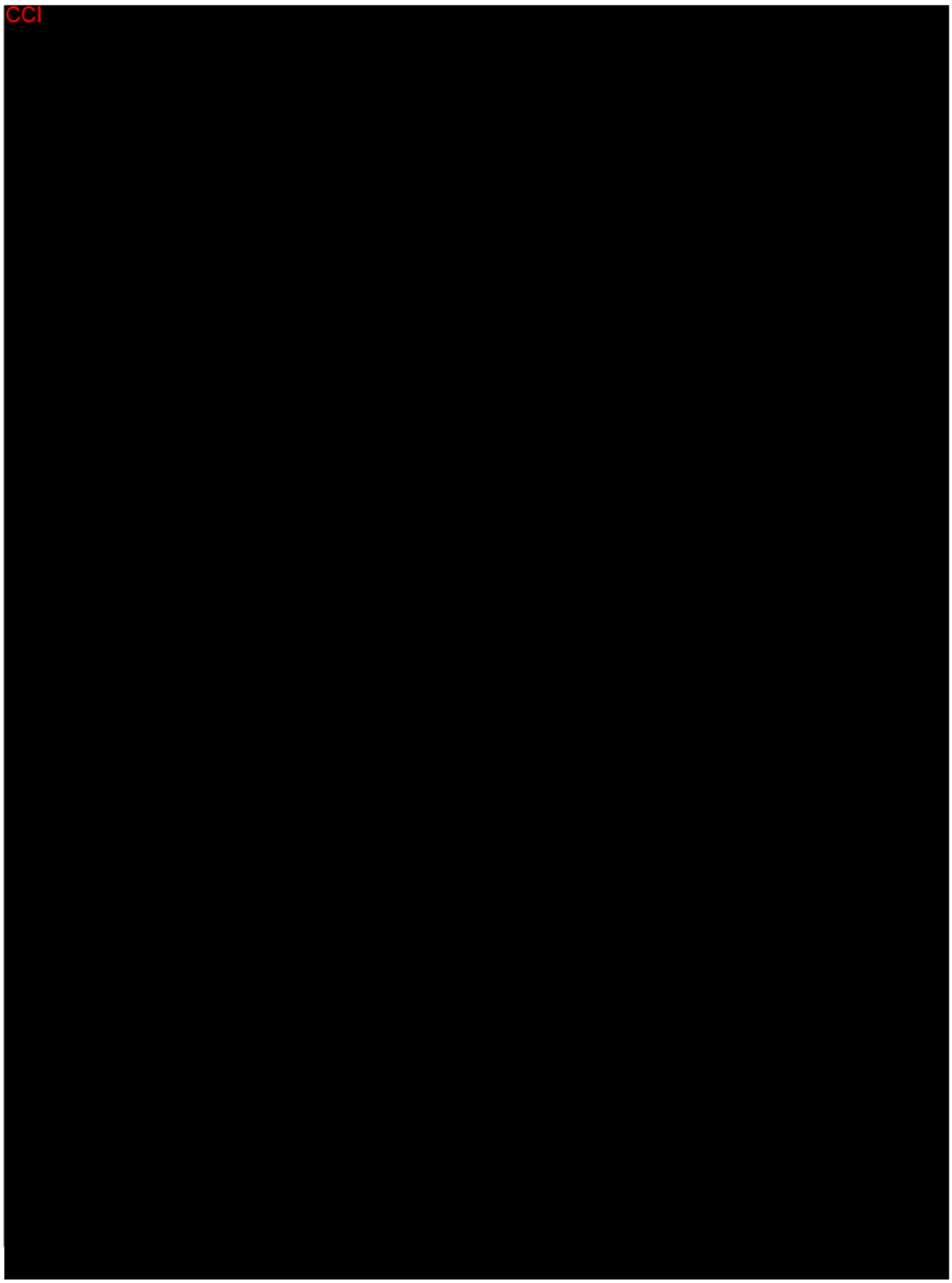
For information about terms and conditions for using the questionnaire, please contact the Quality of Life Unit, EORTC Data Center.

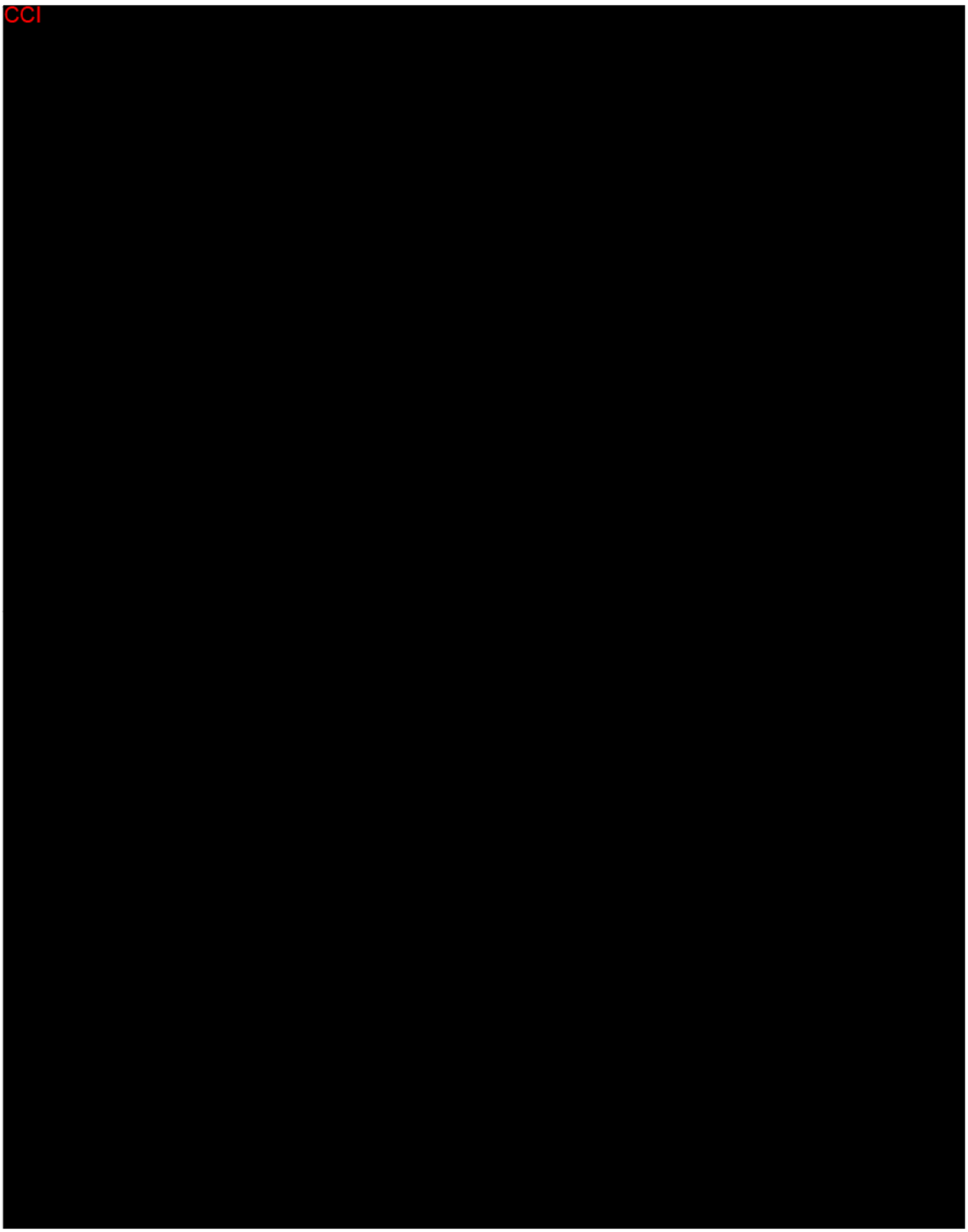
QL Coordinator,
Quality of Life Unit,
EORTC Data Center,
Avenue E Mounier 83 - B11,
1200 Brussels,
BELGIUM

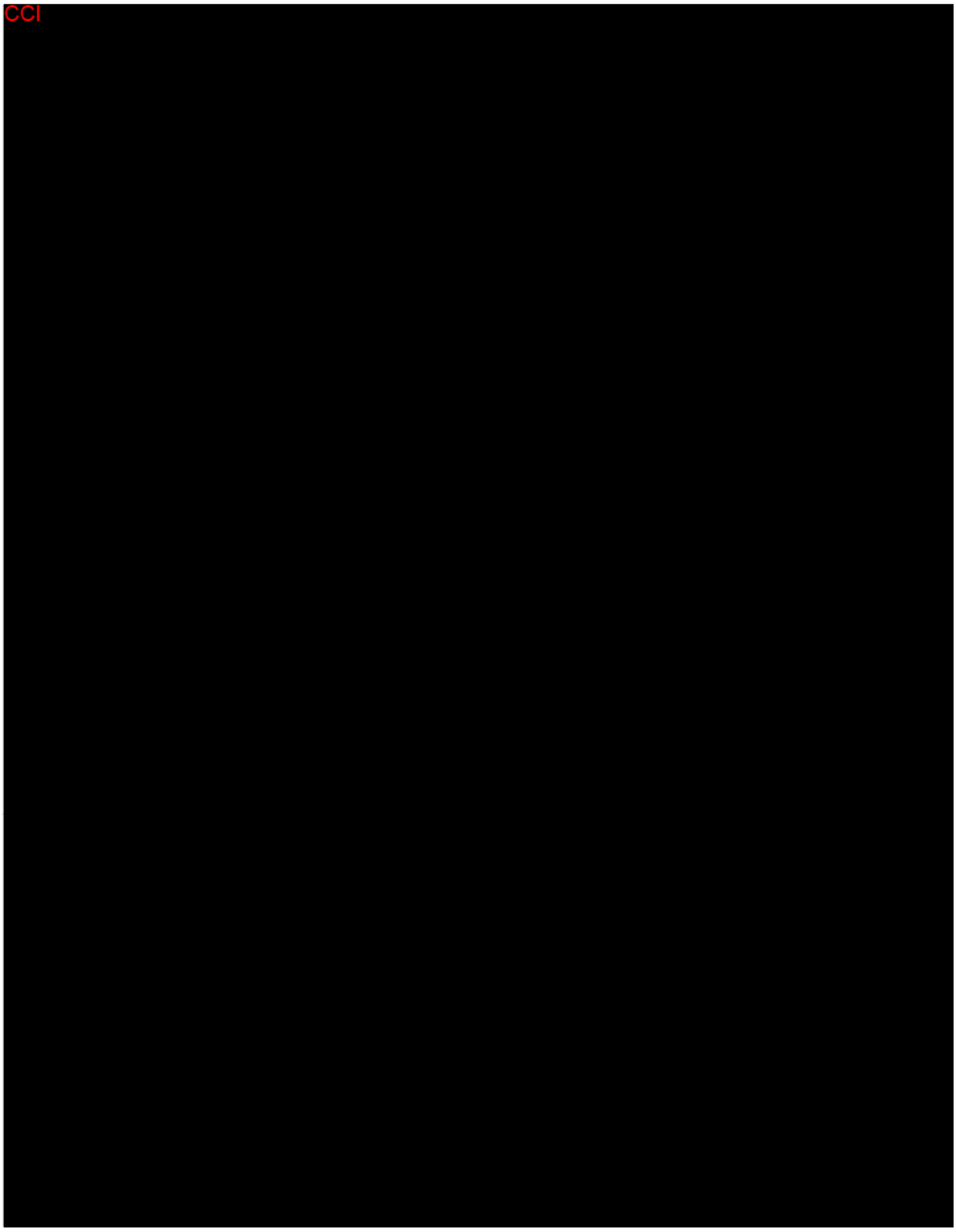
Tel: +32 2 774 1611
Fax: +32 2 779 4568
Email: abo@eortc.be

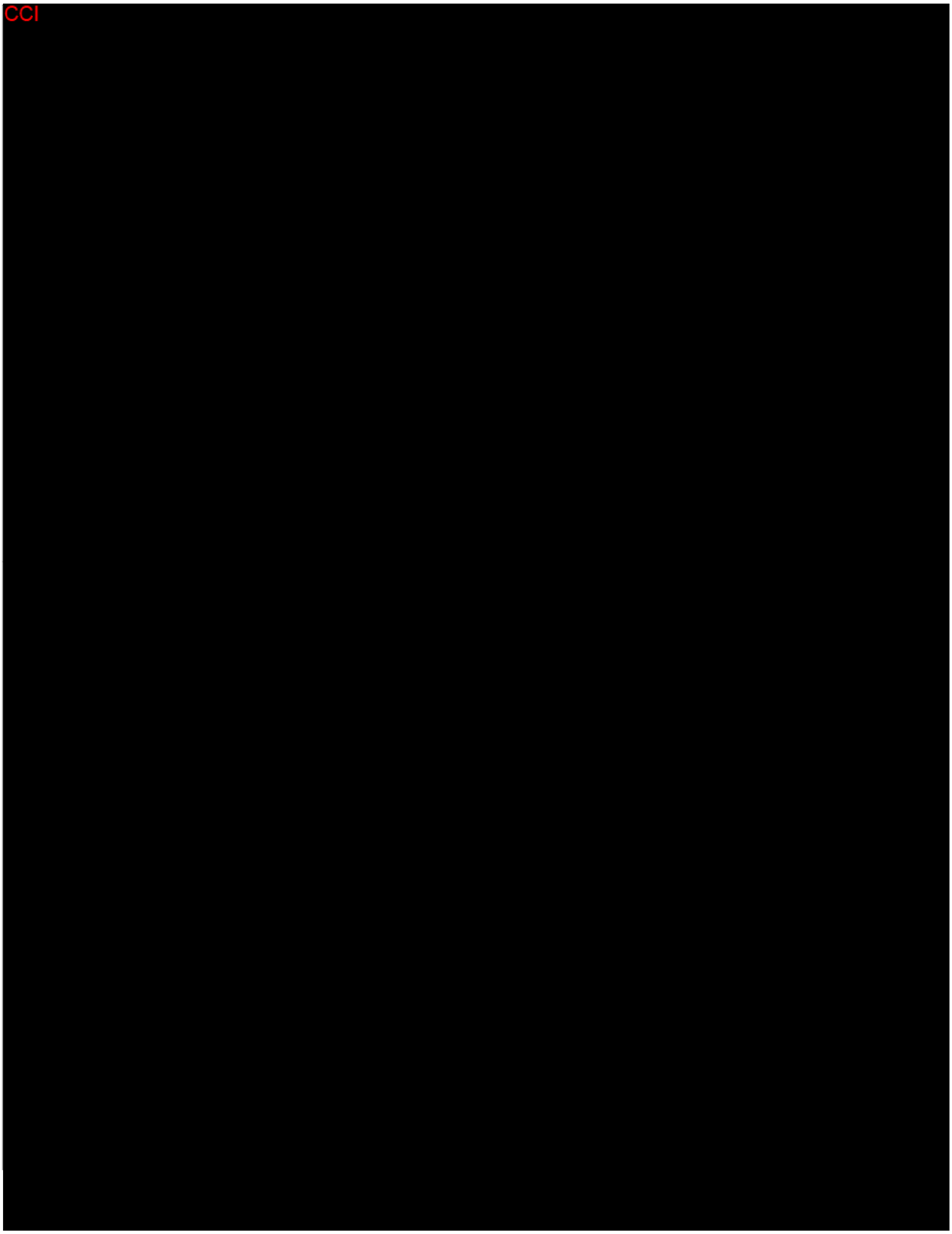
Scoring procedures


CCI


CCI

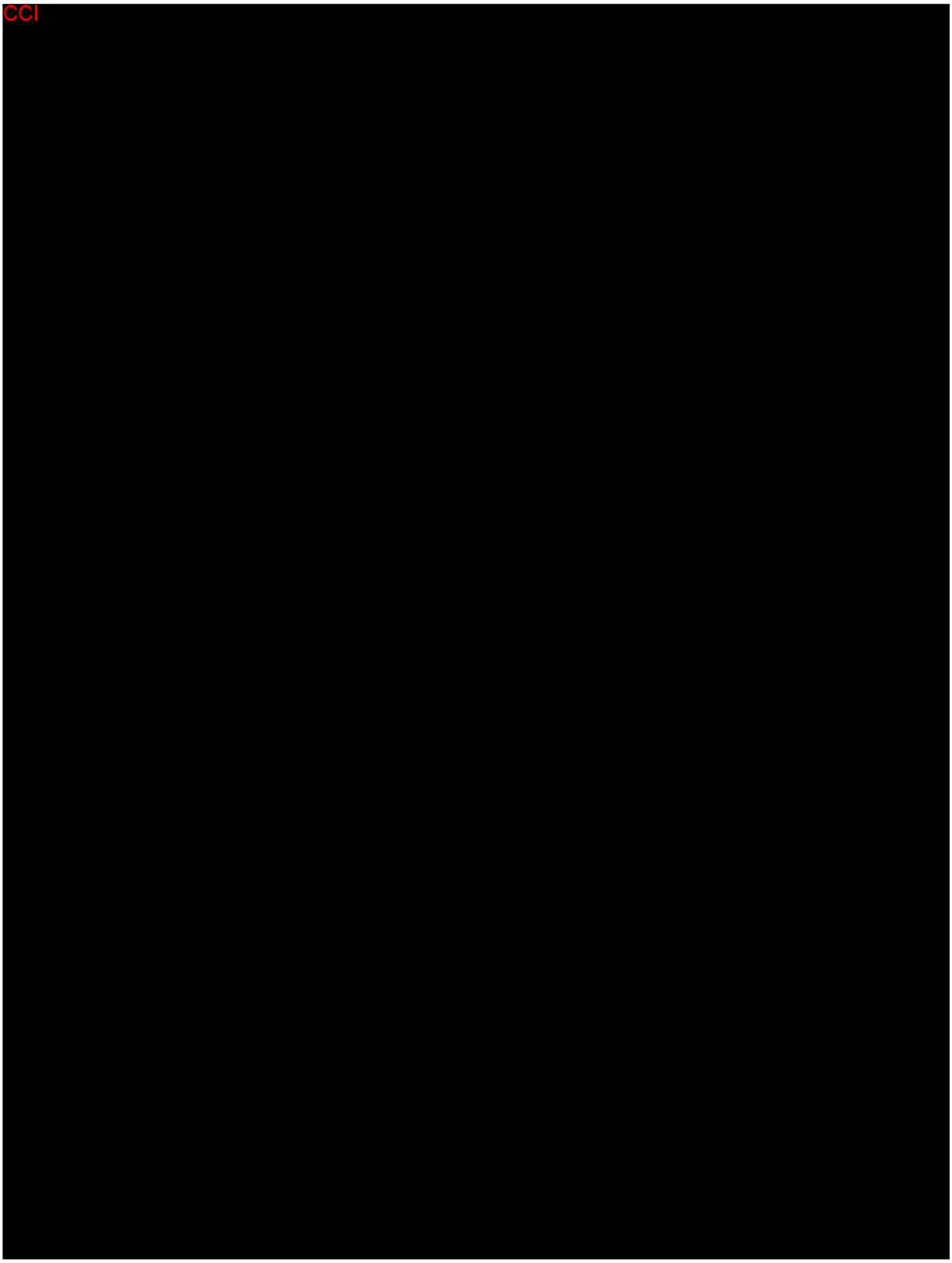

CCI


CCI

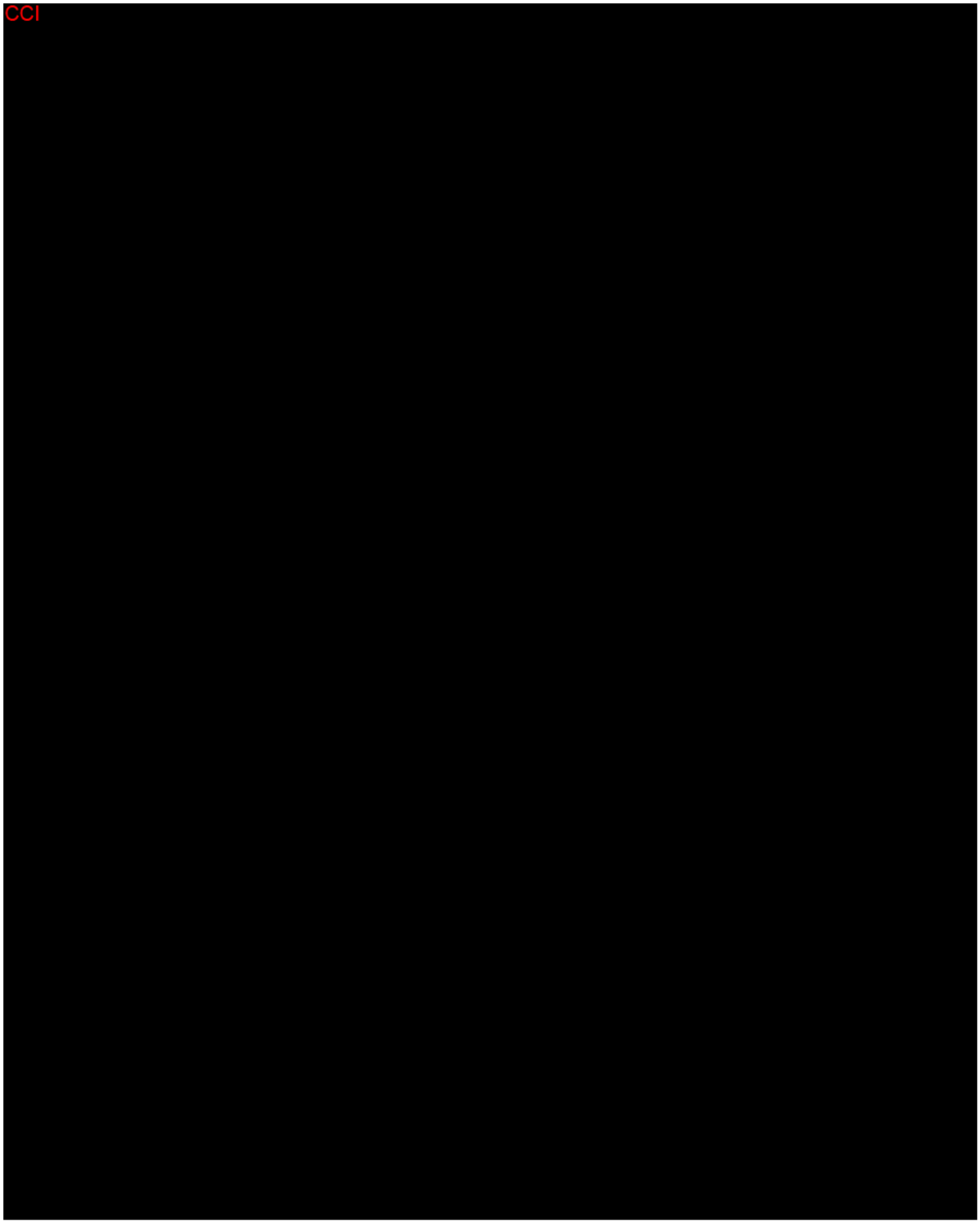

CCI

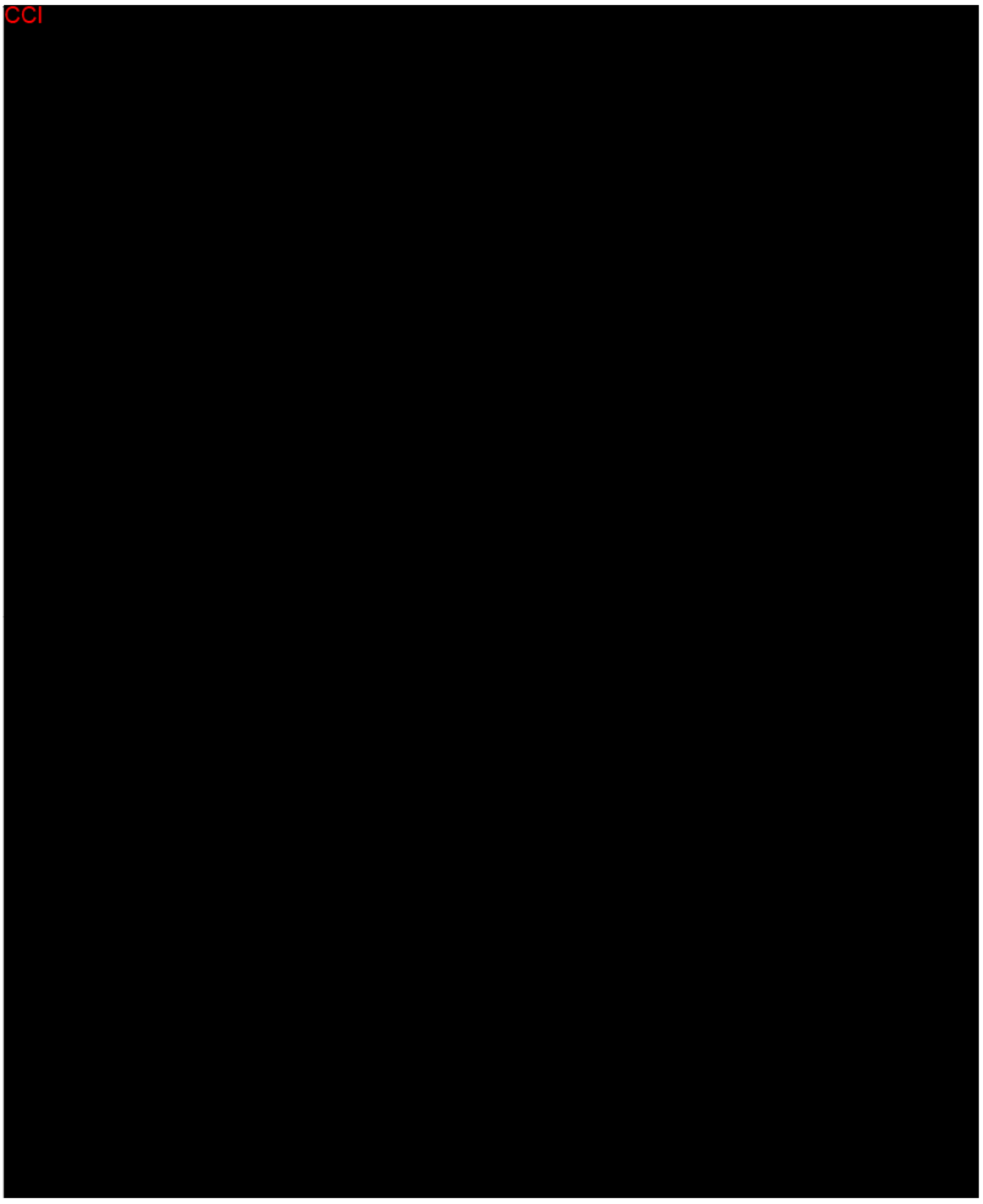

CCI

CCI

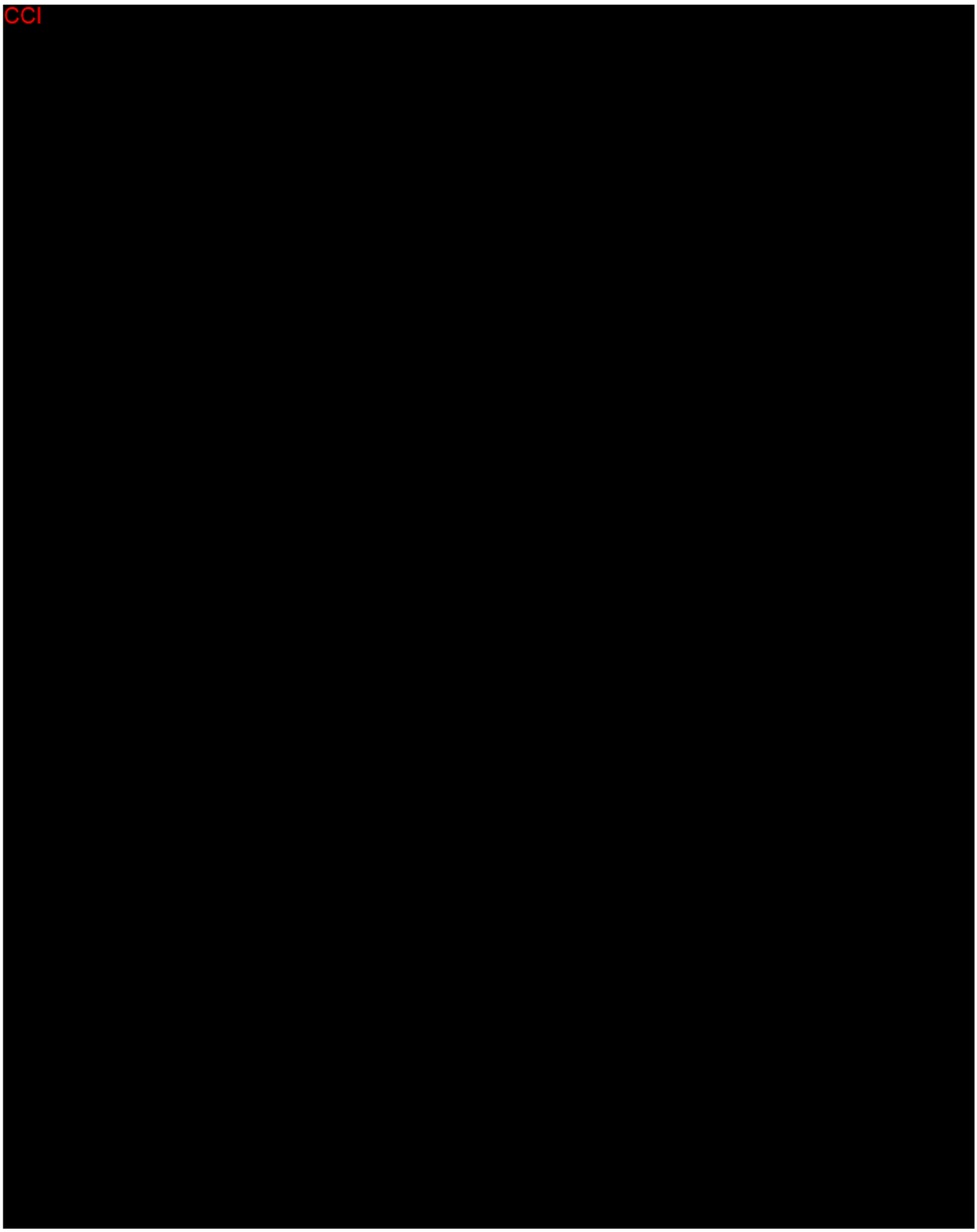


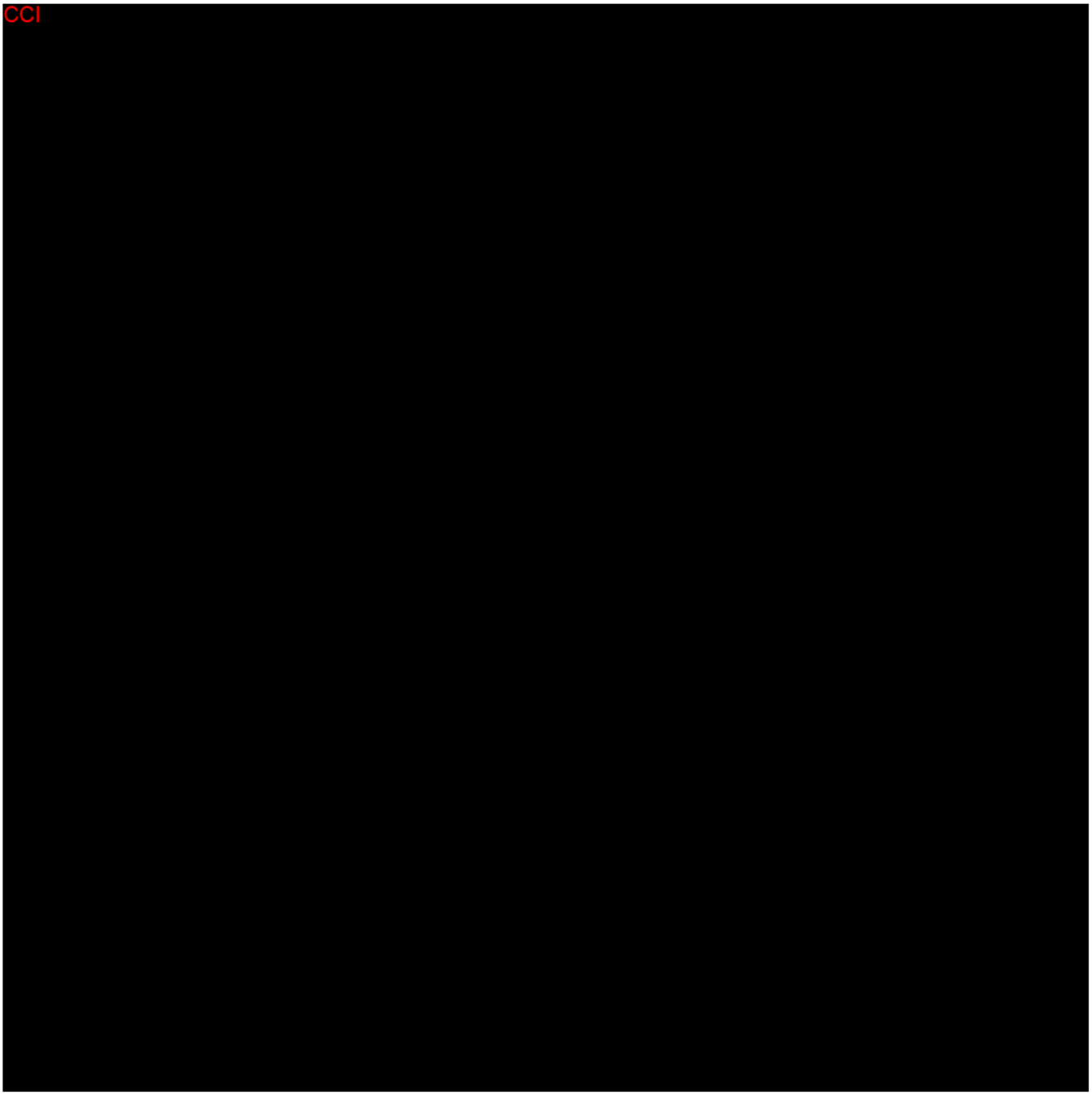
CCI




CCI

EORTC QLQ-NMIBC24 Scoring Manual




CCI

7. REFERENCES

1. Food and Drug Administration. BCG-unresponsive nonmuscle invasive bladder cancer: developing drugs and biologics for treatment: guidance for industry. Silver Spring, MD. 2018.
2. Steinberg, Gary, et al (2000) Efficacy and Safety of Valrubicin for the treatment of BCG refractory carcinoma in situ of the bladder. The Journal of Urology, 163:3, 761-767, March 2000, DOI: https://www.auajournals.org/doi/10.1016/S0022-5347%2805%2967799-3.
3. Osoba D, Rodrigues G, Myles J, Zee B, Pater J.(1998) Interpreting the significance of changes in health-related quality-of-life scores. J Clin Oncol;16(1):139–44.