Official Title of Study: A First-in-Human, Multicenter, Phase 1/2, Open-Label Study of XTX202 in Patients with Advanced Solid Tumors

Document Type: Statistical Analysis Plan

Document Date: 23 August 2023

NTC Number: NCT05052268

Statistical Analysis Plan (SAP)

Protocol Title:	A First-in-Human, Multicenter, Phase 1/2, Open-Label Study of XTX202 in Patients with Advanced Solid Tumors
Protocol Version No./Date:	5.0 (Amendment 4)/20-Apr-2023
eCRF Version No./Date:	3.0 / 23-Jan-2023
SAP Version No./Date:	3.0 / 23-Aug-2023

1.0 Approvals

Sponsor	
Sponsor Name:	Xilio Development, Inc.
Representative/ Title:	[redacted]
Signature /Date:	[redacted] This space reserved for eSignature block (if applicable).
ICON	
Biostatistician / Title:	[redacted]
Signature /Date:	[redacted] This space reserved for eSignature block (if applicable).

(NOTE: Electronic Signatures should only be used if all parties have the ability to eSign.)

2.0 Change History

Version	Change Log
1.0 (09JUNE2022)	Created as new
2.0 (15FEB2023)	Minor updates in advance of the FDA fast-track approval submission: 1. Calculation for hours and minutes into days corrected (see Section 10.7.4) 2. Removed text for displaying certain lab parameters in British imperial units 3. Updated the list of chemistry parameters that will be coded (see Section 10.8.7.1) 4. QTcF CTCAE grading updated for Grade 1 to include 450 msec (see Section 10.8.9) 5. SEGS and BANDS are not collected and so the text to replace missing neutrophil values with SEGS and BANDS was removed 6. Elaborated on the imputation for laboratory values of the form "< x" and "> x" (see Section 10.13.2) 7. Spider lots for tumor lesions was updated to specify percent change from baseline (see Section 13.3.2) 8. Disposition table text updated to include both reasons for completion and discontinuation from treatment and/or study (see Section 13.1) 9. Text added for potentially clinically important vital signs to specify that summaries are for post-baseline results (see Section 13.6.5)
3.0 (23AUG2023)	Updates after protocol versions 4.0 (amendment 3) and 5.0 (amendment 4): 1. Details for Part 1b added to Sections 7.0, 8.0, 9.0, 13.0 and 13.3. 2. Intent-To-Treat Analysis Set added for progression-free survival and overall survival (see Section 11.6). Response Evaluable Analysis Set updated to specify that analysis will be on assigned dose level (see Section 11.5). Details for the sensitivity analysis for progression-free survival on the All-Treated Analysis Set added to Sections 11.2, 13.3.5 and 13.3.6 respectively. Details also added to changes from protocol (see Section 6.1.2). 3. Update to the duration of exposure definition, to use 21 days or end of treatment as the upper bound of exposure (see Section 10.7.1) 4. Abbreviations not used in Section 15.0 removed. 5. Medications and therapies that start after the last administration of study treatment are not included in tables (see Section 10.5). 6. Summaries of treatment-emergent adverse events leading to study discontinuation removed as no longer collected on the case report form (see Section 13.6.1).

3.0 Table of Contents

1.0 Approvals 1
2.0 Change History 2
3.0 Table of Contents 3
4.0 Purpose 6
5.0 Scope 6
6.0 Introduction 6
6.1 Changes from Protocol 6
6.1.1 Number of Completed Doses 6
6.1.2 Intent-To-Treat Analysis Set 6
7.0 Study Objectives 6
7.1 Phase 1 (Part 1A and Part 1B) 7
7.1.1 Primary 7
7.1.2 Secondary 7
7.1.3 Exploratory 7
7.2 Phase 2 7
7.2.1 Primary 7
7.2.2 Secondary 7
7.2.3 Exploratory 7
8.0 Study Design 7
8.1 Phase 1: XTX202 Monotherapy Dose Escalation 8
8.1.1 Part 1a (Dose Escalation) 8
8.1.2 Part 1b (Pharmacodynamic Tumor Evaluation) 9
8.2 Phase 2: XTX202 Monotherapy Dosed at RP2D 9
8.2.1 Part 2a: Subjects with RCC 9
8.2.2 Part 2b: Subjects with Melanoma 10
8.3 Study Treatment 10
8.4 Sample Size Considerations 10
8.4.1 Phase 1 10
8.4.2 Phase 2 10
8.5 Randomization 11
9.0 Study Endpoints 11
9.1 Phase 1 (Part 1A and Part 1B) 11
9.1.1 Primary 11
9.1.2 Secondary 11
9.1.3 Exploratory (All Phases) 11
9.2 Phase 2 11
9.2.1 Primary 11
9.2.2 Secondary 11
10.0 Conventions and Derivations 12
10.1 Study Day 12
10.2 Baseline 12
10.3 Change from Baseline 12
10.4 Medical and Cancer History 12
10.5 Prior and Concomitant Medications and Therapies 12
10.6 Derivation of Efficacy Variables 13
10.6.1 Lesion Response Assessment per RECIST Version 1.1 13
10.6.2 Best Overall Response 15
10.6.3 Objective Response Rate 15
10.6.4 Disease Control Rate 16
10.6.5 Duration of Response 16
10.6.6 Progression-Free Survival 16
10.6.7 Overall Survival 16
10.6.8 Censoring Rules for Time-to-Event Endpoints 16

10.6.9 Measurable Disease at Baseline 17
10.7 Treatment Exposure Variables 17
10.7.1 Duration of Exposure 17
10.7.2 Number of Doses Completed 18
10.7.3 Number of Cycles Started 18
10.7.4 Dose Interruption 18
10.7.5 Dose Adjustment 18
10.8 Safety Variables 18
10.8.1 Assessment of Safety 18
10.8.2 Severity of Adverse Events 19
10.8.3 Treatment-Related Adverse Events 19
10.8.4 Treatment-Emergent Adverse Event 19
10.8.5 Dose-Limiting Toxicity 19
10.8.6 Events of Clinical Interest 19
10.8.7 Clinical Laboratory Parameters 20
10.8.8 Physical Examination and Vital Signs 20
10.8.9 Electrocardiograms 21
10.8.10 Eastern Cooperative Oncology Group Performance Status 22
10.8.11 Prior Anticancer Therapy and Subsequent New Anticancer Therapy 22
10.9 Pharmacodynamic Biomarkers and Immunogenicity Assessment 22
10.10 Other Derived Variables 23
10.10.1 Age 23
10.10.2 Body Mass Index 23
10.10.3 Differentials of Blood Cell Count 23
10.10.4 Worst On-Treatment Reference Range/Grade 24
10.10.5 Last On-Treatment Value 24
10.11 Handling of Partial or Missing Dates 24
10.11.1 Adverse Event and Concomitant Medication Partial Start and Stop Dates 24
10.11.2 Partial Dates of Diagnosis, Disease History and Death 25
10.12 Handling of Missing Adverse Event Severity and Relationship 25
10.13 Handling of Other Missing Data 25
10.13.1 Missing Absolute Differential 25
10.13.2 Laboratory Values with Character Symbols 26
10.14 Handling of Unscheduled Visits and Unscheduled Assessments 26
11.0 Analysis Sets 26
11.1 DLT-Evaluable Analysis Set 26
11.2 All-Treated Analysis Set 26
11.3 PK Analysis Set 26
11.4 Pharmacodynamic Analysis Set 26
11.5 Response-Evaluable Analysis Set 26
11.6 Intent-To-Treat Analysis Set 27
12.0 Interim Analysis 27
13.0 Statistical Methods 27
13.1 Subject Disposition 27
13.2 Demographics and Baseline Characteristics 28
13.2.1 Demographics 28
13.2.2 Prior and Concomitant Medications, Therapies and Procedures 28
13.2.3 Medical and Cancer History 28
13.3 Efficacy Analyses 28
13.3.1 Hypothesis Testing Strategy 29
13.3.2 Best Overall Response and Objective Response Rate 29
13.3.3 Disease Control Rate 29
13.3.4 Duration of Response 29
13.3.5 Progression-Free Survival 29
13.3.6 Overall Survival 29

13.4 Treatments 29
13.4.1 Extent of Study Drug Exposure 29
13.5 Important Protocol Deviations 30
13.6 Safety Analyses 31
13.6.1 Adverse Events 31
13.6.2 Deaths 32
13.6.3 Dose-Limiting Toxicity 32
13.6.4 Laboratory Data 32
13.6.5 Vital Signs 32
13.6.6 Physical Examinations, ECGs, and Other Observations Related to Safety 33
13.7 Pharmacodynamic Analyses and Immunogenicity 33
13.8 Pharmacokinetic Analyses 33
13.9 Methods for Handling Dropouts and Missing Data 33
13.10 Multiplicity 33
13.11 Impact by COVID-19 33
14.0 References 34
15.0 Glossary of Abbreviations 35

4.0 Purpose

The Statistical Analysis Plan (SAP) describes the statistical methods to be used during the reporting and analyses of data collected under Xilio Development, Inc. Protocol XTX202-01/02-001.

5.0 Scope

The SAP outlines the following:

- Study Objectives
- Study Design
- Study Endpoints
- Applicable Study Definitions
- Statistical Methods

See Glossary of Abbreviations for a list of abbreviations used throughout this document. The list of the mock tables, figures, and listings (TFLs) depicting the analyses described in this SAP are presented in a separate document.

6.0 Introduction

This SAP should be read in conjunction with the study protocol and electronic case report form (eCRF). This version of the plan has been developed using the protocol and eCRF version specified in the title page. Any further changes to the protocol or eCRF may necessitate updates to the SAP.

Changes following approval of the first signed version of the SAP will be tracked in the SAP Change Log; the final version of the SAP will be issued for sponsor approval prior to database lock. Any deviations from the final version of the SAP will be documented in the final Clinical Study Report.

6.1 Changes from Protocol

6.1.1 Number of Completed Doses

In Section 12.5, the protocol defines the number of completed doses as:

“For the number of doses started, if a patient has any record of starting administration number “X,” the patient will be counted as completed administration X-1”

In Section 10.7.2 within this SAP, after agreement between ICON and Xilio, the number of completed doses was updated to:

“The number of non-missing actual doses administered as captured in the ‘XTX202 Administration’ eCRF page where Actual Dose = Planned Dose.”

6.1.2 Intent-To-Treat Analysis Set

Since protocol version 3 (amendment 2), once a recommended phase 2 dose (RP2D) is defined, subjects remaining on a lower dose may be escalated to the RP2D, ‘if recommended by their treating investigator’.

Consequently, for the reporting of the progression free survival (PFS) and overall survival (OS) a new population, the Intent-to-Treat (ITT) Analysis Set, has been defined within the SAP. The ITT Analysis Set includes all subjects who have received at least one dose of XTX202. Subjects will be tabulated based on their assigned dose level (DL). This Analysis Set will be defined as the primary population while the All-treated Analysis Set will be used for the sensitivity analysis of PFS.

7.0 Study Objectives

Study objectives of Phase 1 and 2 of the study are outlined in this section. Refer to Section 8.0 for study design details.

7.1 Phase 1 (Part 1A and Part 1B)

7.1.1 Primary

- To evaluate the safety and tolerability of XTX202 monotherapy in subjects with advanced solid tumors.
- To determine the RP2D(s) and schedule of XTX202 monotherapy.

7.1.2 Secondary

- To characterize the pharmacokinetic (PK) profile of XTX202.
- To evaluate the immunogenicity of XTX202.
- To evaluate the preliminary antitumor activity of XTX202 monotherapy in subjects with advanced solid tumors.

[redacted]

7.2 Phase 2

7.2.1 Primary

- To evaluate the efficacy of XTX202 monotherapy in subjects with metastatic renal cell carcinoma (RCC) and subjects with unresectable or metastatic melanoma.

7.2.2 Secondary

- To evaluate the safety and tolerability of XTX202 monotherapy in subjects with metastatic RCC and subjects with unresectable or metastatic melanoma
- To characterize the PK profile of XTX202 in subjects with metastatic RCC and subjects with unresectable or metastatic melanoma
- To evaluate the immunogenicity of XTX202 in subjects with metastatic RCC and subjects with unresectable or metastatic melanoma
- To further evaluate the efficacy of XTX202 monotherapy in subjects with metastatic RCC and subjects with unresectable or metastatic melanoma

[redacted]

8.0 Study Design

This is a first-in-human, Phase 1/2, multicenter, open-label study designed to evaluate the safety, tolerability and efficacy of XTX202, a tumor selective, engineered interleukin-2 (IL-2) prodrug with its activity masked to keep it inactive until selectively activated in vivo in the tumor microenvironment, as monotherapy in subjects with advanced solid tumors.

This study consists of two monotherapy Phases:

Phase 1

- **Part 1a**
  Part 1a will examine XTX202 monotherapy in an accelerated and standard 3+3 dose escalation design to determine the RP2D(s) for Phase 2.

- **Part 1b**
  Based on the results of Part 1a, subjects with select advanced solid tumors will be enrolled in Part 1b, which will evaluate XTX202 monotherapy in relation to specific PaD biomarkers; subjects in Part 1b will be required to have fresh tumor biopsies predose and postdose. [redacted]

Phase 2

Phase 2 will examine the RP2D of XTX202 in select indications to further evaluate the efficacy of XTX202. These will be in 2 disease-specific expansion cohorts:

- **Part 2a**
  Part 2a will include subjects with metastatic RCC who have received a tyrosine kinase inhibitor (TKI) therapy and also have been treated and progressed on an anti-programmed cell death-protein 1 (anti-PD-1) therapy;
- **Part 2b**
  Part 2b will include subjects with unresectable or metastatic melanoma who have received immune-checkpoint therapy with an anti-PD-1 therapy and a cytotoxic T lymphocyte associated protein 4 (CTLA-4) therapy.

8.1 Phase 1: XTX202 Monotherapy Dose Escalation

8.1.1 Part 1a (Dose Escalation)

Part 1a will evaluate ascending doses of XTX202 monotherapy administered via intravenous (IV) infusion every 3 weeks (Q3W). It will employ a single-subject accelerated 100% dose escalation design. If a Grade ≥ 2 adverse event (AE) or any dose-limiting toxicity (DLT) is observed, then the study will transition to a 3+3 dose escalation that will include a 40% increase over the previous dose. [redacted]

Beginning with DL 6 in the 100% dose escalation design (8.0 mg/kg), the study will automatically transition to a conventional 3+3 design with a 40% dose increase independent of whether a Grade ≥ 2 AE or DLT is observed at the preceding DL. See Table 4 in the protocol for illustrations.

A subject will be enrolled at DL1 and monitored for 21 days postdose (the DLT Observation Period). If no Grade ≥ 2 AE or DLT is observed, then dose escalation will proceed to the next DL. Dose escalation will continue with a naïve single subject at each DL up to DL 6 (8.0 mg/kg) in the 100% dose escalation design as long as no subject experiences a Grade ≥ 2 AE or DLT. If a Grade ≥ 2 AE or a DLT is observed in any subject, dose escalation will proceed with a 40% increase over the previous dose and transition to a conventional 3+3 design. Grade ≥ 2 AEs assessed by the SRC to have an etiology completely unrelated to study drug will not trigger a transition to 3+3 dose escalation.

[redacted]

Starting at DL6 (8.0 mg/kg) in the 100% dose escalation design, Phase 1 will transition to a conventional 3+3 design with a 40% dose escalation independent of whether any Grade ≥ 2 AE or DLT is observed at the preceding DLs. Under the 3+3 design, if no subjects experience a DLT, escalation will continue to the next DL; if 1 of the first 3 subjects at a given DL has a DLT, then an additional 3 subjects will be enrolled at that DL; if ≥ 2 subjects have a DLT during the DLT Observation Period (21 days), dose escalation will be stopped and the DL immediately below this level will be considered the maximum tolerated dose (MTD).

After all naïve subjects within a given DL have been followed for at least 21 days after first dose, the SRC will review all accumulated safety data and available laboratory and PK data. Initiation of the next DL will commence upon recommendation by the SRC and agreement from the Sponsor.

Dose escalation will continue until either an MTD is determined and/or the highest RP2D is defined.

[redacted]s may be enrolled at any DL with < 2 DLTs. These additional subjects will not be considered for defining the MTD but will contribute to the overall safety assessment and PK/PaD profile at that DL. The RP2D(s) selected for evaluation in Phase 2 will be based upon review of the cumulative safety, PK, PaD, and available preliminary antitumor activity observed.

Once a RP2D is defined based on SRC recommendation, any subject remaining on treatment at a DL lower than the RP2D may have their dose escalated to a RP2D, if recommended by their treating investigator.

8.1.1.1 DLT Criteria

See Section 10.8.5.

8.1.2 Part 1b (Pharmacodynamic Tumor Evaluation)

Based on the results from Part 1a, Part 1b will be initiated to further examine XTX202 as monotherapy in subjects with select advanced solid tumors and to further characterize XTX202. Part 1b will require mandatory fresh tumor biopsies predose and postdose to fully characterize the PaD profile of XTX202.

[redacted]

Part 1b may be initiated at any DL successfully cleared in Part 1a based upon review of the cumulative safety, PK, PaD, and preliminary antitumor activity, following approval from the SRC. Part 1b may enroll up to a total of approximately [redacted] (including those in the "window of opportunity" subcohort).

Once a RP2D is defined based on SRC recommendation, any subject remaining on treatment in Part 1b at a DL lower than a RP2D may have their dose escalated to the RP2D, if recommended by their treating investigator.

8.2 Phase 2: XTX202 Monotherapy Dosed at RP2D

8.2.1 Part 2a: Subjects with RCC

Part 2a will include subjects with metastatic RCC who have received a TK1 therapy and also have been treated and progressed on an anti-PD-1 therapy. This part will employ an optimal Simon's 2-stage design; subjects will be treated with XTX202 monotherapy at the RP2D until progression of disease, unacceptable toxicity, 3 months after complete response (CR), or 24 months of total study therapy.

8.2.2 Part 2b: Subjects with Melanoma

Part 2b will include subjects with unresectable or metastatic melanoma who have received immune checkpoint therapy with an anti-PD-1 therapy. This part will employ an optimal Simon's 2-stage design; subjects will be treated with XTX202 monotherapy at the RP2D until progression of disease, unacceptable toxicity, 3 months after CR, or 24 months of total study therapy.

Note: Based on the January 2023 Food and Drug Administration (FDA) draft guidance, "Optimizing the Dosage of Human Prescription Drugs and Biological Products for the Treatment of Oncologic Diseases," (US Food and Drug Administration, 2023) the SRC may recommend more than 1 dose to be evaluated in the Phase 2 portion of the study in order to identify the optimal XTX202 dose for further development. The SRC may recommend each RP2D based on the totality of the data (safety, efficacy, and available PK and PaD data), and Phase 2 may open in parallel with ongoing dose escalation in Part 1a prior to reaching MTD. Once more than 1 RP2D is defined, subjects in Phase 2 will be allocated to a DL by the Sponsor.

8.3 Study Treatment

Study treatment, or Investigational Medical Product: XTX202;

Dosage and Mode of Administration: XTX202 will be administered by IV infusion at a starting dose of 0.27 mg/kg Q3W.

8.4 Sample Size Considerations

It is anticipated that the study may enroll up to approximately [redacted] across all parts of the study:

Phase 1 - Dose Escalation and PaD expansion: Up to [redacted]

- Part 1a – Dose Escalation: up to [redacted]
- Part 1b – PaD dose expansion: up to [redacted]

Phase 2 - Dose Expansion: up to [redacted]

- Part 2a: up to [redacted] with metastatic RCC
- Part 2b: up to [redacted] with unresectable or metastatic melanoma
- Note: In Phase 2, subjects who discontinue study prior to their first efficacy assessment due to reasons other than progressive disease (PD) or treatment-emergent adverse events (TEAEs) may be replaced, therefore the actual number of subjects enrolled may be larger.
- Note: The SRC may recommend more than 1 RP2D to be evaluated in the Phase 2 portion of the study to further characterize safety, efficacy, PK, PaD and identify the optimal XTX202 dose for further development. [redacted]

8.4.1 Phase 1

No formal hypothesis testing is used in sample size planning for Part 1a and Part 1b. The planned number of evaluable subjects is derived from the dose-escalation design. The final sample size may vary depending on the total number of DLs that will be evaluated, subject replacement for DLT evaluation, if applicable, and expansion from 3 to 6 subjects at a given DL if a DLT is observed.

8.4.2 Phase 2

Part 2a and Part 2b will each enroll up to [redacted] at the RP2D in an optimal Simon's 2-stage design (Simon, 1989) to exclude an objective response rate (ORR) < 5%. The null hypothesis that P <= 0.05 versus the alternative that P >= 0.20 under treatment will be tested in each individual cohort using a one-sided type I error rate of < 0.05 and 80% power, with a probability of early termination assumed to be 0.418. Part

2a and 2b will be independently assessed. For either Part, after evaluating the first [redacted] for efficacy in Stage 1, the respective cohort will be terminated if no subject responds (either CR or partial response, PR). If either Part 2a or 2b progresses to Stage 2, an additional [redacted] will be treated in that cohort, for a total of up to [redacted] in each Part. If at least [redacted] respond out of the total [redacted] in a particular cohort, there will be demonstration of efficacy of XTX202 within that respective population and the corresponding null hypothesis will be rejected for the respective Part.

8.5 Randomization

Randomization and blinding method are not applicable as this is a single-arm, open-label study.

9.0 Study Endpoints

Study endpoints of Phase 1 and Phase 2 of the study are outlined in this section.

9.1 Phase 1 (Part 1A and Part 1B)

9.1.1 Primary

- Incidence of DLTs (Part 1a only)
- Incidence of TEAEs and changes in clinical laboratory values

9.1.2 Secondary

- Plasma concentrations of XTX202 (total and intact), including maximum observed serum concentration (C_{max}), time of maximum observed concentration (T_{max}), trough concentration (C_{trough}), area under the curve (AUC), half-life ($t_{1/2}$), systemic clearance (CL), and volume of distribution (Vd)
- Antidrug antibody (ADA) occurrence and titer in serum
- Investigator-assessed ORR per Response Evaluation Criteria in Solid Tumors (RECIST) version 1.1

[redacted]

9.2 Phase 2

9.2.1 Primary

- Investigator-assessed ORR per RECIST version 1.1

9.2.2 Secondary

- Incidence of TEAEs and changes in clinical laboratory values
- Plasma concentrations of XTX202 (total and intact), including C_{max}, T_{max}, C_{trough}, AUC, $t_{1/2}$, CL, and Vd
- Incidence and persistence of ADAs (including neutralizing ADAs) and titers, and their potential impact on PK, activity, and safety associated with XTX202

- Duration of response (DOR), defined as the time from first documented confirmed response to first documented disease progression
- Disease control rate (DCR), defined as the percent of subjects who achieve a confirmed CR, PR, or stable disease (SD)
- PFS, defined as the time from first dose to first documented disease progression or death
- OS, defined as the time from first dose to death due to any cause

10.0 Conventions and Derivations

10.1 Study Day

All study days on or after the first administration of study treatment (XTX202) will be calculated as date of assessment minus date of first administration of study treatment plus 1, i.e.,

Date of assessment – date of first administration of study treatment + 1

Study days that occur before the first administration of study treatment will be calculated as date of assessment minus date of first administration of study treatment, i.e.,

Date of assessment – date of first administration of study treatment

Date of first administration of study treatment is the date of Cycle 1 Day 1 on which study treatment is initiated.

Every effort will be made to avoid missing and/or incomplete dates. In cases of missing and/or incomplete dates no study days will be calculated.

10.2 Baseline

Unless otherwise specified, baseline will be defined as the last available measurement taken prior to the first administration of study treatment in each part of the study.

For applicable variables that include a date and time, both date and time should be used in the calculations.

10.3 Change from Baseline

Change from baseline value is defined as the value at any given post-baseline timepoint minus the baseline value where both values are non-missing, i.e.,

Post-baseline value – baseline value

10.4 Medical and Cancer History

Time since initial diagnosis and time since most recent progression/recurrence will be calculated in years from the informed consent date. Derivations are as follows:

Time since initial diagnosis = (Informed Consent date - the date of initial diagnosis + 1) / 365.25.

Time since most recent progression/recurrence = (Informed Consent date - the date of most recent progression/recurrence + 1) / 365.25.

10.5 Prior and Concomitant Medications and Therapies

Prior medications and therapies are defined as medications and therapies with a stop date prior to the first administration of study treatment.

Concomitant medications and therapies are defined as medications and therapies other than the study treatment that are ongoing or with a stop date on or after the first administration of study treatment. Medications and therapies that start after the last administration of study treatment are not included.

All concomitant medications received within 28 days prior to the first dose of study treatment and up to 90 days after the last dose of study intervention at the Safety Follow-up Visit will be recorded. All concomitant medications administered during serious adverse events (SAEs) or Events of Clinical Interest (ECIs) are to be recorded.

Use of all prior and concomitant medications, including any change in treatment, will be recorded, categorized and summarized according to the latest version of the WHO (World Health Organization) Drug Dictionary (WHODrug Global B3 01MAR2021 or later). Each medication will be assigned a preferred term (PT).

Refer to section 10.12 for how to handle partial or missing dates in the assessment of whether or not a treatment was taken prior to or concomitantly with the study treatment.

10.6 Derivation of Efficacy Variables

Response will be based on Investigator assessment according to RECIST version 1.1 and will be reported using the following response categories: CR, PR, SD, PD and not evaluable (NE).

The first occurrence of CR or PR should be confirmed at a subsequent scan at least 4 weeks after the initial response determination.

For each subject, the best overall response (BOR) is defined as the best overall response across all efficacy assessments, recorded from the start of treatment until death, documented disease progression, or the date of start of subsequent therapy, whichever occurs first, taking into account any requirement for confirmation. BOR will be used to inform ORR (confirmed CR and PR).

Disease assessment by computed tomography (CT) or magnetic resonance imaging (MRI) as applicable for disease type and anatomical regions of disease involvement will be conducted using a consistent modality to enable response assessment according to RECIST version 1.1 (Eisenhauer, et al., 2009).

Imaging will be performed at baseline (within 4 weeks prior to first dose) and during the study as outlined in the Schedule of Assessments. Ongoing tumor evaluations to the same areas of disease involvement using the same method of scanning will be performed every 3 cycles (9 weeks ± 5 days) from Cycle 1, Day 1 (C1D1) during the first 12 months, and then every 4 cycles (12 weeks ± 7 days) thereafter until disease progression or the start of subsequent new anticancer therapy, including in subjects who discontinue treatment for reasons other than disease progression. Unscheduled imaging is permitted if clinically indicated per the discretion of the Investigator.

Adjunctive imaging by positron emission tomography is permitted.

The modality chosen (CT or MRI) to evaluate each individual subject should remain consistent for disease evaluation throughout the duration of the study. Prior to treatment initiation, target and non-target lesions will be identified by the Investigator per RECIST version 1.1. During the study, disease response and progression will be determined by Investigator assessment per RECIST version 1.1 at the time points indicated in the Schedule of Assessments. Subjects who achieve a CR or PR should undergo a follow-up response confirmation assessment (CT or MRI) ≥ 4 weeks after the first response. Subjects will be expected to continue to follow the response assessment schedule as outlined in the Schedule of Assessments.

10.6.1 Lesion Response Assessment per RECIST Version 1.1

The primary endpoint in the efficacy expansion cohorts (Part 2a and 2b of Phase 2) is the ORR that is based on all radiographic scans evaluated by the investigator according to RECIST version 1.1.

When evaluating ORR, BOR will be used that is obtained among all tumor assessment visits after start of study treatment until death, documented disease progression, or the date of start of subsequent therapy, whichever occurs first, taking into account the following requirement for confirmation (see also Table 1 below).

Based on the assessment of target lesions (TLs), non-target lesions (NTLs) and the appearance of any new lesions (NLs) an overall tumor response is determined. It is this overall tumor response that is used in

the derivation of efficacy analyses. Efficacy endpoints can only be determined for subjects that have at least one post-baseline assessment response.

Efficacy will be assessed using the following endpoints: BOR, ORR, DCR, DOR, PFS, and OS.

The evaluation of BOR in order of best to worst is categorized as CR, PR, SD, PD, and NE. PD is not considered as a positive response.

Many of the endpoints rely on a confirmatory response. When CR or PR is observed, the responses may be claimed only if the criteria are confirmed at a subsequent timepoint that falls on or after 4 weeks (i.e., 28 days) of the initial evaluation. Subsequent documentation of CR may provide confirmation of a previously identified CR with no more than two intervening NE responses (e.g., CR, NE, NE, CR). Subsequent documentation of CR may provide confirmation of a previously identified PR (i.e. confirmed PR) even with an intervening SD Time Point Response (TPR) assessment and/or no more than two intervening consecutive NE responses. SD requires a minimum duration of 6 weeks (i.e., 42 days) from the start of treatment. Where an overall response is reported as SD and no overall response date is provided, it should be assumed that the SD duration has not met the minimum 6 weeks.

For CR and PR, confirmation of the response according to RECIST version 1.1 will be required, preferably no later than at the regularly scheduled 9-week (3-cycle) assessment interval, but no sooner than 4 weeks after the initial documentation of CR or PR. Confirmation of PR can be confirmed at an assessment later than the next assessment after the initial documentation of PR. SD does not require a confirmatory response. A BOR of SD requires that a Time Point Overall Response (TPOR) of SD has been determined at a timepoint at least 6 weeks (42 days) after start of study treatment. PD also does not require confirmation.

TPOR will be determined based upon the component TPRs of TLs, NTLs and the appearance or absence of NLs, according to RECIST version 1.1. TPOR as reported by the Investigator based on radiographic images and physical findings will be reviewed and evaluated for BOR and ORR in accordance with RECIST version 1.1.

When evaluating objective response (CR or PR), SD is not considered. PR following PD is not considered for BOR, and therefore neither is it considered for ORR.

Table 1: Confirmed Response based on Subsequent Assessments*

First Time Point Response*	Second Time Point Response	Confirmed Response (Best Response)**
PD	No further evaluation	PD
NE	PD	PD
CR	PD	SD or PD (1)
PR	PD	SD or PD (1)
SD	PD	SD or PD (1)
CR	CR	CR
CR	NE **	SD or NE (2)
CR	ND (remains on treatment)	uCR
PR	CR	PR
PR	PR	PR
PR	SD (3) **	SD
PR	NE **	SD or NE (2)
PR	ND (remains on treatment)	uPR
SD	CR	SD
SD	PR	SD
SD	SD	SD
SD	NE	SD or NE (2)
NE	CR	SD
NE	PR	SD
NE	SD	SD
NE	NE	NE

* Subsequent documentation of CR may provide confirmation of a previously identified CR for subjects where the second integrated response was NE. Subsequent documentation of PR may provide confirmation of a previously identified PR for subjects where the second integrated response was NE or SD. If the third TPR confirms the CR (or PR) then the Confirmed Response will be CR (or PR). For this study, multiple intervening NE is allowed between CRs/PRs. For example: CR NE NE CR = CR; PR NE NE PR = PR. Additionally, multiple SD is allowed between PRs (e.g., PR SD SD PR = PR).

** A Best Response of SD can only be made at a time point at least 6 weeks (42 days) after start of study treatment, and any tumor assessment indicating SD before this time period will have a Best Response of NE unless PD is identified.

(1) Best response will be SD if the first TPR is at least 6 weeks (42 days) after start of study treatment. Otherwise, the best response will be PD.

(2) Best response will be SD if the first TPR is at least 6 weeks (42 days) after start of study treatment. Otherwise, the best response will be NE.

(3) TPR is SD if the increase from the first to the second assessment does not qualify for PD.

For subjects with unconfirmed CR/PR (uCR/uPR) and who subsequently drop off the study, the BOR will be SD, assuming the minimum duration of SD (42 days) has been met. Otherwise the BOR will be NE.

10.6.2 Best Overall Response

BOR is defined as the best assessment response as documented by the investigator per RECIST version 1.1 through the course of the evaluations, with a follow-up confirmatory response.

10.6.3 Objective Response Rate

ORR is defined as the proportion of subjects with a confirmed BOR of CR or PR up until PD, death or subsequent anticancer therapy, according to RECIST version 1.1.

Subjects with confirmed objective response (CR or PR) will be classified as responders, and non-responders will include subjects without a confirmed objective response, with SD or PD. Subjects with inadequate data for tumor assessment (e.g., no follow-up assessments) will be considered as non-responders in the assessment of response rate.

10.6.4 Disease Control Rate

DCR is defined as the proportion of subjects with a confirmed BOR of CR, PR or SD according to RECIST version 1.1.

10.6.5 Duration of Response

DOR, in months, is defined as the time from the first documentation of a subsequently confirmed objective response (CR or PR) to the date of the first documentation of radiographic disease progression according to RECIST version 1.1, or death due to any cause, whichever occurs first.

DOR will be calculated only for the subgroup of subjects with a confirmed objective response (CR or PR).

DOR = (min (date of first progression, date of death) – date of first CR or PR + 1) / 30.4375

The date of response or progression will be the date of assessment recorded for the tumor evaluation for each applicable timepoint.

Refer to section 10.6.8 for censoring rules.

10.6.6 Progression-Free Survival

PFS, in months, is defined as from first administration of study treatment until first documentation of radiographic disease progression according to RECIST version 1.1, or death due to any cause, whichever occurs first.

PFS = (Date of first progression or death – date of first dose of study treatment + 1) / 30.4375

The date of response or progression will be the date of assessment recorded for the tumor evaluation for each applicable timepoint.

Refer to section 10.6.8 for censoring rules.

10.6.7 Overall Survival

OS, in months, is defined as the time from first administration of study treatment to death due to any cause.

OS = (Date of death – date of first dose of study treatment + 1) / 30.4375

After the 30-day EOT Visit and 90-day Safety Follow-up Visit subjects will be followed quarterly for survival. Date of death will be recorded on the 'Death Details' eCRF page. For subjects without a record of death, OS will be censored at the date they were last known alive. Refer to section 10.6.8 for censoring rules.

10.6.8 Censoring Rules for Time-to-Event Endpoints

The following applies to the analysis of OS:

- For subjects alive, censoring for survival will be done on the date of the last on-study follow-up, including survival follow up assessment, that the subject is reported to be alive. Subjects who have no on-study data will be censored on the date of first administration of study treatment (Day 1).

The following censoring rules will apply to event dates for time-to-event endpoints that are based on radiographic evaluations, i.e., DOR and PFS:

- Endpoints will be censored on the date of the first dose of study treatment with duration of 1 day under the following scenarios (apply to PFS only):
    - Baseline disease assessment inadequate to apply RECIST version 1.1;
    - No disease assessments are performed during study treatment, except in the event of early death (≤ 18 weeks [+ 5 days to allow for a late visit = 131 days] with no tumor assessment, see further details below for death as an event); or
    - All disease assessments performed during study treatment result in the conclusion of NE.

- Endpoints will be censored on the date of the last evaluable disease assessment under the following scenarios (apply to PFS and DOR):
    - PD or death occurs after ≥ 2 consecutive tumor assessments that are missed or result in the conclusion of NE (i.e., for the previous evaluable assessment within 12 months: >= 18 weeks (+ 5 days to allow for a late visit = 131 days) or for previous evaluable assessment after 12 months: >= 24 weeks (+ 7 days to allow for a late visit = 175 days);
    - Subject administered alternative anticancer treatment prior to documented PD;
    - Subject lost to follow-up;
    - Subject withdrawal of consent for follow-up; or
    - Subject continues on study without PD at the time of data cut-off or End of Study.

10.6.9 Measurable Disease at Baseline

Measurable disease at baseline for the purpose of the Response-Evaluable Analysis Set (see Section 11.5) will be defined as subjects who have at least one target lesion at baseline from either of the two categories below, with a minimum length of:

1. Tumor lesions (Non-nodal): at least 10mm in longest diameter
2. Pathological lymph nodes (Nodal): at least 15mm in short diameter (longest perpendicular)

10.7 Treatment Exposure Variables

Table 2 presents some of the treatment exposure definitions.

Table 2: Definitions

Variable	Definition
Duration of exposure (weeks)	(the last dose date– the first dose date + 21) / 7
Number of doses completed	The number of non-missing actual doses administered as captured in the 'XTX202 Administration' eCRF page where Actual Dose = Planned Dose.
Cumulative dose received (mg/kg)	The total amount of the study drug a subject receives during the study, calculated as: Sum of [Actual Dose Administered] as recorded on the 'XTX202 Administration' eCRF page.
Cumulative planned dose (mg/kg)	The total amount of the study drug a subject is planned to receive during the study, calculated as: Sum of [Planned Dose] as recorded on the 'XTX202 Administration' eCRF page.
Number of Cycles started	A subject is considered to have started a cycle if they received at least one dose (actual dose > 0) of study treatment in that cycle, per the 'XTX202 Administration' eCRF page.
Relative dose intensity (%)	Relative dose intensity will be calculated as: (Cumulative dose received / Cumulative planned dose) * 100.

10.7.1 Duration of Exposure

Duration of exposure (weeks) is defined as the time from the first dose of study treatment to the last dose of study treatment, plus 21 days:

(Date of last administration of study treatment – date of first administration of study treatment + 21) / 7

If EOT occurs prior to the end of the duration above, then duration stops at EOT, and duration of exposure will be:

(Date of EOT – date of first administration of study treatment + 1) / 7

Treatment-free periods within each cycle will be included.

10.7.2 Number of Doses Completed

The number of non-missing actual doses administered as captured in the 'XTX202 Administration' eCRF page where Actual Dose = Planned Dose.

10.7.3 Number of Cycles Started

A subject is considered to have started a cycle if they received at least one dose (actual dose > 0) of study treatment in that cycle, per the 'XTX202 Administration' eCRF page. Cycle information can be found using the Visit variables from the electronic data capture (EDC) system.

10.7.4 Dose Interruption

Dose interruptions and the duration of interruption are recorded on the 'XTX202 Administration' eCRF pages as 'Was the Infusion interrupted?' = 'Yes' and 'If the Infusion was interrupted, how long was the interruption?' respectively. Duration will be summarized in minutes.

Dose interruption per subject (minutes) will be calculated as the sum of interruptions per administration in minutes.

If collected in hours or days, the value will be converted into minutes as follows:

Minutes = Hours * 60.

Minutes = Days * 1440.

10.7.5 Dose Adjustment

Dose adjustments of study treatment are captured as 'Was the dose adjusted' = 'Yes' on the 'XTX202 Administration' eCRF pages. Checks will be done to confirm there is a difference in 'Actual Dose Administered' as compared against 'Planned Dose'.

10.8 Safety Variables

10.8.1 Assessment of Safety

Safety will be assessed through the monitoring of AEs (including DLTs), physical examinations (PEs), and regular clinical laboratory evaluations. Subjects will be monitored continuously for toxicity while on treatment until the Safety Follow-up Visit (90 days after the last dose).

Dose modifications will be applied for AEs assessed as related to XTX202 and of particular severity according to the guidelines set forth in the study protocol (Section 10.2).

AEs will be coded to a PT and classified to a System Organ Class (SOC) according to the latest version of the Medical Dictionary for Regulatory Activities (MedDRA version 24.1 or later). Events that are not coded at the time of the data snapshot for the SRC review will be included and presented using the AE verbatim term (VT), i.e., Investigator recorded AE term (in uppercase letters) as the PT and using "UNCODED" for the SOC.

SAEs will be recorded on the appropriate eCRF within 24 hours of the Investigator's first knowledge of the event, even if the experience does not appear to be related to XTX202, from the time of signing the Informed Consent Form (ICF) through the Safety Follow-up Visit after the last dose of XTX202. SAEs that are considered related to study therapy occurring more than 90 (+/- 7) days after the last dose of study therapy should be reported.

10.8.2 Severity of Adverse Events

Severity of AEs will be assessed using the National Cancer Institute Common Terminology Criteria for Adverse Events (NCI-CTCAE) version 5.0. Events not graded or with a missing severity will not be imputed.

10.8.3 Treatment-Related Adverse Events

The relationship of AEs to study treatment is recorded on the 'Relationship to Study Treatment (XTX202)' field on the 'Adverse Events' eCRF. If the relationship is missing, it will be assumed to be treatment-related.

10.8.4 Treatment-Emergent Adverse Event

A TEAE is an AE that emerges or worsens in the period from the first dose of study treatment to the Safety Follow-up visit after the last dose of study treatment. For the purpose of analysis, AEs that start or increase in severity on or after the first dose of study drug and no later than 90 days after the last dose of study drug will be considered treatment-emergent.

Refer to Section 10.11 for how to handle partial or missing AE dates in the assessment of whether or not an event is treatment-emergent.

10.8.5 Dose-Limiting Toxicity

DLTs will be evaluated during Phase 1, Part 1a (monotherapy dose escalation). All toxicities (AEs) will be graded using the NCI CTCAE version 5.0 based on Investigator assessment.

DLTs for Part 1a are defined as the following:

- Any treatment-related Grade ≥ 3 toxicity
- Any Grade febrile neutropenia
- The following nonhematologic exceptions:
  - Grade 3 nausea or vomiting lasting < 3 days
  - Grade 3 fatigue lasting < 7 days
- Any treatment-related toxicity that results in a treatment delay of ≥ 7 days

See Section 6.4.1 of the protocol for more details.

10.8.5.1 DLT Observation Period

The DLT Observation Period will include the first cycle of XTX202 and will run for approximately 21 days, beginning at C1D1 and ending just prior to the second dose of the study drug at C2D1. For any subject in a dose hold prior to C2D1, the DLT Observation Period may be extended up to 7 days (28 days in total) to confirm the evaluation of potentially ongoing DLTs.

For the analysis, the DLT observation period will include all DLTs up to Day 28, or prior to the start date of C2D1 (if C2D1 occurs prior to day 28).

10.8.6 Events of Clinical Interest

Selected nonserious and serious AEs also known as ECIs will be reported, which include:

1. An overdose of investigational study medication whether or not associated with clinical symptoms or abnormal laboratory results
2. Drug-induced liver injury
3. Cytokine release syndrome (CRS)
4. Capillary leak syndrome (CLS)

See Section 11.2.2 of the protocol for more details.

ECIs are recorded in the 'Is this Event of Clinical Interest?' field on the 'Adverse Events' eCRF.

10.8.7 Clinical Laboratory Parameters

All parameters will be tested by local laboratories. Clinical laboratory data will be entered into the EDC system and converted to International System (SI) units for analysis. Normal ranges will be merged in during dataset programming. Baseline values are reported any time prior to start of treatment on Cycle 1 Day 1 in each phase.

All laboratory data will be presented in International System (SI) units.

A full list of safety laboratory parameters is given in Section 8.1.9 of the study protocol.

10.8.7.1 CTCAE Coding of Laboratory Data

NCI CTCAE v5.0 will be used to assign grades to laboratory data.

Note that grades are applied based only on the numeric SI value of the parameter assessed; clinical signs and symptoms are not considered. For example, Grade 4 hypoglycemia will be assigned based solely on the value of the glucose measurement, and acidosis will not be considered. Where categories are only distinguished by clinical signs or symptoms, the lowest of the possible grades will be assigned.

NCI CTCAE grades will be applied for the following lab parameters:

- Hematology: hemoglobin, WBC, lymphocyte, neutrophils, and platelets.
- Chemistry: Alanine Aminotransferase (ALT), albumin, alkaline phosphatase (ALP), amylase, Aspartate Aminotransferase (AST), total bilirubin, calcium, creatinine, glucose, lipase, phosphate, magnesium, potassium, sodium.

Laboratory measurements that are within their institutional limits of normal and are not graded as 1-4, per the CTCAE, will be summarized as "Grade 0," which is defined as normal.

10.8.7.2 Potential Hepatotoxicity (Hy's Law)

Cases of potential hepatotoxicity are defined by Hy's law if they meet the following criteria concurrently within the same lab sample:

- AST and/or ALT ≥ 3 × upper limit of normal (ULN)
- Total Bilirubin ≥ 2 × ULN
- ALP ≤ 2 × ULN

10.8.7.3 Pregnancy Testing

For subjects who are women of child-bearing potential, a serum pregnancy test will be performed by the local laboratory at Screening, a urine or serum pregnancy test will be performed prior to each subsequent cycle of therapy, and at End of Treatment. Pregnancy tests will also be done whenever pregnancy is suspected during the study.

10.8.8 Physical Examination and Vital Signs

Complete PEs will be conducted during screening, at the End of Treatment visit and other timepoints as per the Schedule of Assessments in the protocol. PEs will also be conducted as clinically indicated. During study treatment PEs will be symptom-directed. Abnormal results of PEs of body systems at screening examination will be captured on the Medical History eCRF pages. Any new clinically significant abnormality from baseline during study after screening examination will be collected as AEs and recorded on the 'Adverse Events' eCRF pages.

A PE will include an assessment of general appearance, skin, head, neck, throat, lymph nodes, cardiovascular, neurological, thyroid, musculoskeletal/extremities, respiratory and abdomen.

The vital signs collected on this study will include height (cm), weight (kg), body temperature (°C), sitting (or semi-recumbent for 5 minutes) systolic and diastolic blood pressure (mmHg), pulse rate (beats/min), pulse oximetry (%) and respiratory rate (breaths/min). These will be entered on the 'Vital Signs' eCRF pages. Baseline vital signs are the last collected any time prior to start of treatment on Cycle 1 Day 1 in each phase.

Abnormal vital signs are defined as:

- Systolic blood pressure (SBP) > 140 mmHg,
- Diastolic blood pressure (DBP) > 90 mmHg,
- Heart rate (or pulse rate) < 60 beats/min or > 100 beats/min, or
- Temperature > 37.5°C.

Potentially clinically important vital sign results are defined as increases in blood pressure and weight changes, in various degrees:

- SBP (>= 160 mmHg, >= 180 mmHg, >= 200 mmHg)
- DBP (>= 100 mmHg, >= 120 mmHg)
- Weight increased >= 10% from baseline
- Weight decreased >= 10% from baseline

10.8.9 Electrocardiograms

Single or triplicate 12-Lead electrocardiograms (ECGs) will be performed at scheduled timepoints as per the Schedule of Assessments in the protocol. For each planned visit for which a triplicate ECG is obtained, the average of the 3 values should be calculated for each parameter at each timepoint. For triplicate ECGs, baseline is defined as the mean of the set(s) of the triplicate readings obtained prior to first administration of study treatment. If two sets of triplicate ECGs were obtained at baseline, then the baseline represents the mean of all ECGs. The average values will be used for the ECG tables.

The following ECG parameters will be entered on the 'ECG (12-Lead)' eCRF pages:

- Mean heart rate (beats/min)
- Mean P-R interval (msec)
- Mean QRS duration (msec)
- Mean QT interval (msec)
- Mean QTcF using Fridericia's formula (msec)
- Interpretation of the ECG ('Normal', 'Abnormal, Not Clinically Significant', 'Abnormal, Clinically Significant')

QTcF will be graded per CTCAE v5.0. The grades are as follows:

- Grade 0: value < 450 msec
- Grade 1: value ≥ 450 to ≤ 480 msec
- Grade 2: value > 480 to ≤ 500 msec
- Grade 3: value > 500 msec or increase from baseline > 60 msec
- Grade 4: torsade de pointes, polymorphic ventricular tachycardia, signs/symptoms of serious arrhythmia. Grade 4 events will be reviewed by the study Medical Monitor from clinical data. As it

is based on clinical judgement and not based on QTcF numerical results, Grade 4 QTcF will not be presented in the summary table.

The following categories for change from baseline in QTcF are defined by the International Council for Harmonization of Technical Requirements for Pharmaceuticals for Human Use (ICH) E14:

- ≤ 30 msec
- > 30 to ≤ 60 msec
- > 60 msec

Baseline values are reported any time prior to start of treatment on Cycle 1 Day 1 in each phase.

10.8.10 Eastern Cooperative Oncology Group Performance Status

Eastern Cooperative Oncology Group (ECOG) performance status (grade or score ranging from 0-5) will be collected at scheduled timepoints as per the Schedule of Assessments in the protocol and reported on the 'ECOG Performance Status' eCRF pages.

10.8.11 Prior Anticancer Therapy and Subsequent New Anticancer Therapy

Prior anticancer therapy and subsequent new anticancer therapy will be collected on the 'Prior Anticancer Therapies' and 'Subsequent New Anticancer Therapy' forms of the eCRF respectively. For summaries of radiation therapy only, these will be selected as 'Radiation' from the 'Therapy Type/Name (Type of Therapy Received)' field of the 'Prior anticancer therapy' eCRF.

All summary details will be taken directly from the eCRF (see Section 13.2.2).

10.9 Pharmacodynamic Biomarkers and Immunogenicity Assessment

Serum of blood samples will be collected for the assessment of immunogenicity. Peripheral blood will be collected for immunophenotyping analysis to assess T cell subsets and their activation states and for the analysis of immune-related gene expression. Changes in peripheral blood monocyte subpopulations, serum cytokines, chemokines, tumor tissue cellular subpopulations and biomarker expression will also be examined.

In both Phase 1 and Phase 2, archival tumor samples or fresh tumor biopsies will be collected to characterize potential biomarkers of response.

A variety of factors that may potentially predict clinical response to XTX202 will be assessed in peripheral blood and in tumor specimens. Data from these assessments will be evaluated for associations with biological effects and response or safety associated with XTX202.

Biomarker samples will be collected for the following purposes:

1. Peripheral blood for immune cell subset characterization
2. Peripheral blood for gene expression profiling
3. Tumor tissue specimens for gene expression and TIL analysis
4. Serum/plasma biomarkers

The following serum-based tumor markers will be assessed for potential relationship to the response or safety associated with XTX202:

- Carcinoembryonic Antigen (CEA)
- Cancer Antigen 125 (CA-125)
- Cancer Antigen 15-3 (CA 15-3)
- Cancer Antigen 19-9 (CA 19-9)

- Cancer antigen 27-29 (CA 27-29)
- α-fetoprotein
- Prostate-Specific Antigen (PSA)
- Bladder tumor antigen
- Cytokeratin fragment 21-1
- Des-gamma-carboxy prothrombin
- Fibrin/fibrinogen
- Gastrin
- HE4
- Neuron-specific enolase
- Soluble mesothelin-related peptides
- Thyroglobulin
- Other

Serum of blood samples collected will also be used for the assessment of immunogenicity, specifically, used for the analysis of ADAs, including neutralizing ADAs, titers and their potential impact on PK, activity and safety associated with XTX202.

10.10 Other Derived Variables

10.10.1 Age

Age in years at informed consent is collected in the 'Demographics' page in the eCRF. If age is missing, it will be calculated as follows:

(Date of informed consent – birth year) / 365.2425

In addition to being analyzed as a continuous variable, baseline age will be categorized into the following groups:

- < 65 years
- 65 to <75 years
- >= 75 years

10.10.2 Body Mass Index

Body mass index (BMI) will be calculated based on baseline weight and baseline height using the following formula:

$$BMI = \text{Weight (kg)} / \text{Height (m)}^2$$

10.10.3 Differentials of Blood Cell Count

When both absolute and % differentials of any blood cell count, such as WBC, are collected for a particular visit then data will be handled as in the following:

- If absolute differential is reported (or absolute and % differentials) then report the absolute
- If only % differential is reported, then calculate absolute differential from % differential and total WBC count using the following formula:

  Absolute differential = % differential x total WBC count

10.10.4 Worst On-Treatment Reference Range/Grade

The worst on-treatment reference range or grade will be derived by taking the highest grading or reference range, including unscheduled visits, on or after the first dose of study drug and no later than 90 days after the last dose of study drug.

For reference ranges, the ranges will be classified from worst to best as high, low and then Normal.

10.10.5 Last On-Treatment Value

The last on-treatment value will be derived by taking the latest value, including unscheduled visits, on or after the first dose of study drug and no later than 90 days after the last dose of study drug.

10.11 Handling of Partial or Missing Dates

For the purpose of assigning AEs as 'treatment-emergent' and medications and therapies as 'prior' or 'concomitant', missing and partial start and stop dates will be imputed for the purpose of analysis. Imputed dates will not be presented in data listings.

In general, an AE will be considered treatment-emergent unless there is evidence in the partial dates available, that it was not treatment-emergent. In particular, in cases of missing start dates of AEs, these will be considered treatment-emergent, unless the stop date of the AE is prior to the first administration of study treatment. In the case of partially missing start dates, the AE will be considered treatment-emergent, unless the information from the partial dates clearly shows that the AE was not treatment-emergent.

In general, medications and therapies will be considered current or concomitant, unless there is evidence in the partial dates available that the medical condition or treatment was stopped prior to the first administration of study treatment.

10.11.1 Adverse Event and Concomitant Medication Partial Start and Stop Dates

The following rules will be applied to impute missing start and stop dates in appropriate data types (e.g., AEs or concomitant medications):

Start Date
If the start date is completely missing (i.e., the day, month, and year are all unknown) the start date will be set to the date of first dose of study medication.

Missing Day Only

- If the month and year of the incomplete date are the same as the month and year of the first dose date, then the day of the first dose date will be assigned to the missing day.
- If either the year is before the year of the first dose date or if years are the same but the month is before the month of the first dose date, then the last day of the month will be assigned to the missing day.
- If either the year is after the year of the first dose date or if both years are the same but the month is after the month of the first dose date, then the first day of the month will be assigned to the missing day.

Missing Month Only

- The day will be treated as also missing and both month and day will be replaced according to the below procedure.

Missing Day and Month

- If the year of the incomplete date is the same as the year of the first dose date, then the day and month of the first dose date will be assigned to the missing fields.
- If the year of the incomplete date is not the same as the year of the first dose date, then January 1 will be assigned to the missing fields.

- If the stop date is complete and the imputed start date is after the stop date, then the start date will be imputed using the stop date.

Stop Date

Missing Day Only

- If the month and year of the incomplete date are the same as the month and year of the last visit date, then the day of the last visit date will be assigned to the missing day.
- If either the year is before the year of the last visit date or if both years are the same but the month is before the month of the last visit date, then the last day of the month will be assigned to the missing day.
- If either the year is after the year of the last visit date or if both years are the same but the month is after the month of the last visit date, then the first day of the month will be assigned to the missing day.

Missing Month Only

- The day will be treated as missing and both month and day will be replaced according to the below procedure.

Missing Day and Month

- If the year of the incomplete date is the same as the year of the last visit date, then the day and month of the last visit date will be assigned to the missing fields.
- If the year of the incomplete date is before the year of the last visit date, then December 31 will be assigned to the missing fields.
- If the year of the incomplete date is after the year of the last visit date, then last visit date will be assigned to the missing fields.

Medications with completely missing start dates will be listed in both prior medications and concomitant medications.

10.11.2 Partial Dates of Diagnosis, Disease History and Death

The partial start date for diagnosis of metastatic disease and prior disease history will be assigned to 15th day of the month (if only day is missing) or July 1st (if both month and day are missing).

If a death date is missing month or day, the following imputation method will be used:

- If only the death year is known and the last date the subject was known to be alive is in the same year, the subject's last alive date will be used as the death date. If the last alive date occurs in a previous year, the missing death month and day will be imputed as the first day and month of the year (01JAN).
- If the year and month are known and the last alive date is in the same year and month as the death date, then the last alive date will be used. If the last alive date is in a month prior to the death month, then the death day will be imputed as the first day of the month.

For death dates, every effort will be made to confirm the full death date prior to database lock.

10.12 Handling of Missing Adverse Event Severity and Relationship

If Severity is missing the event will be assumed to be missing. If relationship is missing the event will be assumed to be Related.

10.13 Handling of Other Missing Data

10.13.1 Missing Absolute Differential

See Section 10.10.3.

10.13.2 Laboratory Values with Character Symbols

In general, missing laboratory data will not be imputed. However, laboratory values of the form of "< x" (i.e., below the lower limit of quantification) and "> x" (i.e., above the upper limit of quantification) will be imputed to x - 0.001 and x + 0.001 respectively for the calculation of summary statistics and comparing to normal ranges. For example, a value recorded as "<10" will be imputed to 9.999 and a value collected as ">0.2" will be imputed to 0.201. If the original values are collected as SI units, then they will still be displayed as "< x" or "> x" in the listings. Otherwise, due to the listings only displaying the SI unit values, they will be presented as the imputed value after conversion to the SI unit values.

10.14 Handling of Unscheduled Visits and Unscheduled Assessments

Scheduled (planned) and unscheduled visits data will be included in listings.

If scheduled and unscheduled assessments are available for the same parameter within a single visit, then only the last scheduled assessment will be used in the analysis and unscheduled assessments will not be mapped to scheduled ones no matter if these are within or outside the visit window. Unscheduled assessments will be summarized separately from the scheduled.

The visit window allowed for each particular parameter is specified in the Schedule of Assessments in the protocol.

11.0 Analysis Sets

11.1 DLT-Evaluable Analysis Set

All subjects in Phase 1 (Dose Escalation Part 1a) who received at least 1 dose of XTX202 and either completed the DLT observation period or experienced a DLT during this period. This analysis set will be used to assess the tolerability of XTX202 in the dose escalation Phase. Over the entire dose escalation Phase, subjects will be grouped based on the highest DL each subject received.

11.2 All-Treated Analysis Set

All subjects who received any amount of XTX202, with treatment group based on the highest DL received. This analysis set will be the primary analysis set for all safety endpoints, excluding DLT evaluation, as well as efficacy sensitivity analysis for PFS.

11.3 PK Analysis Set

All enrolled subjects who received at least 1 dose of XTX202 and have at least 1 postdose measurement of 1 PK analyte without protocol deviations or events affecting the validity of the PK results. This analysis set will be used for the analysis of PK parameters, which will be described in a separate PK analysis plan.

The same blood sample for PK analysis will be used for immunogenicity testing. Plasma will be used for PK analysis and serum will be used for immunogenicity testing.

11.4 Pharmacodynamic Analysis Set

All enrolled subjects who received at least 1 dose of XTX202 and have at least 1 postdose PaD or biomarker measurement. This analysis set will be used for the analysis of PaD endpoints as further defined in the protocol.

11.5 Response-Evaluable Analysis Set

All subjects with measurable disease at baseline who received any amount of XTX202 and had at least 1 post-Baseline response assessment or discontinued treatment due to disease progression (including death caused by disease progression) prior to the first efficacy evaluation. Treatment group will be based on the initial dose received. This analysis set will be the primary analysis set for ORR, DCR and DOR.

11.6 Intent-To-Treat Analysis Set

All subjects who received any amount of XTX202, with treatment group based on the initial dose received. This analysis set will be the primary analysis set for OS and PFS.

12.0 Interim Analysis

During the Phase 1 Dose-Escalation, study data will be continuously evaluated for safety or MTD/RP2D reviews. At the end of Phase 1, reports summarizing the Safety, PK, PaD data will be produced and an interim analysis report may be generated. Interim informal evaluations of the data may also be performed prior to the end of Phase 1.

The data cut-off date for interim analysis during Phase 1 will be the timepoint when all subjects completed at least their first two 3-week cycles, i.e., 6 weeks after the last subject in Phase 1 has received their first administration of XTX202.

For each of the Phase 2 efficacy expansion cohorts, an interim analysis is planned after the first 17 subjects are enrolled in Stage 1 and have completed 3 months' follow-up as per the Optimal Simon's 2-stage design (See Section 8.4.2). Efficacy endpoints will be evaluated according to RECIST version 1.1, and enrollment will be terminated for the respective part if there are no responders; otherwise, the study will continue into Stage 2 with additional 18 subjects enrolled and treated.

13.0 Statistical Methods

All data will be presented in descriptive summaries, which will be tabulated by Phase (Part 1a and Part 1b) and DL cohort (highest DL or initial DL dependent on analysis set)and in some cases overall. Phase 2 will also be summarized by tumor type (Part 2a and Part 2b) as collected as 'Melanoma' or 'Renal Cell Carcinoma' respectively on the 'Primary Cancer Diagnosis' field of the 'Cancer History' eCRF page. If multiple RP2Ds are selected from Phase 1, then Phase 2 will also be presented by RP2D.

All data collected during the study will be presented in data listing unless otherwise specified. Data listings will be sorted by descending DL and then by subject within DL.

Screen failures will be excluded from all tables and listings except those summarizing screen failures or subject disposition.

Categorical data will be summarized using counts and percentages, with the number of subjects in each analysis set as the denominator for percentages, unless otherwise specified. Percentages will be rounded to one decimal place except that 100% will be displayed without any decimal places. Percentage will not be displayed for zero counts. Counts of missing observations will be included in the denominator and presented in a relevant category, such as 'Not Reported', unless otherwise specified.

Continuous data will be summarized using the number of observations (n), mean, standard deviation (SD), median, first and third quartile (Q1, Q3), minimum and maximum (Min, Max). Minimum and maximum will be rounded to the precision of the original value. Mean, median and quartiles will be rounded to 1 decimal place more than the precision of the original value. The SD will be rounded to 2 decimal places more than the precision of the original value.

Unless otherwise specified, missing data will not be imputed.

All statistical analyses will be performed using SAS® Version 9.4 or higher.

13.1 Subject Disposition

The number of subjects screened, screen-failed and enrolled will be summarized. For all enrolled and treated subjects, the number and percentage of subjects having completed or prematurely discontinued the study treatment (XTX202) and/or completed or discontinued from the study will be presented along with a breakdown of the corresponding reasons for completion and/or discontinuation from treatment and/or the

study. The number and percentage of subjects in each of the analysis sets will also be summarized. Subjects enrolled at each site will also be summarized based on all enrolled subjects.

Enrollment information such as date of informed consent signed, details pertaining to DL assignments, date of completion of study treatment and completion of the study, and reasons for treatment discontinuation and/or withdrawal from the study will be presented in listings. For all enrolled subjects, details for analysis sets to which they belong will also be presented in listings.

Note: All enrolled subjects is defined as all subjects who have a non-missing enrollment date in the eCRF and complete screening i.e. not screen-failures.

13.2 Demographics and Baseline Characteristics

All summaries in this section will be summarized in the All-Treated Analysis Set unless otherwise specified.

13.2.1 Demographics

Demographic and PE results will be summarized using descriptive statistics, by study part and DL cohort within each Part. Age (years), age categories, gender, child-bearing potential (females), ethnicity, race, height (cm), weight (kg), BMI (kg/m^2), and ECOG will be summarized at baseline.

13.2.2 Prior and Concomitant Medications, Therapies and Procedures

The count and percentage of subjects using at least one concomitant medication will be summarized. Concomitant medications will also be summarized by Anatomical Therapeutic Chemical (ATC) Classification system (Level 4) and WHO Drug Dictionary PT. Prior medications will also be summarized in a similar way.

Concomitant procedures will be tabulated by primary reason procedure was performed.

The count and percentage of subjects that used at least one prior systemic anticancer therapy will be summarized. Type of therapy, intent, number of regimens, number of cycles, best overall response by RECIST version 1.1 and subjects who progressed on prior anticancer therapy will also be presented. A similar table will be produced for radiation therapy only.

The count and percentage of subjects that used at least one subsequent new systemic anticancer therapy will be summarized. Therapy type, therapy setting, number of regimens and response will also be presented.

13.2.3 Medical and Cancer History

The count and percentage of subjects with any medical history will be summarized. Medical history will also be summarized by SOC and PT.

Current cancer history will be presented by primary cancer diagnosis, time since initial cancer diagnosis (years), stage at study entry and time since most recent progression/recurrence (years).

13.3 Efficacy Analyses

The efficacy endpoints of ORR, DOR and DCR will be summarized for the Response-Evaluable Analysis Set. OS and PFS will be summarized for the ITT Analysis Set. A sensitivity for PFS will be done on the All-Treated Analysis Set. Phase 1 Part 1a and Part 1b will be tabulated separately and presented by DL and overall. Phase 2 will be presented by tumor type (Part 2a and 2b) and overall. If multiple RP2Ds are selected from Phase 1, then Phase 2 will also be presented by RP2D. If within Phase I there are more than 3 subjects within a tumor type, then efficacy will also be summarized by tumor type for Phase I. Subjects from Part 1 treated at RP2D with either melanoma or RCC can be included in the respective tumor cohort and overall in the final analysis.

Note that OS will not be summarized for Phase 1.

Uncertainty of estimates associated with categorical variables will be assessed using the two-sided 95% Clopper-Pearson confidence interval (Clopper CJ & Pearson ES, 1934). Time-to-event analysis will be

performed using the Kaplan-Meier (KM) method with the associated two-sided 95% CIs estimated using the Brookmeyer and Crowley method (Brookmeyer R & Crowley J, 1982).

No inferential statistical analysis is planned for comparisons between tumor types or DL cohorts.

All derived efficacy data will also be listed by phase, descending initial DL and then by subject within DL.

13.3.1 Hypothesis Testing Strategy

No hypothesis will be formally tested, and the statistical analysis will be primarily descriptive. However, in this study hypothesis testing is used for future study design purposes, but not for inferential purposes.

13.3.2 Best Overall Response and Objective Response Rate

Responses categorized per Investigator assessment using RECIST version 1.1 will be tabulated. The number and percentage of subjects in each category will be summarized. ORR will be presented, with associated exact binomial 95% CI (Clopper CJ & Pearson ES, 1934). Subjects who cannot be assessed for response will be counted as NE.

Swimmer plots for the time to BOR, Spider Plots for percent change from baseline in target lesion size and Waterfall Plots for maximum percent change from baseline in target lesion size will also be presented.

13.3.3 Disease Control Rate

The DCR will be summarized overall, and at 3, 6, 9 and 12 months by frequency and percentage. Exact binomial 95% CI for DCR will also be displayed.

13.3.4 Duration of Response

DOR will be summarized for the subset of subjects in the Response-Evaluable Analysis Set with a confirmed response (CR or PR). The number and percentage of subjects with an event and censored, and reasons censored will be presented. The distribution of DOR will be estimated using the KM method. The 25th, median, and 75th percentiles with 95% CI (Brookmeyer R & Crowley J, 1982) will be calculated. The proportion of subjects with a response at 3, 6, 9 and 12 months will be presented, defined as the KM estimate of DOR at 3, 6, 9 and 12 months respectively. The 95% CI for the survival function estimated at these pre-specified timepoints will be derived using the log-log transformation, with standard error estimated using the Greenwood's formula (Kalbfleisch & Prentice, 1980). A KM plot will be presented, with the number of subjects at-risk displayed on the plot.

13.3.5 Progression-Free Survival

PFS will be estimated using the KM method and analyzed using the same methods as for DOR using the ITT Analysis Set. For Phase 1, Parts 1A and 1B, a sensitivity analysis will also be performed based on the All-Treated Analysis Set, where DL will be based on the highest dose received. Only the tables will be repeated for the All-Treated Analysis Set.

13.3.6 Overall Survival

OS will be estimated using the KM method and analyzed using the same methods as for DOR for Phase 2 only using the ITT Analysis Set.

13.4 Treatments

13.4.1 Extent of Study Drug Exposure

Study treatment exposure will be summarized descriptively for the All-Treated Analysis Set, including the following variables:

- Number of subjects receiving at least one dose
- Duration of exposure (weeks)

- Number of doses completed
- Number of cycles started
- Cumulative dose received (mg/kg)
- Relative dose intensity (%)
- Number of subjects with at least one dose not administered with reasons for dose not administered
- Number of subjects with at least one interruption with reasons for interruption
- Number of dose interruptions per subject
- Duration of dose interruptions per subject (minutes)
- Number of subjects with at least one dose adjustment with reasons for adjustment
- Number of dose adjustments per subject

See section 10.6.9 for derivations of treatment exposure variables.

Separate listings will be used to present exposure data by subject for study treatment administration.

13.5 Important Protocol Deviations

Protocol deviations will be identified, including, but not limited to:

- Subjects that were enrolled even though they did not meet all eligibility criteria
- Subjects who received the wrong study treatment or an incorrect dose
- Subjects who took concomitant treatments specifically prohibited by the protocol, such as:
  - Other investigational agents
  - Other systemic or topical antineoplastic therapy
  - Systemic glucocorticoids

The reason for each protocol deviation will also be identified.

Per ICON processes, important protocol deviations data will be entered into our PREDICTIVV Study Operation (PSO), the Clinical Trials Management System (CTMS). The study team and the Sponsor will conduct ongoing reviews of the deviation data from these systems, and reviews of the resulting set by excluding subjects with protocol deviations deemed important throughout the study, and by adjusting the deviation criteria as appropriate. The resulting set must be finalized at the post-freeze data review meeting (or earlier), prior to database lock.

Protocol deviations will be classified into important or not important protocol deviations based on their possible impact on the study results in a data review meeting prior to the database lock.

A tabulation of the number and percentage of subjects with at least one important protocol deviation, by deviation type and deviation category will be provided for each study Part. For each category and deviation subjects will only be counted once even if they have multiple deviations in the category or deviation. The number and percentage of subjects with no important protocol deviations will also be summarized for the All-Treated Analysis Set.

A Data listing by subject will present all protocol deviations including whether deviation is an important protocol deviation. A listing of all Coronavirus Disease-2019 (COVID-19) related protocol deviations will also be provided.

COVID-19 related protocol deviations will be defined in CTMS using a standard naming convention. These deviation descriptions will start with ‘Covid 19:’. The ICON programming team will use this standard naming convention to flag for COVID-19 deviations within our analysis datasets.

13.6 Safety Analyses

Safety analyses will be performed for the All-Treated Analysis Set, unless otherwise stated.

The primary safety endpoints are type, incidence, severity, seriousness and relationship to treatment of AEs evaluated until the end of treatment, and change from baseline in the parameters of vital signs, laboratory in blood and urine, PE, ECGs, ECOG performance status, and biomarkers, immunogenicity in blood.

13.6.1 Adverse Events

Safety will be assessed primarily based on AEs. All summaries of AEs will be based on TEAE unless otherwise indicated. TEAEs will be summarized by SOC and PT, according to severity as per NCI CTCAE version 5.0 and its relationship to study treatment.

A breakdown of the number and percentage of subjects reporting each TEAE, categorized by SOC and PT coded according to the latest version of MedDRA dictionary, and the maximum severity graded according to the NCI CTCAE guidelines, will be presented. A summary of treatment-emergent serious adverse events (TESAEs), categorized by SOC and PT will be presented. Note that counting of TEAEs/TESAEs will be by subject and number of events. Subjects will only be counted once within each SOC or PT category and total number of events within each PT and SOC will be presented.

The number and percentage of subjects who report AEs will be summarized as follows:

- TEAEs overall summary including: Any TEAE, Relationship to Study Treatment, NCI CTCAE v5.0 Toxicity Grade 3 or above, Related and Grade 3 or above, Leading to Death, Leading to Death and Related, Seriousness, Serious and Related, Leading to Discontinuation of Study Treatment, Related and Leading to Discontinuation of Study Treatment, Leading to Dose Reduction, Related and Leading to Dose Reduction, Leading to Dose Interruption, Related and Leading to Dose Interruption, DLTs, ECIs, and Related ECIs
- TEAEs by SOC and PT
- TEAEs with CTCAE Toxicity Grade >= 3 by SOC and PT
- Treatment-Related TEAEs by SOC and PT
- Treatment-Related TEAEs with CTCAE Toxicity Grade >= 3 by SOC and PT
- TEAEs with an Outcome of Death by SOC and PT
- Treatment-Related TEAEs with an Outcome of Death by SOC and PT
- TESAEs by SOC and PT
- Treatment-Related TESAEs by SOC and PT
- TEAEs Leading to Discontinuation of Study Treatment by SOC and PT
- Treatment-Related TEAEs Leading to Discontinuation of Study Treatment by SOC and PT
- TEAEs Resulting in Any Dose Reduction by SOC and PT
- Treatment-Related TEAEs Resulting in Any Dose Reduction by SOC and PT
- TEAEs Resulting in Any Dose Interruption by SOC and PT
- Treatment-Related TEAEs Resulting in Any Dose Interruption by SOC and PT
- Treatment-Emergent ECIs by SOC and PT
- Treatment-Related Treatment-Emergent ECIs by SOC and PT

For grade summaries, subjects with multiple events within a particular SOC or PT category will be counted at their most severe event under that SOC or PT category.

Events will be sorted by descending frequency of the total across all cohorts, by SOC and then by PT within SOC for summaries.

All AEs will be listed by subject, including information for onset, duration, severity, seriousness, relationship to study treatment, outcome and actions taken. The following will be listed separately: DLTs, SAEs, ECIs and Death Records.

If an AE is coded using more than one PT for a single VT, then the PTs will be grouped together within the same VT line for the listings only.

13.6.2 Deaths

Number and percentages for each primary cause of death will be summarized in a table from the 'Death Details' page of the eCRF.

13.6.3 Dose-Limiting Toxicity

DLT-Evaluable Analysis Set will be used for determination of the MTD during the DLT Observation Period. The number and percentage of subjects who experience a DLT will be tabulated by SOC and PT. Individual subject data will be reported in data listing, including the description of DLT event, severity and relationship to XTX202. Only DLTs in the DLT observation period will be included in tables (see Section 10.8.5.1).

13.6.4 Laboratory Data

All laboratory results for hematology, coagulation and chemistry will be presented for change from baseline will be summarized in tables and presented in listings. Table summaries will be presented at each scheduled assessment.

Shift tables for hematology and chemistry tests will be provided for the maximum post-baseline grade (see Section 10.10.4) where CTCAE grading is available, including data from any unscheduled post-baseline timepoint. These tables will compare the NCI CTCAE Toxicity Grade of the baseline measure relative to the maximum post-baseline value that is non-missing. Only subjects with post-baseline laboratory values will be included in these analyses. Results of urinalysis and other tests that are not part of NCI CTCAE, v5.0 will be listed only. Shift tables for coagulation will be provided for the maximum post-baseline reference range (See Section 10.10.4) including data from any unscheduled post-baseline timepoints.

Individual subject hematology, coagulation, chemistry and urinalysis values will be presented in data listings, which will include data from both scheduled and unscheduled timepoints. Results with CTCAE Toxicity Grade ≥ 3 or Abnormal will also be listed separately.

13.6.4.1 Pregnancy Testing

Only listing will be presented for pregnancy test results.

13.6.4.2 Thyroid

Thyroid function results will be summarized at each scheduled assessment visit and also using shift tables to worst on-treatment reference range (see Section 10.10.4) including data from any unscheduled post-baseline timepoints.

13.6.4.3 Potential Hepatotoxicity (Hy's Law)

Subjects with potential hepatotoxicity (as defined by Hy's law) during the treatment period will be summarized in listings only, where ALT/AST/Bilirubin and ALP will be presented at all visits for subjects meeting Hy's law criteria (see Section 10.8.7.2) .

13.6.5 Vital Signs

All vital signs and changes from baseline through the last cycle will be summarized.

Post-baseline abnormal and potentially clinically important vital sign results (see Section 10.8.8) will also be summarized.

For baseline, the last timepoint prior to first dose on C1D1 will be used. If missing, the data at the Screening visit will be used as baseline.

13.6.6 Physical Examinations, ECGs, and Other Observations Related to Safety

Additional safety assessments include PEs, ECG measurements and ECOG performance status.

13.6.6.1 Physical Examinations

Only listings will be provided for PE results.

13.6.6.2 ECGs

Single or average of triplicate 12-lead ECGs parameters will be summarized. A summary of ECG parameters including heart rate (beats/min), P-R interval (msec), QRS duration (msec), QT interval (msec) and QTcF interval (msec) will be presented for each planned visit. Change from baseline will be presented for QTcF only.

The number and percentage of subjects who have abnormal single or averaged triplicate QTcF values during the study will be summarized based on the ICH E14 Category and by the CTCAE v5.0 Grade (see Section 10.8.9). Shift table will be presented to summarize change in single or averaged triplicate QTcF values from baseline to worst CTCAE Grade.

All single or averaged triplicate ECG data will be provided in data listing, including their overall interpretation.

13.6.6.3 ECOG Performance Status

Shift tables comparing ECOG data from baseline to last on-treatment ECOG score will be presented (see Section 10.10.5).

13.7 Pharmacodynamic Analyses and Immunogenicity

PD and Immunogenicity analyses will be based on the PD and PK Analysis Sets. These will be performed by an external vendor (CATO-SMS). Relevant analysis plan and analysis outputs (TFLs) will be provided by this vendor.

13.8 Pharmacokinetic Analyses

All PK analyses will be based on the PK Analysis Set. These will be performed by an external vendor (CATO-SMS). Relevant analysis plan and analysis outputs (TFLs) will be provided by this vendor.

13.9 Methods for Handling Dropouts and Missing Data

Data of subjects who are replaced will not be included in the summary analyses; data of replaced subjects may be presented in data listing if appropriate.

Sensitivity analyses of response parameters including all subjects (with replaced subjects analyzed as non-responders) may be generated as deemed appropriate per study results.

13.10 Multiplicity

There will be no adjustment made for multiplicity as there are no formal statistical tests planned.

13.11 Impact by COVID-19

A sensitivity analysis may be carried out at the end of the study if the number of subjects impacted by the COVID-19 pandemic is large (e.g., >10% missed a dose or assessment due to COVID-19) and thought to be impacting the safety and efficacy endpoints. This sensitivity analysis may exclude the subjects affected by the pandemic.

Data on COVID-19 infection (test performed and result) will be listed.

14.0 References

Brookmeyer R & Crowley J. (1982). Confidence interval for the median survival time. *Biometrics*, 38:29-41.

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

Eisenhauer, E. A., Therasse, P., Bogaerts, J., Schwartz, L. H., Sargent, D., Ford, R., . . . Mooney, M. (2009). New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *European journal of cancer*, 228-247.

Kalbfleisch, J. D., & Prentice, R. L. (1980). *The statistical analysis of failure time data.* New York: John Wiley.

Simon, R. (1989). Optimal two-stage designs for phase II clinical trials. *Control Clin Trials*, 1-10.

US Food and Drug Administration. (2023). *Optimizing the dosage of human prescription drugs and biological products for the treatment of oncologic diseases.*

15.0 Glossary of Abbreviations

ADA	Antidrug Antibody
AE	Adverse Event
ALP	Alkaline phosphatase
ALT	Alanine Aminotransferase
AST	Aspartate Aminotransferase
AUC	Area Under the Curve
BOR	Best Overall Response
C1D1	Cycle 1, Day 1
CL	Systemic Clearance
CLS	Capillary Leak Syndrome
C_{max}	Maximum Observed Serum Concentration
COVID-19	Coronavirus Disease-2019
CR	Complete Response
CT	Computed Tomography
CTCAE	Common Terminology Criteria for Adverse Events
CTLA-4	Cytotoxic T Lymphocyte Associated Protein 4
C_{trough}	Trough Concentration
DCR	Disease Control Rate
DL	Dose Level
DLT	Dose-Limiting Toxicity
DOR	Duration of Response
ECG	Electrocardiogram
ECI	Events of Clinical Interest
ECOG	Eastern Cooperative Oncology Group
eCRF	Electronic Case Report Form
EDC	Electronic Data Capture
EOT	End-of-Treatment
FDA	Food and Drug Administration
ICH	International Council for Harmonization of Technical Requirements for Pharmaceuticals for Human Use
ITT	Intent-To-Treat
IL-2	Interleukin-2
IV	Intravenous
KM	Kaplan-Meier
MRI	Magnetic Resonance Imaging
MTD	Maximum Tolerated Dose

MedDRA	Medical Dictionary for Regulatory Activities
NCI-CTCAE	National Cancer Institute Common Terminology Criteria for Adverse Events
NE	Not Evaluable
NTL	Non-Target Lesion
ORR	Objective Response Rate
OS	Overall Survival
PaD	Pharmacodynamic(s)
PD	Progressive Disease
PD-1	Programmed Cell Death-Protein 1
PE	Physical Examination
PFS	Progression-Free Survival
PK	Pharmacokinetic(s)
PR	Partial Response
PT	Preferred Term
Q3W	Every 3 Weeks
RCC	Renal Cell Carcinoma
RECIST	Response Evaluation Criteria in Solid Tumors
RP2D	Recommended Phase 2 Dose
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SD	Stable Disease
SRC	Safety Review Committee
$t_{1/2}$	Half-Life
TEAE	Treatment-Emergent Adverse Event
TESAE	Treatment-Emergent Serious Adverse Event
TFL	Table, Figures, and Listings.
TIL	Tumor-Infiltrating Lymphocyte
TKI	Tyrosine Kinase Inhibitor
TL	Target Lesion
T_{max}	Time of Maximum Observed Concentration
TPRO	Time Point Overall Response
TPR	Time Point Response
uCR	Unconfirmed Complete Response
ULN	Upper Limit of Normal
uPR	Unconfirmed Partial Response
Vd	Volume of Distribution
WHO	World Health Organization