Statistical Analysis Plan – XT-150-2-0204

TITLE: A Double-Blind, Placebo-Controlled Assessment of the Tolerability and Efficacy of XT-150 for the Treatment of Moderate to Severe Pain Due to Osteoarthritis of the Knee

STUDY SPONSOR:

Xalud Therapeutics, Inc.
2120 University Avenue, Suite 532
Berkeley CA 94704
USA

SPONSOR CONTACT:

[redacted]

Document History:

Document	Version	Version Date
Statistical Analysis Plan – XT-150-2-0204	Final 2.0	11 Oct 2021
	Final 1.0	20 Nov 2020

CONFIDENTIALITY STATEMENT:

All information concerning the study drug supplied by the Sponsor in connection with this study, and not previously published, is considered confidential and proprietary information. This information includes the Investigator's Brochure, clinical protocol, subject information and informed consent, and case report forms. This confidential information shall remain the sole property of the Sponsor, shall not be disclosed to others without prior written consent from the Sponsor and shall not be used except in the performance of this study.

The information developed during the conduct of this clinical study is also considered confidential and will be used by the Sponsor in connection with the development of the study drug. This information may be disclosed as deemed necessary by the Sponsor.

This Statistical Analysis Plan has been prepared by [redacted] for the XT-150-2-0204 study (Xalud Therapeutics, Inc.) The contents of this procedure are confidential and reproduction, in part or whole, is not permitted without prior written permission from [redacted] and the Sponsor.

DocuSigned by:
[redacted]
Signing Reason: I am the author of this document
Signing Time: 11-Oct-2021 | 23:09:40 ACDT
BFCD598BC92B4F0494B97845786FF6E0

11-Oct-2021 | 23:09:45 ACDT

Prepared by: [redacted] Date

DocuSigned by:
[redacted]
Signing Reason: I have reviewed this document
Signing Time: 12-Oct-2021 | 10:06:23 ACDT
E96ABDB3EAE44B008862EB136B39B348

12-Oct-2021 | 10:06:27 ACDT

Reviewed by: [redacted] Date

DocuSigned by:
[redacted]
Signer Name [redacted]
Signing Reason: I approve this document
Signing Time: 11-Oct-2021 | 16:42:44 PDT
C7567E0BFE5D41D9870B876B8F76505E

11-Oct-2021 | 16:42:51 PDT

Approved by: [redacted] Date

[redacted]

TABLE OF CONTENTS

[redacted]

1. SCOPE OF THE STATISTICAL ANALYSIS PLAN

This Statistical Analysis Plan is an adjunct to Protocol No. XT-150-2-0204 (Xalud Therapeutics, Inc.) dated 25 Mar 2020, Amendment 2. The Statistical Analysis Plan details the procedures for the statistical methods used in the presentation and analysis of the clinical data.

[redacted]

Revision of this SAP will not be required for any subsequent amendments to the protocol which do not change the analyses described in this SAP. Revisions to specifics such as blood sample collection times will not require an update to the SAP.

2. ABBREVIATIONS

ADaM	Analysis Data Model
AE	Adverse event
BMI	Body mass index
BPI	Brief pain inventory of intensity and interference
CDISC	Clinical Data Interchange Standards Consortium
CRF	Case report form
ICH	International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use
ITT	Intent-to-Treat
KOOS	Knee outcome and osteoarthritis scores
MedDRA	Medical Dictionary for Regulatory Activities
NCI CTCAE	National Cancer Institute Common Terminology Criteria for Adverse Events
OA	Osteoarthritis
QA	Quality assurance
QC	Quality control
ROM	Range of motion
SAE	Serious adverse event
SAP	Statistical analysis plan
SDTM	Standard Data Tabulation Model
SOP	Standard operating procedure

TEAE Treatment emergent adverse event

WOMAC The Western Ontario and McMaster Universities Osteoarthritis Index

3. RELEVANT [redacted] STANDARD OPERATING PROCEDURES AND GUIDANCE

The following [redacted] SOPs and regulatory guidance are relevant to this Statistical Analysis Plan:

- [redacted]
- [redacted]
- [redacted]
- [redacted]
- ICH: Statistical Principles for Clinical Trials (ICH E9, 5 February 1998); Addendum on Estimands and Sensitivity Analysis in Clinical Trials (ICH E9(R1), 20 November 2019)

4. INTRODUCTION

4.1. Study Overview

This is a Phase 2 safety, tolerability, and efficacy study of XT-150 in adult participants with pain due to osteoarthritis of the knee. The study will be conducted in 2 Stages, A and B.

Stage A is placebo controlled for six months following intra-articular injection in the Index Knee, which is defined by Kellgren-Lawrence grading and participant-reported pain. The primary efficacy endpoint is 30% response rate in WOMAC pain at Day 180.

Stage B continues for an additional six months of safety and efficacy monitoring during which participants can elect for a second injection to the Index Knee. Active 150 µg or 450 µg XT-150 will be randomly chosen for the second injection.

At six months of exposure, after a single injection dose, the double blinded, placebo-controlled portion of this randomized RCT will end, and an analysis described below will be performed of both efficacy and safety. Subjects continue in the study in a blinded fashion for an additional six months of follow-up. During the second 6-month period, subjects may request a second injection, which will be 450 µg or 150 µg (this second dose having been randomized previously as described in the treatment dose sequence, section 4.3).

4.2. Study Objectives

Safety:

- To confirm the safety and tolerability of intra-articularly-injected XT-150

Efficacy:

- To establish analgesic efficacy of intra-articularly-injected XT-150

4.3. Study Design

Participants will provide informed consent and meet all study eligibility criteria before any study procedures are initiated. Baseline confirmation of study eligibility will be completed the day before or day of study drug administration. The Index Knee for study drug injection will be the OA-affected joint identified by Kellgren-Lawrence grade of 2 or 3; and WOMAC Pain Subscale $\geq$ 8 on a scale of 20.

Study drug will be administered by intra-articular (IA) injection into the joint space of the Index Knee.

Participants will undergo the study in 2 stages, A and B, over a 12-month period:

Stage A: Placebo-controlled for 6 months

Stage B: Continued follow up for 6 months with the option of randomly receiving a single 150 µg or 450 µg injection as the second injection to the Index Knee (knee dosed in Stage A)

In Stage A (6 months duration), about 270 participants will be randomly enrolled into 1 of 6 treatment groups at dose levels (approximately 90 participants/ dose level) to receive a 1 mL single injection containing 150 µg XT-150, 450 µg XT-150, or placebo on Day 0.

In Stage B (further 6 months duration), participants will have continued follow up for 6 months with the option of receiving a pre-randomized at study start single 150 µg or 450 µg injection as the second injection (1 mL) to the Index Knee (knee dosed in Stage A) at any time up to Day 330.

Safety parameters and efficacy endpoints for each participant will be evaluated monthly from the first XT-150 dose (Stage A, Day 0) for about 1 year (end of Stage B, Day 360).

4.4. Study Treatment

Treatments in Stage A:

- Placebo (1 mL phosphate-buffered saline for injection)
- 150 µg XT-150 (1 mL)
- 450 µg XT-150 (1 mL)

Treatments in Stage B:

- 150 µg XT-150 (1 mL)
- 450 µg XT-150 (1 mL)

At enrolment in Stage A, participants will be randomly enrolled into 1 of 6 treatment sequences (approximately 45 participants/ group) to receive one of 3 treatments in Stage A, and, optionally, one of 2 treatments in Stage B.

At any time in Stage B up to Day 330, inclusive, participants can elect for a single injection to the knee dosed in Stage A (Index Knee). The second injection dose will be either 150 µg or 450 µg XT-150, according to the randomized treatment sequence allocated on enrolment in Stage A. There will be no placebo comparator in Stage B.

Treatment sequences for randomization:

- XT-150 (150 µg) in Stage A, followed by XT-150 (150 µg) in Stage B
- XT-150 (150 µg) in Stage A, followed by XT-150 (450 µg) in Stage B
- XT-150 (450 µg) in Stage A, followed by XT-150 (150 µg) in Stage B
- XT-150 (450 µg) in Stage A, followed by XT-150 (450 µg) in Stage B
- Placebo in Stage A, followed by XT-150 (150 µg) in Stage B
- Placebo in Stage A, followed by XT-150 (450 µg) in Stage B

4.5. Sample Size

A total of 270 subjects will be randomized to receive one of 6 treatment sequences in a 1:1:1:1:1:1 ratio. This will result in subjects randomly receiving one of the three treatments in Stage A in a 1:1:1 ratio The sample size of 90 subjects per treatment in Stage A provides more than 90% power to detect a greater responder rate in at least one XT-150 group compared to placebo and 85% power to detect both arms superior to placebo, assuming the WOMAC Pain Score at Day 180 responder rate is 30% in each XT-150 group and 10% in the placebo group.

Treatment Effect That Can Be Detected

Placebo response rate (%)	XT-150 response rate (%) in both treatment arms	Statistical Power (%)	
		To detect at least one treatment arm superior to placebo	To detect both treatment arms superior to placebo
10	25	78	57
	30	95	85
	35	99	97
15	25	38	18
	30	70	46
	35	91	77

4.6. Study Endpoints

Safety:

Safety measures to be assessed throughout the study from signing of the informed consent form (ICF).

- Adverse events and concomitant medications used
- Vital signs (temperature and the collection site [i.e., oral, rectal, temporal, or tympanic], heart rate, blood pressure, and respiratory rate)
- Physical examination including all major organ systems
- Injection site examination

Safety assessments conducted after completion of the study by all participants include serum assay results for anti-IL-10 antibody.

Efficacy:

The primary efficacy outcome measure is whether the subject achieves at least a 30% improvement from baseline in WOMAC Pain Score (obtained from the KOOS questionnaire) at Day 180, as recorded on a 0 to 20 point scale with 20 being the maximum (worst) score.

Secondary efficacy outcome measures are:

- Change from baseline in the WOMAC Function Score (from KOOS questionnaire) at Day 180
- Change from baseline in the WOMAC Pain Score (from KOOS questionnaire) at Day 180 (this endpoint was inadvertently omitted from the protocol)
- Change from baseline in patient response at Day 180, using a scale of 1 – 5, to the OA question: "Considering all the ways the OA in your knee affects you, how are you doing today?", with 1 being very good (asymptomatic and no limitation of normal activities) to 5, very poor (very severe, intolerable symptoms and inability to carry out normal activities)

Additional endpoints include:

- [redacted]
- Total and individual parameter scores for the Brief Pain Inventory of Intensity and Interference (BPI) (reported values and change from baseline)
- The Knee Outcome and Osteoarthritis Scores (KOOS) (reported values and change from baseline)
- [redacted]
- [redacted]

[redacted]

5. DATA LISTINGS

5.1. Sources of Data

Data sets containing raw data will be provided by the Data Management group, extracted from the clinical study database of data entered into a Case Report Form. The clinical study database will contain the data from Screening through to the Follow-up Safety Visit of the study.

Data sets (SDTM and ADaM) will be generated from the clinical study database of data entered into a Case Report Form, along with the following data received from other sources:

- Randomisation, as provided by the responsible [redacted] Biostatistician [redacted]
- Anti-IL-10 antibodies: to be provided by [redacted] to the unblinded [redacted] Biostatistician

All data entered into the CRF or provided from other sources as described above will be presented (explicitly or implicitly) in data listings or figures as described in Appendix 1.

5.2. Randomization and Subject Identification Code

After informed consent has been obtained, participants will be screened for study eligibility before enrollment. A total of 270 subjects will be randomized to the following 6 treatment sequences in a 1:1:1:1:1:1 ratio:

- XT-150 (150 µg) in Stage A, followed by XT-150 (150 µg) in Stage B
- XT-150 (150 µg) in Stage A, followed by XT-150 (450 µg) in Stage B
- XT-150 (450 µg) in Stage A, followed by XT-150 (150 µg) in Stage B
- XT-150 (450 µg) in Stage A, followed by XT-150 (450 µg) in Stage B
- Placebo in Stage A, followed by XT-150 (150 µg) in Stage B
- Placebo in Stage A, followed by XT-150 (450 µg) in Stage B

The randomization schedule will be prepared by unblinded statisticians [redacted] using SAS version 9.4 (or higher) statistical software package and will be maintained under controlled access. A copy of the randomization schedule will be provided to unblinded site pharmacy staff for the purposes of dispensing the study drugs.

Subjects will be assigned a randomisation number with the format 'Rnnn', where 'nnn' is the sequential subject id, from 001 to 270.

In the listings, subjects will be identified by randomisation number, with dose of XT-150 (treatment) as a secondary grouping identifier. Screening number will be shown only on the listing of informed consent data.

5.3. Maintaining the Study Blind

In this double-blind study, all personnel involved, i.e., physicians, site staff, and participants will remain blinded at all times, except in an emergency where knowledge of the randomisation code is required to provide appropriate treatment for an adverse event. The Investigator has the sole responsibility for determining if unblinding of a subject's treatment assignment is warranted for emergency reasons.

Preparation of the randomisation schedule in [redacted] was performed by staff members [redacted] who have no responsibility for monitoring and data management of this study.

The listings for inclusion in an appendix to the report will be prepared after the database has been locked and the study has been unblinded.

5.4. Assessment Time Point Identifiers

In the individual data listings, scheduled assessment time points will be identified as follows:

- Scheduled safety assessments will be identified by stage, study day, and also nominal study time point, where relevant.
- Assessments at Screening and End of Study may be identified as such for both study day and time point, as appropriate.

5.5. Data Derived by Calculation

The following data fields for inclusion in the data listings will be derived by calculation as per the Clinical Data Interchange Standards Consortium (CDISC) standard:

Adverse Events:

- Treatment for adverse event, assigned as
  Non-Treatment-Emergent if
  (Onset Date) < (Date of Stage A Dose)

  Treatment for Stage A if
  (Onset Date) >= (Date of Stage A Dose) &
  (Onset Date) < (Date of Stage B Dose);

  Treatment for Stage B if
  (Onset Date) >= (Date of Stage B Dose)

- Adverse event time since Stage A dose (in days), calculated as
  (Onset Date) – (Date of Stage A Dose)
- Adverse event time since Stage B dose (in days), calculated as
  (Onset Date) – (Date of Stage B Dose) if (Onset Date) >= (Date of Stage B Dose)
- Adverse event time since Stage B dose is Not Applicable if
  if (Onset Date) < (Date of Stage B Dose)
- Adverse event duration (in days), calculated as
  (Resolution Date– Onset Date)

Safety Assessments Change from Baseline:

- Baseline for vital signs and clinical laboratory parameters will be the most recent assessment prior to first dose [which may be an unscheduled assessment, a repeat of a scheduled Baseline assessment, or the screening assessment if separate Baseline assessments are not performed]

Clinical Safety Laboratory Assessments:

- If a clinical laboratory result is above the upper limit of normal, then Out of Range flag will be set to "H"
- If a clinical laboratory result is below the lower limit of normal, then Out of Range flag will be set to "L"
- Clinical Laboratory parameters that are outside the reference range and medically assessed as clinically significant will be identified as "Abnormal CS"
- Clinical Laboratory parameters that are outside the reference range and not medically assessed as clinically significant will be identified as "Abnormal NCS"

Efficacy:

- Baseline for WOMAC Pain score, WOMAC Function Score, and Patient response to OA question will be the most recent assessment prior to first dose
- 30% improvement in WOMAC Pain Score (yes/no) is defined as a reduction of ≥ 30% in the WOMAC Pain Score where reduction is defined as the relative change from the baseline value. [redacted]

[redacted]

5.6. Handling of Missing Data

For adverse events, with unknown intensity (severity) or unknown relationship to study treatment, these will be imputed as follows:

- If the intensity (severity) of an adverse event is unknown/missing, the intensity will be imputed for the summary of adverse events as being "Grade 3: Severe".
- If the relationship to investigational product of an adverse event is unknown/missing, the relationship will be imputed for the summary of adverse events as being the least favourable relationship, i.e., "Related".

For adverse events, where either the onset time or resolution time is unknown, time since first dose and duration will be imputed as follows:

- If onset date is unknown, and it cannot be confirmed that onset was prior to the start of dose administration, then the AE will be classified as treatment-emergent, with unknown time since first dose.
- If either the date of onset or the date of resolution is unknown, then duration will be shown as unknown.

For WOMAC, the scores will be imputed as follows:

- WOMAC Pain score will be calculated as the average of the corresponding non-missing sub items, multiplied by 5. In cases where 2 or more questions are missing, the WOMAC Pain score will not be imputed and will be set to missing
- WOMAC Function score will be calculated as the average of the corresponding non-missing sub scores, multiplied by 17. In cases where 4 or more questions are missing, the WOMAC Pain score will not be imputed and will be set to missing

5.7. Use of Abbreviations

Any abbreviations used in data listings will be included in a key in a footnote, as appropriate.

6. ANALYSIS POPULATIONS

The Intent-to-Treat (ITT) Population will comprise all enrolled participants who received the treatment injection, analyzed according to randomized Stage A treatment. The ITT Population will be the primary analysis population for efficacy.

The Safety Population will comprise all participants who receive any amount of study drug, analyzed according to treatment actually received.

Participants who are assigned a randomization number but withdraw prior to dosing will not be included in the safety analysis set.

7. SUBJECT DISPOSITION/BACKGROUND DATA

7.1. Subject Disposition

Subject disposition and administration procedures recorded in the CRF are as follows:

- Written informed consent
- Randomization
- Administration of IP
- Study completion /discontinuation
- Protocol deviations
- Additional comments

Details of participation and inclusion in analysis populations will be listed by subject. Completion status will be summarized by dose level / treatment.

Details of administration of IP will be listed by subject. Exposure to IP will be summarized by dose level / treatment.

Protocol deviations will be listed by subject. A summary of protocol deviations will be prepared for inclusion in the clinical study report.

7.2. Baseline and Eligibility Assessments

Baseline and eligibility assessments are as follows:

- Demographic details
- History of osteoarthritis of the knee
- Other medical and surgical history
- Physical examination
- Focused analgesia selection test
- Serology (HIV/Hepatitis B and C)
- Pregnancy Tests

7.3. Baseline Data Analysis

Demographics will be listed individually by subject and summarized by dose level / treatment and overall.

Medical history and physical examination data at baseline and results of tests performed for eligibility will be listed by subject.

Baseline assessments of on-study measures will be listed and summarized along with post-dose assessments, as described below.

8. SAFETY DATA

8.1. Safety Assessments

Safety assessments are as follows:

- Adverse events and concomitant medications used:
  - Continuous monitoring throughout the study period
- Vital signs
  - (heart rate, systolic and diastolic blood pressure, respiratory rate, temperature and site collected)
- Clinical laboratory tests
  - Hematology: hemoglobin, hematocrit, total and differential leukocyte count, red blood cell count, platelet count
  - Serum chemistry: bilirubin (total and direct), alkaline phosphatase, aspartate aminotransferase (AST), alanine aminotransferase (ALT)
  - Coagulation: prothrombin time [PT] / international normalized ratio [INR] and either partial thromboplastin time [PTT] or activated partial thromboplastin time [aPTT]
- Physical examination including all major organ systems

8.2. Safety Data Presentation and Analysis

All clinical safety and tolerability data will be listed for each subject.

Safety data will be summarized using descriptive statistics. Where summaries include changes from baseline, the relevant baseline value will be determined as described in Section 5.5 above.

All adverse events will be listed using MedDRA terms, as coded per the Data Management Plan. The number and percentage of subjects with at least one treatment emergent AE and the number of treatment emergent AEs reported will be summarized by dose level or treatment, grouped according to system organ class and preferred term, using descriptive statistics. The order for system organ class will be according to the Internationally Agreed Order, as per the guidance document MedDRA® *Data Retrieval and Presentation: Points to Consider*, specific to the MedDRA version used for this study. Summaries of TEAEs will also be presented by severity (graded according to the NCI CTCAE version 5.0) and by relationship to investigational product. In these summaries, subjects will be counted at most once per MedDRA term, for the AE of highest severity or least favorable relationship. Summaries will also be presented for SAEs and for AEs leading to study withdrawal.

Clinical laboratory parameters for hematology, serum chemistry, and coagulation will be listed. Summaries of reported results, including changes from baseline and medical

assessments, will be presented by treatment group and study timepoint. Individual subject profiles will be presented for any laboratory parameters with at least one post-dose value outside the laboratory's reference ranges and deemed clinically significant.

Vital signs will be listed and summarized by dose level / treatment and study timepoint. A summary of changes from pre-dose baseline will also be presented.

Changes in physical examination findings over time will be listed by body system for each subject. It is noted that any untoward findings identified on physical examinations after the administration of the first dose of study medication will be captured as an adverse event if those findings meet the definition of an adverse event as defined in the protocol.

Injection site examinations will be listed and summarized by dose level / treatment and study timepoint.

Any detection of anti-IL-10 antibodies will be summarized by treatment and timepoint.

9. EFFICACY DATA

9.1. Efficacy Assessments

Efficacy assessments will be recorded throughout the study, during Stage A and Stage B at each of the following study visits:

Stage A: nominally Days 7, 30, 60, 90, 120, 150, and 180.

Stage B: nominally Days 210, 240, 270, 300, 330 and 360

The primary efficacy outcome measure is whether the subject achieves at least a 30% improvement from baseline in WOMAC Pain Score (obtained from the KOOS questionnaire) at Day 180, as recorded on a 0 to 20 point scale with 20 being the maximum (worst) score.

Secondary efficacy outcome measures are:

- Change from baseline in the WOMAC Function Score (from KOOS questionnaire) at Day 180
- Change from baseline in the WOMAC Pain Score (from KOOS questionnaire) at Day 180 (this endpoint was inadvertently omitted from the protocol)
- Change from baseline in patient response at Day 180, using a scale of 1 – 5, to the OA question: "Considering all the ways the OA in your knee affects you, how are you doing today?", with 1 being very good (asymptomatic and no limitation of normal activities) to 5, very poor (very severe, intolerable symptoms and inability to carry out normal activities)

Additional endpoints include:

[redacted]

- Total and individual parameter scores for the Brief Pain Inventory of Intensity and Interference (BPI) (reported values and change from baseline)
- The Knee Outcome and Osteoarthritis Scores (KOOS) (reported values and change from baseline)
- [redacted]

[redacted]

9.2. Efficacy Data Presentation

Efficacy assessments will be listed for each subject.

The WOMAC Pain Score, including changes from baseline will be summarized by study visit and dose level / treatment.

The number of subjects achieving at least a 30% improvement from baseline in WOMAC Pain Score will be summarized by study visit and dose level / treatment. [redacted]

Reported values and calculated changes from baseline in WOMAC Function Score, in total and individual BPI, and in KOOS will be summarized by study visit and overall.

For the patient response, the number of subjects in each of the response categories will be summarized by study visit and dose level / treatment. The change from baseline in patient response will be summarized by study visit and dose level / treatment.

[redacted]

9.3. Efficacy Analysis

All inferential analyses of efficacy will be based on the first 180 days after dose administration (Stage A) and on the analysis of time from Day 0 until patient request for a second injection.

For the statistical analysis of the primary efficacy endpoint, a logistic regression (response: 30% improvement at Day 180 (yes/no)) will be performed to compare 150 µg XT-150 with placebo and 450 µg XT-150 with placebo statistically. The model will include a fixed effect for treatment and the subject's baseline WOMAC pain score as a covariate. From this model, the 150 µg XT-150 to placebo and 450 µg XT-150 to placebo odds ratio of reduction and its two-sided 95% CI will be calculated. In this analysis, a subject who does not provide a response at Day 180 will be considered a non-responder (i.e., improvement will be set to "no").

In the event that one or more treatment arms has no responders, which would cause numerical difficulties in the logistic regression model, the statistical analysis will add 2 successes and 2 failures to each treatment group. This based on the recommendation in Agresti and Caffo, "Simple and Effective Confidence Intervals for Proportions and Differences of Proportions Result from Adding Two Successes and Two Failures," The American Statistician, Vol. 54, No. 4 (Nov., 2000), pp. 280-288.

[redacted]

The change from baseline in WOMAC Function Score will be analyzed using a repeated measures model including fixed effects for treatment (i.e., 150 µg XT-150, 450 µg XT-150, placebo), study day (i.e., Day 7, 30, 60, 90, 120, 150, 180) and treatment by study day interaction. In addition, the subject's baseline score will be included as a covariate. From the model, the treatment differences between 150 µg XT-150 and placebo, and 450 µg XT-150 and placebo will be estimated for the average effect of the first 180 days after dose administration (i.e., Stage A). Additionally, p-values and corresponding two-sided 95% confidence intervals will be provided.

The change from baseline in WOMAC Pain Score will be analyzed similarly.

Change from baseline in total and individual BPI, change from baseline in KOOS, and change from baseline in the patient response will be analyzed similarly.

[redacted]

9.4. Adjustment for Multiplicity

Two endpoints will be subject to multiplicity adjustment: The primary endpoint (achievement of a 30% reduction in WOMAC pain) [redacted] This will be accomplished by first testing for significance of the two active doses of XT-150 versus placebo with respect to the primary endpoint at the 0.025 level, one-sided. This addresses the multiplicity of doses with respect to the primary endpoint. Then a Dunnett test for the 50% endpoint will be performed at the alpha remaining after having done the first Dunnett test. Using only the remaining alpha for this additional endpoint continues to control the overall false positive rate at 0.025, one-sided. The numerical value of this remaining alpha depends on the primary (30% endpoint) Dunnett test as follows:

- If both doses are significantly better than placebo, then the remaining alpha is the full 0.025 level.
- If exactly one of the two active doses is significantly better than placebo, then the alpha remaining will be determined by the software performing the primary Dunnett test. (If the statistical analysis was of Gaussian group means, with N=90 per group, the remaining alpha would be 0.0115 (one-sided); the actual alpha may be different, because the model includes a covariate and the sample sizes may not be exactly 90 per group.)
- If neither active dose is significantly better than placebo, then no alpha remains and no inferential testing [redacted] will be performed. [redacted]

Achievement of a 30% improvement [redacted] in WOMAC Pain score at Day 180 will be analyzed using a logistic regression model as described in Section 9.3 to preserve a familywise error rate of 0.05 (two-sided),

10. CHANGES FROM THE ANALYSIS PLANNED IN THE PROTOCOL

[redacted]

Although the protocol called for step-down statistical testing using the Dunnett approach, the Dunnett testing will not use step-down testing, because it would exhaust the 'alpha' available for additional statistical testing. By not using the step-down approach, it is permissible to continue statistical testing provided that at least one of the treatment arms is statistically significantly different from the control group, with respect to the primary endpoint.

[redacted]

11. DATA LISTINGS AND SUMMARY TABLES

[redacted]

11.2. Listings and Tables for Clinical Study Report

The data listings and summary tables, and associated figures, planned to be generated from the study data are listed in Appendix 1 and Appendix 2, respectively. The numbering and titles may vary in the final presentation, depending on the amount of data to be presented.

Listings, tables and figures will be provided to the Sponsor in a form suitable for inclusion as appendices to the Clinical Study Report, in RTF or PDF format. The layout will be landscape A4 size, with a margin of 1 inch. The default font for listings and tables will be Courier New 8 pt.

In the summary tables other than those included in the main text of the report, special characters and formatting will not be used. For example, units will be shown as ug (µg) and m^2 (m^2).

Data in listings will be ordered by subject, then date/time.

11.3. Data Transfer to the Study Sponsor

At the conclusion of the study, [redacted] will provide the Sponsor with an electronic copy of the listings, tables and figures, along with analysis data sets in SAS transport format and XPT format. This may be provided by email or by upload, as determined by the Sponsor.

Details of the statistical analyses will be retained in the [redacted] study file for reference.

12. GENERAL CONSIDERATIONS FOR DATA MANAGEMENT AND ANALYSIS

12.1. Analysis Packages

SAS 9.4 or higher (SAS Institute Inc., Cary, NC, USA) will be used for generating data listings and summary tables and associated figures, and for performing statistical analysis.

12.2. Electronic Data Management

All listings, tables and figures for inclusion in the appendices of the clinical study report will be generated using SAS programs.

A copy of final listings, tables and figures will be retained in the [redacted] internal study file. A tracking log will be maintained detailing the date and initials of the staff member responsible for generating each listing, table or figure, and the staff member conducting the quality control review. Any erroneous results identified during subsequent checking process will be corrected by update of the SAS program, and the document recreated and then re-checked. To ensure an accurate data trail is maintained, corrections to electronic documents will be included on the tracking log, detailing the correction along with the date of the new version as well as the initials of the staff member responsible for the change and the verifying staff member. If statistical analysis is repeated with a modified data set, the new data produced will be saved as a new file.

12.3. Archiving

At the conclusion of the study, the final listings, tables and figures will be archived along with the CDISC data sets and QC documentation. Copies will be retained in the [redacted] study file for a minimum of 15 years, and in the [redacted] secure archives in accordance with [redacted] standard operating procedures.

Appendix 1: Planned Data Listings

The following listings and figures of individual subject data are planned to be created, however the numbering and titles may vary, depending on the amount of data to be presented. Data presentation in listings may vary depending on CDISC datasets.

[redacted]

[redacted]

Appendix 2: Planned Summary Tables

The following summary tables are planned to be created, however the numbering and titles may vary, depending on the amount of data to be presented:

[redacted]

[redacted]

[redacted]

Statistical Analysis Plan Addendum

Protocol Title	A Double-Blind, Placebo-Controlled Assessment of the Tolerability and Efficacy of XT-150 for the Treatment of Moderate to Severe Pain Due to Osteoarthritis of the Knee
Protocol Number:	XT-150-2-0204
Protocol Version:	v 2.0
SAP Version:	Final 2.0, 11 Oct 2021
SAP Addendum Version:	15 Nov 2022
Sponsor:	Xalud Therapeutics, Inc. 2120 University Avenue, Suite 532 Berkeley CA 94704
Author:	[redacted]

Signatures / Approvals

Introduction

This Statistical Analysis Plan (SAP) Addendum describes the planned analyses to be done in addition to the primary analyses defined in the primary study SAP. The analyses planned in this SAP addendum were defined after the study database was locked and unblinded. Therefore, these analyses are considered post hoc analyses. However, these analyses can provide important supportive evidence of efficacy for the purposes of planning future studies and for understanding the effectiveness of XT-150 in subjects with moderate to severe pain due to osteoarthritis of the knee.

The focus of the analyses planned in this SAP Addendum is to characterize the efficacy during Stage B of the study (after Day 180 to Day 360) and to assess the effects of the second injection of XT-150.

Due to the exploratory nature of these analyses, no adjustments for multiplicity will be made.

SAS v9.4 or higher (SAS Institute Incl, Cary, NC, USA) will be used for all programming and analyses as well as to produce the summary tables and figures.

Analysis Populations

The following analysis populations will be used for the post hoc efficacy analyses:

> **Intent-to-Treat Population (ITT)**: The Intent-to-Treat population will consist of all randomized subjects who received at least one treatment injection. ITT population is analyzed based on the treatment to which subjects are randomized.
>
> **Modified Intent-to-Treat Population (mITT)**: The Modified Intent-to-Treat population will consist of ITT population who received at least one treatment injection with a baseline WOMAC Pain score of 9-20 (>8). This is the target population of interest who may benefit most from the treatment.

Analysis Sets

We will further separate the analyses into analyses for Stage A and Stage B:

- Stage A (up to and including Day 180) – data on first injection. This is the primary analysis for efficacy.
- Stage B (>Day 180 to Day 360) – data on second injection. The second injection could be taken at any time the patient decided, at Day 180 through Day 330.

Treatment Groups

The following descriptions define the treatment groups (or regimens) that will be used in the analyses:

- Stage A (up to and including Day 180): Subjects were randomized to placebo, 150 µg XT-150 or 450 µg XT-150 during Stage A. For the Stage A analyses each dose of XT-150 will be compared to placebo at Day 180:
    - 150 µg XT-150 vs Placebo
    - 450 µg XT-150 vs Placebo.
- Stage B (> Day 180 to Day 360): Subjects were randomized to receive either 150 µg XT-150 or 450 µg XT-150 during Stage B. The following treatment regimens will be compared for the Stage B analyses:

    - 1 Dose 450 vs 2 Dose 450
        - Placebo to 450 vs
        - 450 to 450
    - All Dose Regimens
        - Placebo to 150
        - 150 to 150
        - 450 to 150
        - Placebo to 450
        - 150 to 450
        - 450 to 450

Baseline

Baseline will be defined as the Day 0 value for both Stage A and Stage B analyses. [redacted]

[redacted]

Planned Analyses

[redacted]

[redacted]

4. Efficacy Analyses

The following efficacy endpoints will be analyzed as described below:

- WOMAC Pain Score Change from Baseline at Day 360
- WOMAC Function Score Change from Baseline at Day 360

[redacted]

- BPI Interference Change from Baseline at Day 360

[redacted]

WOMAC Pain Score Change from Baseline

WOMAC Pain Score Change from Baseline will be summarized with summary statistics by visit and treatment group. Analyses will be performed for both the ITT population and mITT population. [redacted]

A Mixed Model analysis will be conducted and mixed model statistics (LSMean, StErr, LSMeans Diff and 95% CI of LSMeans diff) – generated for Day 180 and Day 360 (covariates include site, baseline WOMAC Pain, K/L and sex). The effect size at Day 180 and Day 360 using Cohen's d will be calculated from the observed mean and standard deviations and presented.

A summary line plot of WOMAC Pain Mean Change from Baseline by visit and treatment group (with Standard Error at each visit) will be generated.

The following additional exploratory analyses will be performed for the WOMAC Pain Change from Baseline using the same MMRM model to compare treatment groups:

[redacted]

- Stage B (Day 360) analyses comparing treatment groups placebo/150 and placebo/450 to assess the difference between these groups and the assumption of combining them for the placebo/active treatment group.

[redacted]

WOMAC Function Score Change from Baseline

WOMAC Function Score Change from Baseline will be summarized with summary statistics by visit and treatment group.

The same Mixed Model analysis used for WOMAC Pain Score will be used, presenting the mixed model statistics (LSMean, StErr, LSMeans Diff and 95% CI of LSMeans diff) – for Day 180 and Day 360 (covariates include site, baseline WOMAC Function, K/L and sex). The effect size at Day 180 and Day 360 using Cohen's d will be calculated and presented.

The same additional exploratory and subgroup analyses to be done for the WOMAC Pain Score change from baseline will be performed for this endpoint.

[redacted]

[redacted]

BPI Interference Change from Baseline

BPI change from baseline will be summarized with summary statistics by visit and treatment group.

The same analyses to be performed for the WOMAC Pain Score change from baseline will be done for this endpoint.

[redacted]