XP-8121-120 Statistical Analysis Plan (SAP) Cover Page

Study Title:	A Phase 2, Multicenter, Non-Randomized, Open-Label, Single Arm, Self-controlled Study of XP-8121 For the Treatment of Adult subjects with Hypothyroidism
Most Recent Document Version Date:	04 April 2024
NCT No.	NCT05823012
IND No.	153502

CONFIDENTIAL

Xeris Pharmaceuticals, Inc.

Protocol Number: XP-8121-120

A Phase 2, Multicenter, Non-Randomized, Open-Label, Single Arm, Self-controlled Study of XP-8121 For the Treatment of Adult Subjects with Hypothyroidism

Statistical Analysis Plan (SAP)

Final Version 2.0

04 April 2024

SAP APPROVAL SIGNATURE PAGE

Protocol Number:	XP-8121-120
Protocol Version and Date:	Version 3.0 – 07 May 2023
Protocol Title:	A Phase 2, Multicenter, Non-Randomized, Open-Label, Single Arm, Self-controlled Study of XP-8121 For the Treatment of Adult Subjects with Hypothyroidism
SAP Version and Date:	Version 2.0 – 04 April 2024

☐ Original Statistical Analysis Plan ☒ Amended Statistical Analysis Plan

Signature below indicates the SAP has been authored by the following:

[redacted] Manager, Biostatistics Atorus Research Inc.	[redacted]

Signatures below indicate the SAP has been reviewed and approved by the following:

[redacted] Executive Director, Biostatistics Atorus Research Inc.	[redacted]
[redacted] Principal Statistical Programmer Atorus Research Inc.	
[redacted] Director, Biostatistics and Statistical Programming Xeris Pharmaceuticals, Inc.	
Valentina Conoscenti, MD Medical Director Xeris Pharmaceuticals, Inc.	

VERSION HISTORY OF FULLY EXECUTED (FINAL) PLANS

Version Number	Date	Version Author	Comments
1.0	30AUG2023	[redacted]	Initial version
2.0	04APR2024		Section 6.5: Clarified the prior medication definition to only include medications which were discontinued prior to treatment. Section 6.7: Wording clarification, no change to methods Sections 6.9.1 and 9.8.1: Corrected the XP-8121 dose volume to be µL rather than mL. Section 6.9.5: Described handling of triplicate values for clarification. Adjusted wording for clinically significant vital signs to be per EDC not investigator's discretion. Section 6.9.5 & 9.8.5: Removed BMI from change from baseline summary since it is only collected at baseline. Section 8.3: Added requirement that no date imputation will be done for medications marked as "Ongoing" Section 9.5: Removed statement of primary hypothesis. Added figure relating dose conversion factor to prior oral levothyroxine dose. Clarified difference between the general dose conversion factor and the final dose conversion factor. Added scatterplot of dose conversion factor over time. Added requirement to perform sensitivity analyses only if not all subjects have primary hypothyroidism Section 9.6: Scatterplot updated to be box plot Section 9.8.1: Scatterplot removed

TABLE OF CONTENTS

1. LIST OF ABBREVIATIONS

Abbreviation or Term	Definition
ATC	Anatomical Therapeutic Chemical
AUC	Area under the concentration vs. time curve
BMI	Body mass index
C	Celsius
CI	Confidence Interval
cm	Centimeter
C_{min}	Minimum observed concentration
C_{max}	Maximum observed concentration
CS	Clinically Significant
CRU	Clinical Research Unit
CTMS	Clinical Trial Management System
C_{ss}	Mean concentration at steady state
C_{trough}	Last concentration
%CV	Coefficient of variation
DBP	Diastolic blood pressure
ECG	Electrocardiogram
eCRF	Electronic Case Report Form
ET	Early Termination
FDA	United States Food and Drug Administration
fT4	Free Thyroxine
ms	milliseconds
MedDRA	Medical Dictionary for Regulatory Activities
PK	Pharmacokinetic
PO	Oral
PT	Preferred Term
QTcB	QT interval using Bazett's formula
QTcF	QT interval using Fridericia's formula
SBP	Systolic blood pressure
SC	Subcutaneous
SD	Standard Deviation
SOC	System Organ Class
T3	Triiodothyronine
T4	Thyroxine
TEAE	Treatment-Emergent Adverse Event
t_{max}	Time to occurrence of C_{max}
TSH	Thyroid Stimulating Hormone
TSQM-9	Treatment Satisfaction Questionnaire for Medication
WHO	World Health Organization
Xeris	Xeris Pharmaceuticals, Inc.

2. INTRODUCTION

Xeris Pharmaceuticals, Inc. (Xeris) has developed a novel formulation of levothyroxine, XP-8121 (levothyroxine sodium for subcutaneous [SC] injection) to mitigate challenges associated with oral (PO) formulations of levothyroxine, which can include interference with absorption (e.g., simultaneous ingestion with calcium, iron, certain foods, other drugs, and conditions that impair gastric acidification or intestinal absorption, such as celiac disease). For this reason, subjects are usually instructed to ingest oral levothyroxine on an empty stomach 30 to 60 minutes before food intake. Further, a once weekly dosing schedule, as provided by XP-8121, could improve adherence that is adversely affected by unintentionally missed daily doses. Finally, SC administration of XP-8121 could alleviate burdens associated with swallowing difficulties or other barriers to taking daily medication.

2.1 Hypothyroidism

The thyroid gland is responsible for the synthesis, storage, and release of the thyroid hormones, thyroxine (T4) and triiodothyronine (T3) [Colucci et al, 2013]. These hormones are nuclear receptor-activating and regulate critical metabolic processes that are vital for normal growth, development, and health throughout life.

Deficiency of circulating T4 or T3 is known as hypothyroidism. Most commonly in developed countries, the primary cause of hypothyroidism is autoimmune destruction of the glands responsible for thyroid hormone synthesis, also known as Hashimoto's disease or autoimmune thyroiditis. Other causes of primary hypothyroidism (thyroid hormone production/release deficits) include other forms of thyroid inflammation and iatrogenic causes (surgery and radiation sequelae and certain medications, iodine deficiency, et al). Uncommonly, thyroid synthesis is impaired at birth (congenital). Hypothyroidism can also result from problems with Thyroid Stimulating Hormone (TSH) production/secretion or action (secondary hypothyroidism, most commonly because of pituitary tumor or trauma) or from hypothalamic dysregulation of pituitary function (tertiary hypothyroidism). Rare syndromes of resistance to the actions of thyroid hormone can also resemble more typical hypothyroidism [Laurent et al, 2018; Colucci et al, 2013; Winther et al, 2016].

Regardless of etiology, inadequate thyroid hormone that manifests clinically should be treated, i.e., thyroid hormone replacement. Treatment is often chronic, most commonly lifelong. The goal of replacement therapy is restoration of the euthyroid state, with reversal of the clinical manifestations of hypothyroidism. In most developed countries, replacement most commonly consists of levothyroxine monotherapy and in most guidelines, TSH is currently recommended as the sole biomarker of replacement adequacy.

2.2 Study Design and Dose Range Rationale

The current study was designed, incorporating United States Food and Drug Administration (FDA) guidance, to provide information on weekly use of XP-8121 in stably treated subjects with hypothyroidism. XP-8121 is a ready-to-use, liquid formulation of levothyroxine sodium intended for SC administration once-weekly. The goals of the study are to determine safe starting doses, safe and feasible titration schedules, and the effective dose range of XP-8121, wherein effectiveness is determined by maintenance of normalized TSH. The oral to injected dose-conversion target was considered the primary objective and serves as the basis for hypothesis testing and sample size determination. Pharmacokinetics will also be characterized in a subset of subjects to determine any

differences from modeled chronic pharmacokinetic (PK) that used data from healthy volunteers. The findings from this study will inform a future clinical efficacy and safety demonstration.

This is a non-randomized, open-label, single arm, self-controlled study (i.e., subjects serve as their own controls based on baseline assessment of thyroid function). The design is appropriate because replacement of deficient T4 with levothyroxine is necessary and sufficient to treat hypothyroidism. A comparator is not necessary for an initial assessment of the ability of SC administered levothyroxine to maintain normalized TSH in stably treated subjects, as efficacy is not being established by this study. Blinding is not necessary since all subjects receive active therapy and there is no comparator. Furthermore, the primary outcome measure, TSH, is objective, and compliance (adherence) with therapy will be assessed.

The XP-8121 initiation dose will correspond to 50% of the target weekly dose of XP-8121 SC, which is defined, per study hypothesis, as the baseline daily oral levothyroxine dose (in µg) multiplied by 4, rounded up to the nearest 50 µg. The lowest dose to be studied is 100 µg. The highest dose to be studied is 1500 µg weekly, selected based on the known safety and toxicity profile of levothyroxine, published use of oral levothyroxine as a weekly dose regimen, and the observed and expected PK profile of XP-8121. This dose is not expected to result in overdosage using the proposed biweekly titration-to-target scheme with careful safety monitoring as described below.

3. STUDY OBJECTIVES

3.1 Primary Objectives

The primary objectives of this study are:

- To assess the safety and tolerability of XP-8121 after once-weekly SC injections in adult subjects with hypothyroidism
- To determine a target dose conversion factor from stably dosed oral levothyroxine to XP-8121 in subjects with hypothyroidism

3.2 Secondary Objectives

Secondary objectives of this study are:

- To determine a dose range of XP-8121 capable of maintaining normalized TSH following a switch from stable replacement therapy with oral levothyroxine
- To evaluate a safe initiation dose of XP-8121 based on the prior stable dose of oral levothyroxine
- To evaluate a safe titration scheme for dosing XP-8121 to TSH target

3.3 Exploratory Objectives

Exploratory objectives of this study are:

- To characterize the dynamics of thyroid hormones and TSH in blood during and after titration (i.e., Titration and Maintenance Periods) with XP-8121
- To characterize the PK of plasma thyroid hormones following multiple once-weekly XP-8121 SC injections in adult subjects with hypothyroidism

- To compare subject treatment satisfaction between daily oral levothyroxine and once-weekly XP-8121 SC injections

4. STUDY DESIGN

4.1 General Design

Only a synopsis of the study is presented herein, full details can be found in the protocol. This is a non-randomized, open-label, single arm, self-controlled study of XP-8121 to determine a target dose conversion factor from stably dosed oral levothyroxine to XP-8121 in subjects with hypothyroidism and to assess the safety and tolerability of XP-8121 after once-weekly SC injections.

This study includes the following periods: Screening, Titration Period, and Maintenance Period. The study will conclude with an End of Maintenance or Early Termination (ET) Visit. The XP-8121 initiation dose will correspond to 50% of the target weekly dose of XP-8121 SC (in µg), which is calculated as the current daily oral levothyroxine dose multiplied by 4, rounded up to the nearest 50 µg.

During the Titration Period, subjects will come into the clinic once a week for at least 2 and up to 8 weeks, as applicable, for dosing administration. Based on free thyroxine (fT4) levels, XP-8121 will be titrated biweekly (every other week) by intervals of 25% of the target dose. TSH and T4 (free and total) will be checked weekly. A safety check will be completed 72 hr. (± 1 day) after the initial dose and each escalation during the Titration Period. If fT4 exceeds the upper limit of normal during titration, the XP-8121 dosage will be lowered and held at the previously tolerated dose.

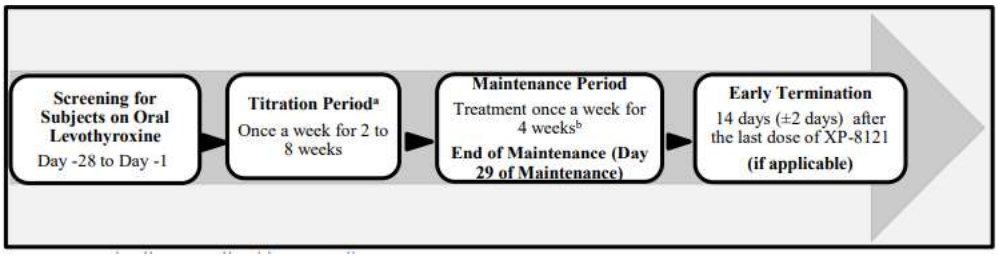
Once subjects have completed the Titration Period, they will continue at the same dose of XP-8121 in the Maintenance Period for 4 weeks, returning to the Clinical Research Unit (CRU) on a weekly basis for dosing. An End of Maintenance Visit will be performed at Day 29 (1 week after the last XP-8121 dose). Subjects not achieving an adequate fT4 at ≤125% of the target dose or the maximum weekly dose of 1500 µg may also continue into the Maintenance Period at the discretion of the Investigator and after conferring with the Medical Monitor. Titration details are outlined in Appendix A of the study protocol.

Subjects completing the Maintenance Period will have been maintained on a stable dose of XP-8121 for a total of approximately 6 weeks: subjects are required to be on a stable dose of XP-8121 for the last 2 weeks of the Titration Period, enter the Maintenance Period on the same dose, and continue at the same dose during the 4-week duration of the Maintenance Period. During Week 4 of the Maintenance Period, a subgroup of subjects (~12) who have been on a stable dose of Synthroid for at least 3 months prior to Screening and who have maintained adequate trough fT4 and total T4 levels in the Maintenance Period will participate in a PK sub-study. Subjects will be housed in the CRU for 9 days and 8 nights to undergo PK sample collections at specified timepoints as outlined in the study protocol.

After completing the Maintenance Period, subjects will complete the study with an End of Maintenance Visit and will subsequently restart their previous daily levothyroxine PO dose. AEs (Adverse Events) that are ongoing at the End of Maintenance Visit will be recorded on the case report form (CRF) and followed until resolution or maximally up to 14 days after the End of Maintenance Visit or ET Visit. Subjects that withdraw or are discontinued will have an ET Visit 14 days (±2 days) after the last dose of XP-8121. Subjects who withdraw or are discontinued, for any reason, will not be replaced.

Approximately 44 subjects who meet all inclusion criteria and exclusion criteria will be enrolled. Approximately 12 subjects will be enrolled in the PK sub-study.

Figure 1: Study Design Schema



a: Titration details are outlined in Appendix A of the study protocol.
b: A subgroup of subjects (~12), who have been on Synthroid for at least 3 months prior to Screening and who have maintained adequate trough T4 (free and total) levels in the Maintenance Period will participate in a PK sub-study. During Week 4 of the Maintenance Period, subjects will be housed in the CRU for 9 days and 8 nights to undergo PK sample collections at specified timelines as outlined in the study protocol.

4.2 Randomization and Treatment Assignments

All enrolled subjects will receive XP-8121 with titrated dosage; randomization is not applicable.

4.3 Blinding/Unblinding

This study is open-label: subjects, investigators, and study staff are aware of treatment identity.

4.4 Sample Size Considerations

Assuming a between-subject coefficient of variation of 0.5 and a within-subject, between-measurements correlation of 0.5, a sample size of 38 subjects will have 80% power to reject the null hypothesis that the dose conversion factor is less than 3.2X or more than 5X, in favor of the alternative hypothesis that the dose conversion factor is between 3.2X and 5X (in another words, within 80% and 125% of the proposed 4X conversion factor derived from Population PK modeling). For the exploratory endpoint of the treatment satisfaction questionnaire for medication (TSQM-9), assuming a SD of differences of 20, a sample size of 34 subjects will have 80% power to detect a difference of 10 in means (domain scores range from 0 to 100) with a 5% two-sided significance level (α=0.05). Taking into consideration non-completers and non-normalizers of TSH at study end, a total of 44 subjects will be enrolled.

5. CHANGES IN THE ANALYSES SPECIFIED IN THE PROTOCOL

Dose-normalized PK parameters have been removed as endpoints.

6. STUDY ENDPOINTS

6.1 Schedule of Assessments

The assessments planned to be collected at each visit are presented in the study protocol.

6.2 Timepoint Definitions

6.2.1 Relative Day

The date on which the first dose of study drug was administered will be considered as relative day 1, while the day before that will be relative day -1 (there is no relative day 0). Relative days will be calculated using the following formula only if the date of the assessment is fully known (i.e., the relative day will be missing if the assessment date is a partial date):

Relative Day = Date of Assessment – Date of First Dose of Study Drug (+1 if Date of Assessment ≥ Date of First Dose of Study Drug).

6.2.2 Visit and Timepoint Windows

Nominal, scheduled visits and timepoints will be used for analysis. Unscheduled visits will be considered for clinically significant or abnormal safety summaries. Scheduled and unscheduled visits will be listed.

6.2.3 Baseline

Two baselines are defined for this study as the last evaluable measurement taken prior to the date/time of the first dose of study drug administration in the Titration Period (prior to any dose of study treatment) and prior to the date/time of the first dose of study drug administration in the Maintenance Period. For electrocardiograms, if a cardiologist interpretation is available on this record, this value will be used as baseline for the appropriate period.

Vital signs (blood pressure, heart rate) baseline will use the average of the triplicate assessments in the last timepoint occurring prior to the first dose of study drug treatment administration for each treatment period.

All change from baseline analyses will use the Titration Period baseline across all records in addition to the Maintenance Period baseline on records only from the Maintenance Period, as applicable.

6.3 Protocol Deviations

Protocol deviations will be documented in the clinical database for subject-specific protocol deviations and the ProSciento Clinical trial Management System (CTMS) for site- and study-specific protocol deviations. The protocol deviations will be reviewed according to the Protocol Deviation Plan. Before database lock; this review will include a confirmation of which deviations are considered to be major. The following representatives from Xeris will participate in the review of the deviations: Biometrics, Clinical Research Operations, Regulatory, Quality and Medical. The database will not be locked until the deviations are finalized.

6.4 Demographics and Baseline Characteristics

The following assessments will be collected at baseline:

- Demographic parameters
  - Age (years) – calculated as: *(Date of Informed Consent – Date of Birth)/365.25*
  - Sex
  - Race
  - Ethnicity

- Vital Signs
  - Weight (kilograms [kg])
  - Height (centimeters [cm])
  - Body Mass Index (BMI; kg/m^2) – calculated using the following formula: *Weight (kg) ÷ (Height (cm)/100)^2*
- Medical History
  - Coded using MedDRA version 26.0
- Disease History
  - Hypothyroidism History Type (e.g., Primary, Secondary, Tertiary)
  - Prior Oral Levothyroxine Brand and Dose

6.5 Prior and Concomitant Therapies

The World Health Organization (WHO) Drug Dictionary version from March 2023 will be used to code the recorded medications to Anatomical Therapeutic Chemical (ATC) level 2 and standardized (generic) medication name.

Prior medications include any recorded medication that was taken and discontinued before the date/time of first dose of study drug administration. Concomitant medications include any recorded medication that was taken on/after the date/time of first dose of study drug administration. See Section 8.3 below for details on how to identify prior vs. concomitant when the start/end dates for the medication usage are partial dates or missing.

6.6 Primary Endpoints

The primary endpoints are as follows:

- Safety endpoints collectively include incidence, frequency, severity, and relationship to study drug of Treatment-Emergent Adverse Events (TEAEs), including clinically significant vital sign measurements, safety laboratory test results, Electrocardiograms (ECGs), and physical examinations.
  - Described in Section 6.9.
- Tolerability assessments include Modified Draize Scale and Injection Site Discomfort questionnaires.
  - Described in Section 6.9.7.
- Geometric mean ratio of the weekly dose of XP-8121 SC (in µg) to the daily dose of oral levothyroxine (i.e., dose conversion factor) as recorded in the concomitant medications electronic case report form (eCRF), and where the XP-8121 dose corresponds to the administration immediately preceding collection of blood TSH used to determine normalization at the end of the Maintenance Period.
  - The final dose conversion factor will be assessed by comparing the ln-transformed final XP-8121 dose that maintains normalized TSH at the end of Maintenance Period to the ln-transformed daily oral levothyroxine dose the subject was taking before the study. The dosage for oral levothyroxine and XP-8121 should use micrograms as a unit. This primary endpoint will be assessed using the Completer population (as defined in Section 7) for subjects with primary hypothyroidism only (see Section 9.5 for sensitivity analyses on other populations/subsets).

6.7 Secondary Endpoints

The secondary endpoints are as follows (only subjects with primary hypothyroidism will be considered for the secondary endpoints):

Endpoint	Applicable to	Timespan	Derivation
Normalized Values	TSH	Throughout the Maintenance Period	Normalized TSH will be defined as values of TSH within the normal range throughout the Maintenance Period
Normalized Values	TSH	End of Maintenance Period	Normalized TSH will be defined as values of TSH within the normal range at the end of the Maintenance Period
Observed Values that are lower than, within, or higher than the normal range	Total thyroxine (T4) fT4 TSH	During the Titration and Maintenance Periods	Observed results at each visit will be compared to the normal range and subjects will be categorized as "Low", "Normal", or "High".

6.8 Exploratory Endpoints

Exploratory endpoints are as follows (only subjects with primary hypothyroidism will be considered for the non-PK exploratory endpoints):

Endpoint	Applicable to	Timespan	Derivation
Observed Values	T4 fT4 TSH	Titration Period Maintenance Period	N/A
Change from Baseline	T4 fT4 TSH	Titration Period Maintenance Period	Value at Timepoint – Value at applicable Baseline Maintenance Period records will also display change from Titration Period Baseline
Time of Maximum Concentration (t_{max}) (hours)	T4 Triiodo-thyronine (T3)	168-hr PK Period for Subjects in PK sub-study	Time in hours of the Maximum Value
Maximum Concentration (C_{max}) (ng/mL)	T4 T3	168-hr PK Period for Subjects in PK sub-study	Maximum concentration
Minimum Concentration (C_{min}) (ng/mL)	T4 T3	168-hr PK Period for Subjects in PK sub-study	Minimum concentration

Endpoint	Applicable to	Timespan	Derivation
Last concentration (C_{trough}) (ng/mL)	T4 T3	168-hr PK Period for Subjects in PK sub-study	Concentration at 168hrs
Mean concentration at steady state (C_{ss})	T4 T3	168-hr PK Period for Subjects in PK sub-study	Calculated as $AUC_{0\text{-}168hrs}$ / 168hrs
$AUC_{0\text{-}168hrs}$	T4 T3	168-hr PK Period for Subjects in PK sub-study	AUC using the observed result over a dosing interval using the "linear-up log-down" approach, which uses the linear trapezoidal method up to time t_{max} and the logarithmic trapezoidal method for all times after t_{max}. The linear trapezoidal method is defined as the sum across all adjacent pairs of timepoints: $\frac{1}{2}(t_{i+1} - t_i)(concentration_{i+1} + concentration_i)$ Where t_i and t_{i+1} are adjacent collection timepoints, and $concentration_i$ and $concentration_{i+1}$ are the respective concentration levels at those times. The logarithmic trapezoidal method is defined as the sum across all adjacent pairs of timepoints: $\frac{(t_{i+1} - t_i)(concentration_i - concentration_{i+1})}{\ln(concentration_i) - \ln(concentration_{i+1})}$ Where t_i, t_{i+1}, $concentration_i$, and $concentration_{i+1}$ are the same as described above. The time 0 value will use the nominal time of collection at time 0 while actual time elapsed will be used for timepoint >0.
Observed Values	Convenience Domain Score of the Treatment Satisfaction Questionnaire for Medication (TSQM-9)	Per collection timepoint	Convenience Score is the transformed total of items 4, 5, and 6 as follows: If all items are non-missing: ([{Sum of Item 4 to Item 6} – 3] divided by 18) * 100. If one item is missing: ([{Sum of non-missing items 4, 5, and 6} - 2] divided by 12) * 100. If two or more items are missing, then do not calculate.

Endpoint	Applicable to	Timespan	Derivation
Observed Values	Overall Satisfaction Domain Score of TSQM-9	Per collection timepoint	Overall Satisfaction is the total of Item 7, 8, and 9. If all items are non-missing: ([{Sum of Item 7 to Item 9} – 3] divided by 14) * 100. If item 7 or 8 is missing: ([{Sum of non-missing Item 7 to Item 9} – 2] divided by 10) * 100. If item 9 is missing: ([{Sum of item 7 and item 8} – 2] divided by 8) *100 If two or more items are missing, then do not calculate.
Observed Values	Effectiveness Domain Score of TSQM-9	Per collection timepoint	Effectiveness is the total of Items 1, 2, and 3. If all items are non-missing: ([{Sum of Item 1 to Item 3} – 3] divided by 18) * 100. If one item is missing: ([{Sum of non-missing Item 1 to Item 3} - 2] divided by 12) * 100. If two or more items are missing, then do not calculate.
Observed Values	Subject Preference Survey	Per collection timepoint	N/A

6.9 Safety Endpoints

6.9.1 Study Drug Exposure and Compliance

Exposure to study drug will be calculated separately for each treatment period and overall.

The duration of XP-8121 exposure will summarize the total time in days while taking XP-8121 in each period and overall. The duration of exposure will be calculated as: *(date of last dose of study drug – date of first dose of study drug +1)*.

The number of doses administered will categorize each subject by total number of XP-8121 doses received within each period and overall.

The total volume (µL) administered will be the sum of the volume (µL) of drug administered in each period and overall.

The total dose (µg) administered will be the sum of the doses (µg) of drug administered in each period and overall.

6.9.2 Adverse Events

AE verbatim terms recorded by the investigator will be coded to system organ class (SOC) and preferred term (PT) using the Medical Dictionary for Regulatory Activities (MedDRA) version 26.0.

A TEAE is defined as any adverse event which occurs on or after the first dose of study drug.

All TEAEs will be assessed by the investigator for severity and for causality. Events with missing severity will be summarized as Severe. Events with missing relationship to study treatment will be summarized as related.

TEAEs marked with a relationship to study treatment of definitely related, probably related, possibly related, or missing will be considered as related TEAEs, and TEAEs summarized by relationship will count each subject once for the maximum relationship within each SOC and PT.

TEAEs will be attributed to treatment periods by comparing the start date/time of the event to the date/time of first dose within each of the treatment periods.

6.9.3 Clinical Laboratory Assessments

Changes from baseline for all hematology and chemistry data will be calculated as: *the visit value minus the baseline value*. Maintenance Period records will also display changes from Titration Period Baseline.

Laboratory results will be classified as Abnormal if outside of the normal range in either direction, as Low if below the lower limit of normal, and as High if above the upper limit of normal.

The investigator will classify relevant laboratory results as clinically significant (CS) in the eCRF.

6.9.4 Electrocardiograms

Changes from baseline for all ECG parameters (heart rate [beats/min], QT interval [ms], PR interval [ms], QRS interval [ms], RR interval [ms], the corrected QT interval using Fridericia's formula [QTcF; ms], and the corrected QT interval using Bazett's formula [QTcB; ms]) will be calculated as: *the visit value minus the baseline value*. Maintenance Period records will also display changes from Titration Period Baseline.

Clinical significance of ECG interpretation will be determined by the investigator and overread by the cardiologist. While both results will be listed, only the overread by the cardiologist will be summarized in tables.

6.9.5 Vital Signs

Vital signs (blood pressure, heart rate) will use the average of the triplicate assessments collected at each timepoint. If the average is derived in the eCRF, then that average will be used for analysis. If a triplicate is collected but the average is not determined in the eCRF, then the average of available values will be derived and used for analysis.

The change from baseline will be calculated as: *the visit value minus the baseline value*, for the following vital sign parameters:

- Systolic blood pressure (SBP; mmHg)
- Diastolic blood pressure (DBP; mmHg)
- Heart rate (beats/minute)
- Respiratory rate (breaths/minute)

- Body temperature (Celsius [C])
- Weight (kg)

Maintenance Period records will also display changes from Titration Period Baseline.

CS values, as collected on the CRF, will be summarized and separately flagged in a listing.

6.9.6 Physical Examination

Physical examination results will be listed.

6.9.7 Local Tolerability

Local tolerability will be assessed using the Modified Draize Scale and Injection Site Discomfort questionnaires following protocol-specified timepoint following SC injections of XP-8121. The Modified Draize Scale assesses Erythema and Edema on a 0 to 4 point scale with 0 being None and 4 being Severe.

Injection site discomfort will be assessed using an 11-point Numeric Rating Scale (NRS) and an Injection Site Discomfort Description and Duration Questionnaire following injection of XP-8121. Injection Site Discomfort Intensity will be categorized into the following groups: Mild = 1 – 3, Moderate = 4 – 6, and Severe = 7 – 10.

Time until onset of discomfort (minutes) and Duration of discomfort (minutes) will be summarized as reported on the eCRF.

7. ANALYSIS SETS

Analysis Set	Definition	Used for Analyses
Safety Population	All subjects who received at least one dose of XP-8121.	Disposition, Demographics, Medical History, Disease History, Concomitant Medications, Deviations, Safety Analyses, Tolerability endpoints Secondary endpoints, Exploratory endpoints (except PK related endpoints)
Completer Population	All subjects who completed the Maintenance Period and exhibit normalized TSH throughout the Maintenance Period while on a stable dose for approximately 6 weeks. Subjects with a reduced dose at any point during the Maintenance Period will not be included.	Dose Conversion Factor analysis, Sensitivity analysis of Change from Baseline of Blood Thyroid Hormone Concentrations Secondary Endpoint
PK Population	All safety population subjects who participated in the PK sub-study, who do not have any major protocol deviations affecting the PK endpoints and for whom a sufficient number of samples is available to determine at least one PK parameter.	PK analyses

8. STATISTICAL METHODS

8.1 General Methodology

Unless otherwise specified, statistical analyses will be 2-tailed and assessed at the 5% significance level.

Tabulations will be produced for all measured data. For categorical variables, summary tabulations of the number and percentage of subjects within each category (with a category for missing data) of each measure will be presented. For continuous variables, the number of subjects (n), mean, standard deviation (SD), median, minimum, and maximum will be presented. Geometric means and coefficient of variation (%CV) may be presented for certain analyses.

Subject listings of all data from the case report forms (CRFs) and any derived variables will be presented.

8.2 Adjustments for Covariates

Not applicable.

8.3 Handling of Dropouts or Missing Data

Drop-out subjects will not be replaced.

In general, there will be no replacement of missing data points. All data recorded on the eCRF for enrolled subjects will be included in data listings that will accompany the clinical study report.

When considering AE and Concomitant Medications, partial start/end dates will be managed in a conservative fashion so that if there is any possibility that the available partial information could reasonably result in the AE being treatment-emergent or the medication being concomitant, then it will be imputed as such. Specifically:

- If the start day is missing but month and year are present:
    - If the known month/year are the same as the month/year of the first dose of treatment, then impute to the date of the first dose of treatment;
    - Otherwise, impute the day to 01.
- If the start day and month are both missing but the start year is present:
    - If the known year is the same as the year of the first dose of treatment, then impute to the date of the first dose of treatment;
    - Otherwise, impute the day and month to 01January.
- If the start date is completely missing, impute to the date of the first dose of treatment, unless the end date suggests it could have started prior to this, in which case impute to 01January of the same year as the end date.
- When imputing a start date, ensure that the new imputed date is on or prior to the end date of the AE or medication.
- If the end date is missing or partial, and the AE/medication is not ongoing, then impute the missing/partial end date as follows:
    - If the end day is missing but month and year are present:

- If the known month/year are the same as the month/year of the first dose of treatment, then impute to the date of the first dose of treatment;
- Otherwise, impute the day to the last day of the month.

- If the end day and month are both missing but the end year is present:
  - If the known year is the same as the year of the first dose of treatment, then impute to the date of the first dose of treatment;
  - Otherwise, impute the day and month to 31December.
- If the end date is completely missing, impute to the date of the first dose of treatment.
- When imputing an end date, ensure that the new imputed date is on or after the start date of the AE or medication.

- Other unexpected combinations of partial date information (e.g., known month but missing day and year) will be managed as needed on a case-by-case basis.

If samples are collected below the limit of quantification or using "<xx", then use the "xx" value divided by 2 for analysis.

The handling of missing data will be discussed throughout sections 6 and 9 of the SAP, as applicable.

8.4 Multi-Center Studies and Pooling of Centers

Not Applicable.

8.5 Multiple Comparisons/Multiplicity

Not Applicable.

8.6 Interim Analyses and Data Monitoring Committees

Not Applicable.

8.7 Examination of Subgroups

Additional subgroup analysis may be performed at the discretion of the sponsor. Details are described in Section 9.9.

9. STATISTICAL ANALYSIS

9.1 Subject Disposition and Analysis Sets

Analysis sets will be summarized by subject counts and percentages.

Subjects screened and screen failures (with reasons) and those who enroll, complete, or discontinue early (with reasons) will be summarized with subject counts and percentages. Subject disposition information per the eCRF will be listed. A by-subject listing of those who discontinue early will include dates of treatment start and end, informed consent, study end, and discontinuation reason.

9.2 Protocol Deviations

Subjects with at least one major protocol deviation will be summarized with counts and percentages within deviation category, sub-category, and deviation type.

9.3 Demographics and Baseline Characteristics

Demographic and other baseline characteristics will be summarized.

Descriptive statistics (n, mean, SD, median, minimum, and maximum) will be summarized for age (years), baseline height (cm), baseline weight (kg), and baseline BMI (kg/m^2). The number and percentage of subjects will be summarized for sex, race, ethnicity, hypothyroidism history type, and prior oral levothyroxine dose.

The number and percentage of subjects with at least one medical history event will be summarized within SOC and PT categories. Subjects will be counted only once within each SOC and PT.

9.4 Prior and Concomitant Therapies

The number and percentage of subjects with at least one prior medication will be summarized using WHO ATC Level 2 class and preferred names. Subjects will be counted only once within each ATC class and standardized (generic) medication name. Concomitant medications will be summarized likewise.

9.5 Analysis of Primary Endpoints

The final dose conversion factor will be assessed by comparing the ln-transformed final XP-8121 SC dose (in µg) that maintained normalized TSH throughout the Maintenance Period to the ln-transformed daily oral levothyroxine dose (in µg) each subject was taking before the study, using a linear mixed effects model with treatment (i.e., XP-8121 or oral levothyroxine) as fixed effects, and subject as random effect. A Variance Components (diagonal) covariance matrix will be used. If the model does not converge, then the compound symmetric covariance matrix will be used and defined in a footnote.

Geometric Least Square Means (LSM) by treatment, geometric mean ratios of XP-8121 SC dose vs daily PO levothyroxine dose (i.e. the final dose conversion factor), and the 90% CIs for the final dose conversion factor will be presented for the completer population with primary hypothyroidism.

Individual dose conversion factor, calculated as final XP-8121 SC dose (in µg) divided by prior oral levothyroxine dose (in µg), will also be listed. A line plot will display the final dose conversion factor versus the prior oral levothyroxine dose.

In addition, a scatterplot will be provided with relative day of dose on the x-axis and individual dose conversion factor on the y-axis. Data will be grouped into all subjects in Titration Period, subjects who completed the Maintenance Period with a stable dose and normalized TSH, and all other subjects in Maintenance Period.

Sensitivity analyses will be performed using the same analyses methods defined above but with other analysis populations, if applicable. No plots will be created for sensitivity analyses. Populations used in sensitivity analyses are as follows (the analysis for subjects with any hypothyroidism diagnosis will only be performed if not all subjects had primary hypothyroidism) :

- Subjects with primary hypothyroidism who maintained normal TSH throughout Maintenance Period (allowing for dosing changes during Maintenance Period)
- Subjects with primary hypothyroidism who maintained both normal TSH and fT4 throughout Maintenance Period (allowing for dosing changes during Maintenance Period)
- Subjects with any hypothyroidism diagnosis who maintained normal TSH throughout Maintenance Period (allowing for dosing changes during Maintenance Period)
- Subjects with any hypothyroidism diagnosis who maintained both normal TSH and fT4 throughout Maintenance Period (allowing for dosing changes during Maintenance Period)
- Completer population with primary hypothyroidism who also maintained normal fT4 throughout Maintenance Period
- Completer population with any hypothyroidism diagnosis
- Completer population with any hypothyroidism diagnosis who also maintained normal fT4 throughout Maintenance Period

Analyses of safety endpoints include incidence, frequency, severity, and relationship to study drug of TEAEs. CS vital sign measurements, safety laboratory test results, ECGs, and physical examination findings will be listed and qualitatively analyzed where applicable.

Analyses of the tolerability assessments Modified Draize Scale and Injection Site Discomfort questionnaires are described in Section 9.8.7.

9.6 Analysis of Secondary Endpoints

Secondary endpoints will be analyzed as follows (only subjects with primary hypothyroidism will be considered for the secondary endpoints):

Endpoint	Applicable to	Timespan	Analysis Method
Normalized Values	TSH	Throughout the Maintenance Period/ End of Maintenance Period	The number and percentage of subjects with normal values will be presented along with the number of subjects entering the Maintenance Period. Corresponding 95% CIs will be included using the Clopper-Pearson method. Values will be presented with the denominator comprising subjects entering the Maintenance Period and subjects in the safety population.

Endpoint	Applicable to	Timespan	Analysis Method
Observed Values that are lower than, within, or higher than the normal range	T4 fT4 TSH	During the Titration and Maintenance Periods)	The number and percentage of subjects with results will be summarized by three categories based on the normal range: "Low", "Normal", or "High". A boxplot will be included with the x-axis displaying the starting dose of XP-8121 and the y-axis showing fT4 following the first dose of XP-8121 on the y-axis.

9.7 Analysis of Exploratory Endpoints

Exploratory Endpoints will be summarized as follows (only subjects with primary hypothyroidism will be considered for the non-PK exploratory endpoints):

Endpoint	Applicable to	Timespan	Analysis Method
Observed Values	T4 fT4 TSH	Titration Period Maintenance Period	Summarized by each protocol-specified timepoint using descriptive statistics (n, mean, SD, median, minimum, and maximum). A line plot of the mean observed value will be provided over duration of the study. A spaghetti plot with each subject compiled will be provided over the duration of the study for fT4 and TSH. Repeat analysis using the Completer Population

Endpoint	Applicable to	Timespan	Analysis Method
Change from Baseline	T4 fT4 TSH	Titration Period Maintenance Period	Summarized by each protocol-specified timepoint using descriptive statistics (n, mean, SD, median, minimum, and maximum). A line plot of the mean change from baseline will be provided for each period. Repeat analysis using the Completer Population
Total Plasma Concentrations	T4 T3	168-hr PK Period for Subjects in PK sub-study	Summarized by collection timepoint and period using descriptive statistics (N, mean, SD, SD, %CV, median, minimum, maximum, geometric mean, and geometric %CV). A line plot with mean and SD of concentrations by timepoint will be provided along with a by subject plot of concentration by timepoint.
t_{max} (hours)	T4 T3	168-hr PK Period for Subjects in PK sub-study	Summarized using descriptive statistics (n, mean, SD, %CV, median, minimum, maximum, geometric mean, and geometric %CV).
C_{max} (ng/mL)	T4 T3	168-hr PK Period for Subjects in PK sub-study	Summarized using descriptive statistics (n, mean, SD, %CV, median, minimum, maximum, geometric mean, and geometric %CV).
$AUC_{0\text{-}168hrs}$	T4 T3	168-hr PK Period for Subjects in PK sub-study	Summarized using descriptive statistics (n, mean, SD, %CV, median, minimum, maximum, geometric mean, and geometric %CV).
Minimum Concentration (C_{min}) (ng/mL)	T4 T3	168-hr PK Period for Subjects in PK sub-study	Summarized using descriptive statistics (n, mean, SD, %CV, median, minimum, maximum, geometric mean, and geometric %CV).

Endpoint	Applicable to	Timespan	Analysis Method
Last concentration (C_{trough}) (ng/mL)	T4 T3	168-hr PK Period for Subjects in PK sub-study	Summarized using descriptive statistics (n, mean, SD, %CV, median, minimum, maximum, geometric mean, and geometric %CV).
Mean concentration at steady state (C_{ss})	T4 T3	168-hr PK Period for Subjects in PK sub-study	Summarized using descriptive statistics (n, mean, SD, %CV, median, minimum, maximum, geometric mean, and geometric %CV).
Observed Values	Convenience Domain Score of the TSQM-9	Per collection timepoint	Summarized by each protocol-specified timepoint using descriptive statistics (n, mean, SD, median, minimum, and maximum). The analysis of TSQM-9 will be hierarchically analyzed to ensure control of the family-wise type I error rate at the 0.05 level, with the Convenience Score first, Overall Satisfaction score second, and Effectiveness third. Hypothesis tests for the three domain scores are: H_0: No difference between oral levothyroxine and once-weekly XP-8121 SC treatments H_1: Difference between oral levothyroxine and once-weekly XP-8121 SC treatments. Each domain score will be compared between the two treatments (XP-8121 and Oral Levothyroxine) using a paired t-test at each timepoint. A line plot will be provided with the mean of Oral Levothyroxine as a reference line and the mean 95% Confidence Interval XP-8121 results at each collection point.

Endpoint	Applicable to	Timespan	Analysis Method
Observed Values	Overall Satisfaction Domain Score of TSQM-9	Per collection timepoint	Summarized by each protocol-specified timepoint using descriptive statistics (n, mean, SD, median, minimum, and maximum). A p-value from a paired t-test at each timepoint will be displayed per note above for Convenience score. A line plot will be provided with the mean of Oral Levothyroxine as a reference line and the mean 95% Confidence Interval XP-8121 results at each collection point.
Observed Values	Effectiveness Domain Score of TSQM-9	Per collection timepoint	Summarized by each protocol-specified timepoint using descriptive statistics (n, mean, SD, median, minimum, and maximum). A p-value from a paired t-test at each timepoint will be displayed per note above for Convenience score. A line plot will be provided with the mean of Oral Levothyroxine as a reference line and the mean 95% Confidence Interval XP-8121 results at each collection point.
Observed Values	Subject Preference Survey	Per collection timepoint	The subject counts and percentages for each answer to each question will be summarized.

9.8 Analysis of Safety Endpoints

9.8.1 Study Drug Exposure and Compliance

Subject exposure to treatment will be summarized by treatment period and overall. The duration of exposure, number of doses administered per subject, total volume (µL) administered, and total dose (µg) will be summarized using descriptive statistics (n, mean, SD, median, minimum, and maximum). In addition, the prior oral levothyroxine dose (µg) will be summarized using descriptive statistics (n, mean, SD, median, minimum, and maximum).

Subject dose level by visit will be summarized by subject counts and percentages.

9.8.2 Adverse Events

The number and percentage of subjects with AEs as well as event counts will be summarized by MedDRA-coded SOCs and PTs . The following AE summaries will be presented by treatment period and overall:

- Overall summary of events including
  - Any AEs
  - Any TEAEs
  - Any drug-related TEAEs
  - Any severe TEAEs
  - Any serious TEAEs
  - Any drug-related serious TEAEs
  - Any TEAEs leading to death
  - Any TEAEs leading to premature discontinuation of study drug
  - Any serious TEAEs leading to premature discontinuation of study drug
- Summary of all TEAEs by SOC and PT
- Summary of all TEAEs by SOC, PT, and severity
- Summary of all TEAEs by SOC, PT, and relationship
- Summary of serious TEAEs by SOC and PT
- Summary of TEAEs leading to premature discontinuation of study drug by SOC and PT
- Summary of all TEAEs by descending PT overall
- Summary of TEAEs occurring in 10% of subjects overall

Within each summary, a subject will be counted only once within each SOC and PT for the number and percentage of subjects, but will be counted for each occurrence of the AE for the event counts. TEAEs summarized by severity will count each subject once by maximum severity within each SOC and PT. TEAEs summarized by relationship will count each subject once in the order of Related, Not Related within each SOC and PT.

9.8.3 Clinical Laboratory Assessments

Visit values and changes from baseline to each post-baseline visit will be summarized with descriptive statistics (n, mean, SD, median, minimum, and maximum) for all chemistry, hematology, and urinalysis measures. Box plots of observed results and changes from baseline for chemistry, hematology, and urinalysis at each timepoint will be presented. A data table will be displayed under the plot with the mean value and the number of subjects with data at each timepoint.

All laboratory results will be displayed in a listing, and results that are outside the normal range will be identified on the listing as either high or low and CS. Separate listings for abnormal and CS laboratory results will be provided. Historical laboratory assessments will be listed.

9.8.4 Electrocardiograms

Visit values and changes from baseline to each post-baseline visit will be summarized with descriptive statistics (n, mean, SD, median, minimum, and maximum) by treatment period for all ECG parameters:

heart rate (beats/min), QT interval (ms), PR interval (ms), QRS interval (ms), RR interval (ms), QTcF (ms) and QTcB (ms).

Subject counts and percentages for ECG interpretation and clinical significance as reported on the eCRF by the investigator (and confirmed by cardiologist) will be presented for visit results and shifts from baseline to each visit. Cardiologist's overread will take precedence if it differs from investigator's assessment.

9.8.5 Vital Signs

Visit values and changes from baseline to each post-baseline visit will be summarized with descriptive statistics (n, mean, SD, median, minimum, and maximum) by treatment period for all vital signs: SBP (mmHg), DBP (mmHg), pulse rate (beats/min), respiration rate (breaths/min), temperature (C), and weight (kg).

CS categories for relevant vital signs will be summarized with subject counts and percentages by treatment period. CS results will be summarized by unique subject counts for any CS result, any CS Low result, and any CS High result within each measure and across all measures.

9.8.6 Physical Examination

Physical examination results will be listed.

9.8.7 Local Tolerability

The Modified Draize Scale, Injection Site Discomfort, time until onset (minutes), and duration of discomfort (minutes) will be summarized by descriptive statistics (n, mean, SD, median, minimum, and maximum) by each collection timepoint.

Discomfort description (i.e., pain, irritation, itching, etc.), each severity category for erythema/edema on the Modified Draize Scale, and the category of Injection Site Discomfort Intensity will be summarized by count and percentage by each collection timepoint.

9.9 Subgroup Analyses

If a sufficient number of subjects with secondary or tertiary hypothyroidism are included in the study, additional analyses for the secondary efficacy endpoints and non-PK exploratory endpoints may be performed on all subjects regardless of hypothyroidism diagnosis at the discretion of the sponsor.

10. COMPUTER SOFTWARE

Statistical analyses will be performed using SAS® version 9.4 or higher in a Windows environment, or R version 4.1.1 or higher, or Phoenix WinNonlin™ version 8.3 or higher.

11. REFERENCES

Colucci P, Yue CS, Ducharme M, Benvenga S. A review of the pharmacokinetics of levothyroxine for the treatment of hypothyroidism. Eur Endocrinol. 2013;9(1):40-47.

Laurent I, Tang S, Astere M, et al. Liquid L-thyroxine versus tablet L-thyroxine in subjects on L-thyroxine replacement or suppressive therapy: a meta-analysis. Endocrine. 2018;61(1):28-35.

12. APPENDIX

Not applicable.

13. TABLE/FIGURE/LISTING SHELLS

The mock shells for the tables/figures/listings will be provided in a separate document.