Protocol ID: DCC-3014-03-001

Official Title: A Phase 3, Randomized, Placebo-controlled, Double-blind Study of Vimseltinib to Assess the Efficacy and Safety in Patients With Tenosynovial Giant Cell Tumor (MOTION)

NCT Number: NCT05059262

SAP Approval Date: 09 Aug 2023

STATISTICAL ANALYSIS PLAN

Protocol Number:	**DCC-3014-03-001**
Protocol Title:	**A PHASE 3, RANDOMIZED, PLACEBO-CONTROLLED, DOUBLE-BLIND STUDY OF VIMSELTINIB TO ASSESS THE EFFICACY AND SAFETY IN PATIENTS WITH TENOSYNOVIAL GIANT CELL TUMOR (MOTION)**
IND Number:	131218
EudraCT Number:	2020-004883-25
Indication:	Tenosynovial Giant Cell Tumor
Phase:	3
Investigational Medicinal Product:	Vimseltinib (DCC-3014)
Sponsor:	Deciphera Pharmaceuticals, LLC 200 Smith Street Waltham, MA 02451
Protocol Version	Amendment 3
Date	31 May 2023
SAP Version	Version 2.0
Date	09 Aug 2023

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Table 1: List of Abbreviations

Abbreviation	Definition
AE	Adverse Event
BPI	Brief Pain Inventory
CGIC	Clinician Global Impression of Change
CGIS	Clinician Global Impression of Severity
CI	Confidence Interval
ClinRO	Clinician Reported Outcome
CR	Complete Response
DOR	Duration of Response
ECG	Electrocardiogram
eCRF	Electronic Case Report Form
EQ-5D-5L	5-level EQ-5D
EQ-VAS	EuroQol Visual Analogue Scale
FACT-G	Functional Assessment of Cancer Therapy - General
IRR	Independent Radiologic Review
ITT	Intent-to-Treat
KM	Kaplan-Meier
MedDRA	Medical Dictionary for Regulatory Activities
MCID	Minimally Clinical Important Difference
MMRM	Mixed Model for Repeated Measurements
mRECIST	Modified RECIST
NRS	Numeric Rating Scale
ORR	Objective response rate
PD	Progressive Disease
PGIC	Patient Global Impression of Change
PGIS	Patient Global Impression of Severity
PK	Pharmacokinetic
PR	Partial Response
PRO	Patient Reported Outcome
PROMIS	Patient-reported Outcomes Measurement Information System
PROMIS-PF	PROMIS-Physical Function
PT	Preferred Term
QOL	Quality of life
QTcF	QT interval corrected by Fridericia's formula
RECIST	Response Evaluation Criteria in Solid Tumors

ROM	Range of Motion
SAE	Serious Adverse Event
SAF	Safety Analysis Set
SAP	Statistical analysis plan
SD	Stable Disease
SOC	System Organ Class
TCGT	Tenosynovial Giant Cell Tumor
TDS	Tissue Damage Score
TEAE	Treatment-emergent adverse event
TVS	Tumor Volume Score

1. INTRODUCTION

This statistical analysis plan (SAP) contains definitions of analysis populations and endpoints, and provides a comprehensive description of statistical analyses to be implemented to assess the clinical efficacy, safety and tolerability of vimseltinib in the Phase 3, Randomized, Placebo-controlled, Double-blind Study of Vimseltinib to Assess the Efficacy and Safety in Patients with Tenosynovial Giant Cell Tumor (MOTION) (Protocol DCC-3014-03-001 Amendment 3, dated 31 May 2023).

If there are any differences between the protocol's statistical analysis section and the present document, this SAP will be the primary reference for the study's statistical analyses. Any differences will be highlighted in Section 11.

2. STUDY DESIGN

This is a multicenter, randomized, placebo-controlled study of vimseltinib in participants with Tenosynovial Giant Cell Tumor (TGCT), consisting of 2 parts: Part 1 is double-blinded and Part 2 is open label. Symptomatic participants with histologically confirmed TGCT for whom surgical resection will potentially cause worsening functional limitation or severe morbidity will be eligible. Participants who received anti-CSF1/CSF1R therapy previously (except for imatinib or nilotinib) will be excluded. The study will evaluate efficacy, safety, clinical outcome assessments, PK, and pharmacodynamics of vimseltinib in this population.

The study will consist of a 42-day screening period prior to the first dose of study drug, a Part 1 double-blinded treatment period of 24 weeks (referred to in 28-day cycles) and a Part 2 open-label period until Week 49 (Figure 1). Participants will continue treatment after Week 49 during the extension period. There will also be an End-of-Treatment (EOT) Visit within 7 days after the decision to stop study drug, a Safety Follow-up Visit 30 days (±5 days) after the last dose of study drug, and a Disease Follow-up period of up to 2 years or until initiation of new TGCT treatment or surgery, whichever occurs first. Participants will be allowed to undergo surgical resection only after completion of Part 1.

Approximately 120 participants will be randomized in a 2:1 ratio to receive either vimseltinib at the dose of 30 mg biw (n=80) or placebo (n=40) for 24 weeks. Randomization will be stratified for tumor location (lower limb vs. all other) and region (U.S. vs. non-U.S.).

At Week 25, the primary and secondary endpoints will be assessed, and participants randomized to placebo in Part 1 will have the option to crossover and receive open-label vimseltinib in Part 2 upon completion of Part 1. Participants randomized to placebo in Part 1 with confirmed disease progression by blinded independent radiologic review (IRR) before Week 25 are eligible for early entry into Part 2. Participants randomized to vimseltinib in Part 1 with confirmed disease progression by IRR before Week 25 will discontinue from study, while those without confirmed disease progression by IRR before Week 25 will continue to receive vimseltinib in Part 2 upon completion of Part 1.

3. STUDY OBJECTIVES

3.1. Primary Objective

- To evaluate anti-tumor activity of vimseltinib using RECIST v1.1 by blinded independent radiological review

3.2. Secondary Objectives

- To assess anti-tumor activity of vimseltinib using tumor volume score (TVS) and modified RECIST (mRECIST) by blinded independent radiological review
- To assess the effects of vimseltinib on range of motion (ROM)
- To assess the effects of vimseltinib on physical function, worst stiffness, worst pain, and quality of life (QoL) using patient-reported outcome (PRO) measures
- To assess safety and tolerability of vimseltinib

3.3. Exploratory Objectives

- To assess the correlation of pharmacokinetics (PK) with efficacy and/or safety
- To assess the pharmacodynamic effects of vimseltinib in relation to safety or efficacy
- To assess germline polymorphic variations in genes involved in the metabolism or disposition of vimseltinib or in relation to safety or efficacy
- To assess long-term safety and efficacy of vimseltinib
- To assess the effects of vimseltinib on symptomatic relief and functional assessments

4. STUDY ENDPOINTS

4.1. Primary Endpoint

- Objective response rate (ORR, including complete response [CR] and partial response [PR]) per RECIST v1.1 at Week 25

4.2. Key Secondary Endpoints

- ORR per TVS at Week 25
- Change from baseline in active ROM of the affected joint, relative to a reference standard, at Week 25
- Change from baseline in the Patient-reported Outcomes Measurement Information System (PROMIS) physical function score at Week 25
- Change from baseline in the worst stiffness numeric rating scale (NRS) score at Week 25
- Change from baseline in EQ-VAS (EuroQol Visual Analogue Scale) at Week 25

- Response of at least a 30% improvement in the mean Brief Pain Inventory (BPI) worst pain NRS score without a 30% or greater increase in narcotic analgesic use at Week 25

4.3. Other Secondary Endpoints

- ORR per RECIST v1.1
- ORR assessed by mRECIST at Week 25
- Duration of response (DOR; time from first PR or CR to disease progression or death) assessed using RECIST v1.1, TVS, and mRECIST
- Incidence of TEAEs, treatment-emergent serious adverse events, related TEAEs, dose reductions, dose interruptions, and discontinuation of study drug due to adverse event
- Changes from baseline in laboratory parameters, electrocardiograms (ECGs), and vital signs

4.4. Exploratory Endpoints

Pharmacokinetics

- Correlation of PK with efficacy and/or safety

Pharmacodynamics

- Effects of pharmacodynamics in relation to safety or efficacy

Pharmacogenomics

- Germline polymorphisms in genes involved in the metabolism or disposition of vimseltinib or related to safety and/or efficacy

Efficacy and functional assessments

- ORR at Week 49
- Change from baseline in tissue damage score (TDS)
- Change from baseline in active ROM of the affected joint, relative to a reference standard, at Week 49
- Change from baseline in active ROM of the affected joint, relative to the contralateral joint, at Week 25
- Change from baseline in passive ROM of the affected joint, relative to a reference standard, at Week 25
- Change from baseline in passive ROM of the affected joint, relative to a contralateral joint, at Week 25
- Clinician Reported Outcome (ClinRO) and PRO assessments at Week 13 and Week 49
- Additional ClinRO and PRO assessments:

- Response of at least a 30% improvement in the mean BPI average pain NRS score without a 30% or greater increase in narcotic analgesic use at Week 25
- Patient global impression of change (PGIC) and clinician global impression of change (CGIC) up to Week 25
- Patient global impression of severity (PGIS) and clinician global impression of severity (CGIS) up to Week 25
- GP5 "burden-of-side-effects" question from the Functional Assessment of Cancer Therapy-General (FACT-G) at Week 25: response of 3 ("quite a bit") or 4 ("very much")

5. SAMPLE SIZE DETERMINATION

The sample size selection of approximately 120 participants with TGCT (n~80 vimseltinib, n~40 placebo) was based on considerations for powering the analyses of the primary endpoint, key secondary endpoints, detection of rare safety events and overall exposure to vimseltinib, assuming 15% participant dropout. Participants will be randomized in a 2:1 ratio of vimseltinib versus placebo.

This sample size will have 98% power to detect a statistically significant difference between treatment groups, assuming true ORRs of 35% and 5% in the vimseltinib arm and the placebo arm, respectively, using a two-sided Fisher's exact test at a 5% type I error rate level.

6. RANDOMIZATION AND BLINDING

The study will randomize approximately 120 participants with TGCT in a 2:1 ratio (vimseltinib:placebo). Randomization will be stratified for tumor location (lower limb vs. all other) and region (U.S. vs. non-U.S.).

The present document, a comprehensive statistical analysis plan (SAP), will be approved by the Sponsor prior to the unblinding of study data.

Where required, safety personnel at the Sponsor or designee may be unblinded to a participant's study drug assignment to meet reporting requirements to regulatory agencies. In addition, the Investigator, participant, Sponsor, and study team will be unblinded at the time the participant has disease progression by RECIST v1.1 based on IRR.

Additional Sponsor representatives may be unblinded to some data for the purposes of ensuring adequacy of study conduct, including proper distribution of study drug.

At no point in time before official full study unblinding for the primary analysis will any aggregate efficacy and safety analyses by true treatment assignment be conducted by the Sponsor or Sponsor representatives, unless there is explicit permission to do so, for instance high level efficacy assessment for the purposes of a data monitoring committee (DMC) review to determine the adequacy of the risk/benefit profile of the drug as it pertains to the conduct of the clinical study.

7. ANALYSIS SETS

7.1. Screen Set

The Screen Set consists of all participants who have signed the informed consent form.

7.2. Intent-to-Treat (ITT) Set

The Intent-to-Treat (ITT) Set consists of participants who have been randomized to a study treatment regimen. Analysis will be performed according to the allocated treatment regimen. The ITT Set will be the primary analysis set for all the efficacy endpoints analyses.

7.3. Per-protocol (PP) Analysis Set

The Per-protocol (PP) Analysis Set consists of participants in the ITT Set with at least one postbaseline IRR tumor assessment who have no important protocol deviations (IPDs). IPDs will be pre-specified in a separate document. Participants with IPDs resulting in exclusion from the PP Set will be identified and documented prior to database lock and may include violations of key inclusion/exclusion criteria, noncompliance with study treatment, participant taking the wrong study treatment, or participant receiving prohibited concomitant medications or therapies.

The efficacy analyses performed on the PP Set will be supportive and treatment group will be based on actual treatment received.

7.4. Safety Analysis Set (SAF)

The Safety Set (SAF) consists of participants who have received at least one dose of study treatment. Analysis will be performed according to the treatment regimen actually received.

7.5. Pharmacokinetic Set

The Pharmacokinetic Set consists of participants who received at least one dose of vimseltinib and have at least 1 non-missing PK concentration in plasma reported for vimseltinib or [redacted].

7.6. PRO Set

The PRO Set consists of participants in the ITT Set who have valid baseline and at least one postbaseline PRO assessment.

8. DEFINITIONS AND CONVENTIONS

8.1. Baseline

For all evaluations unless otherwise noted, baseline is defined as the most recent non-missing measurement prior to the first administration of study drug.

For the PROs collected multiple times during the screening period such as the Brief Pain Inventory (BPI) Numeric Rating Scale (NRS) for worst pain and average pain, and the NRS for worst stiffness, baseline will be the average of measurements taken during the period prior to Cycle 1 Day 1 which are the 14 days of assessments closest to and prior to Cycle 1 Day 1. A

participant must have at least 4 assessments available to compute this average, otherwise it is set to missing. For crossover participants and data from Part 2, baseline is defined to most recent measurement prior to first administration of vimseltinib. For BPI and NRS, the crossover baseline is defined as the average of Weeks 2 and 3 of Cycle 6 in Part 1.

8.2. Study Days

Unless otherwise noted, study days of an evaluation are defined as number of days relative to the Cycle 1 Day 1 visit which is designated as Day 1, and the preceding day is Day -1, the day before that is Day -2, etc.

- For assessments on/after Cycle 1 Day 1, study days are calculated as

  Date of assessment – Date of Cycle 1 Day 1 + 1

- For assessments before Cycle 1 Day 1, study days are calculated as

  Date of assessment – Date of Cycle 1 Day 1

For adverse events (AEs), prior or concomitant medications, prior cancer therapies and procedures, prior or concomitant non-drug therapies and procedures, and study administration data, study day is defined from the first dose date and calculated as follows:

- For events on/after first dose date, study days are calculated as

  Date of event – first dose date + 1

- For events before first dose date, study days are calculated as

  Date of event – first dose date

8.3. Analysis Visit Window

For by-visit (or cycle) statistical summaries, data will be analyzed according to the scheduled visit assigned on the eCRF (also referred to as the nominal visit). Otherwise, data collected as unscheduled assessments will be mapped to a scheduled visit using a visit window from Cycle 1 Day 1. The visit window for a specific visit of interest will go from the midpoint between the preceding visit and the visit of interest to the midpoint between the visit of interest and the subsequent scheduled visit.

For PRO, cycles will be mapped based on the expected length of 28 days to determine where assessments will be assigned based on the participant's Day 1 of each cycle.

If there are more than one unscheduled result, visit date will be used to calculate study day and the valid assessment within the visit window closest to the target study day will be used. If two or more assessments are the same distance apart from the target study day, then the earliest assessment will be used. The early termination visit may also be mapped to a derived visit window. Data that are collected from unscheduled visits and cannot be mapped to a scheduled visit will not be included in the by-visit summary tables but will be presented in the listings.

8.4. Data Handling

For adverse events (AEs), handling rules for partial onset and resolution dates are detailed in Section 8.4.2. For the other safety data, unless otherwise stated, only observed data will be used for analyses.

8.4.1. Handling of Repeated and Unscheduled Assessments

In general, for by-visit summaries, data recorded within analysis visit window will be presented. The visit mapping rule is described in Section 8.3. Results that cannot be mapped will not be included in the by-visit summaries however will contribute to the best/worst case value and last on-treatment visit, where applicable.

Tumor assessments at unscheduled visits may be used for the derivation of the primary endpoint and other efficacy endpoints based on visit windows.

The repeated and unscheduled measurements will be presented in the listings.

8.4.2. Handling of Partial Dates for Adverse Events

For AEs with partial or missing start or stop dates, the following imputation approach will be used.

Partial start date:

- If the day of the month is missing, the onset day will be imputed as the first day of the month unless it is the same month and year as the start of study treatment. In this case, the onset date will be imputed as the first dose date of study treatment.
- If the onset day and month are both missing, the day and month will be imputed as January 1, unless the event occurred in the same year as the start of study treatment. In this case, the event onset will be imputed as the first dose date of study treatment to conservatively report the event as treatment emergent.
- A missing onset date will be imputed as the first dose date of study treatment. If the resulting onset date is after a reported date of resolution (AE end date), the onset date will be imputed as the date of resolution.

Partial end date:

- If the day of the month is missing, the end day will be imputed as the last day of the month or death date or data cut-off date (or end of study date), whichever is earliest.
- If the end day and month are both missing, the day and month will be imputed as December 31 or death date or data cut-off date (or end of study date), whichever is earliest.
- A missing end date will be left as missing if the event is ongoing. Otherwise, it will be imputed as the death date or data cut-off date (or end of study date), whichever is earliest.

Imputation of missing or partial dates for adverse events is used only to determine whether an event is treatment-emergent and calculation of duration of adverse events; data listings will present the missing or partial date as recorded in the eCRF.

In the event of a partial date for date of death, imputation will follow the approach described above for AE start date, unless there is an existing date of last contact later than the imputed date, in which case the date of death would be imputed as the day after the date of last contact.

8.4.3. Handling of Partial Dates for Original Diagnosis

For original diagnosis with partial or missing dates:

- No imputation will be done if the year is missing
- If the year is before the informed consent date, the missing day will be imputed as the first day of the month and missing month will be imputed as July
- If the year is the current year of informed consent date, the missing day will be imputed as first day of the month and missing month will be imputed as January

8.4.4. Handling of Partial Dates for Medications

When determining prior or concomitant medications, partial start dates of prior and concomitant medications will be assumed to be the earliest possible date consistent with the partial date.

- Medication start dates with a missing day and non-missing month and year will be imputed as the first day of the non-missing month, except for medications with a start date in the same month and year as the first dose of study treatment, in which case the start date will be imputed as the first dose date.
- Medication start dates with missing day and month and non-missing year will be imputed as the first day of the non-missing year (i.e., 01 January), except for medications with a start date in the same year as the first dose of study treatment, in which case the start date will be imputed as the first dose date.
- A missing medication start date will be imputed as the first dose date.
- Medications that are not ongoing and have a medication stop date with a missing day and non-missing month and year will have their stop date imputed as the last day of the non-missing month, except for medications with a stop date in the same month and year as the last study date, in which case the stop date will be imputed as the last study date.
- Medications that are not ongoing and have a medication stop date with missing day and month and non-missing year will have their stop date imputed as the last day of the non-missing year (i.e., 31 December), except for medications with a stop date in the same year as the last study date, in which case the stop date will be imputed as the last study date.

Imputation of missing or partial dates for medications is used only to determine whether a medication is a prior or concomitant medication. Data listings will present the missing or partial date as recorded in the eCRF.

8.4.5. Handling Partial Dates for Non-drug Therapy/Procedure

Imputation of missing or partial dates for non-drug therapy/procedure will follow the same approach as will be used for medications. This imputation will be used only to determine

whether a therapy/procedure is prior or concomitant therapy/procedure. Data listings will present the missing or partial date as recorded in the eCRF.

9. STATISTICAL CONSIDERATIONS

All analyses described in this plan are considered a priori analyses in that they have been defined prior to locking the database. All other analyses, if any, designed subsequently to locking the database, will be considered post hoc analyses and will be described as exploratory analyses in the clinical study report (CSR) or a separate document.

All summaries and statistical analysis will be performed using SAS v9.4 or higher.

9.1. General Methods

Data collected in this study will be documented using summary tables and participant data listings. Continuous variables will be summarized using descriptive statistics (number of observations, mean, median, standard deviation, minimum, and maximum). Categorical variables will be summarized using frequencies and proportions. Time-to-event data will be summarized via Kaplan-Meier (KM) methodology using the 25th, 50th (median), and 75th percentiles with associated 2-sided 95% confidence intervals (CIs).

Medical history, AEs, and concurrent procedures will be coded using Medical Dictionary for Regulatory Activities (MedDRA) version 26.0 (or higher). Prior and concomitant medications will be coded using the World Health Organization (WHO) Drug Dictionary March 2023 Global B3 or later.

This study comprises 2 periods: the randomized double-blinded period (Part 1) and the open-label portion that occurs after unblinding the study treatment at Week 25 (or earlier if there is a treatment discontinuation due to a centrally confirmed disease progression) (Part 2).

Unless otherwise specified, the ITT and PP Sets are used for efficacy analysis and the SAF is used for the safety analysis.

The primary analysis will be conducted at the earliest when all participants have completed the double blinded treatment period (Part 1).

In general, unless otherwise specified, data summaries will be presented by treatment group (vimseltinib versus placebo) and overall. Between-group comparisons will focus on the comparative performance of vimseltinib versus placebo. All statistical tests will be conducted at a 2-sided significance level of 0.05.

9.2. Participant Disposition

The number and percentage of participants achieving some planned assessment listed below will be summarized by treatment group and overall (if applicable):

- Summary of randomization by country and site
- Summary of randomization stratification
- Summary of analysis sets (ITT, PP, SAF, PK, PRO)

- Participant disposition:
    - Screened
    - Randomized
    - Treated
        - Entered double-blinded treatment phase (Part 1)
            - Discontinued double-blinded treatment
                - Reason for discontinuation of treatment
            - Completed Part 1 treatment
            - Discontinued study during Part 1
                - Reason for discontinuation of study
        - Entered open label phase (Part 2) (received at least one dose of study treatment in Part 2)
            - Ongoing on treatment
            - Discontinued open label phase treatment
                - Reason for discontinuation of treatment
            - Discontinued study during Part 2
                - Reason for discontinuation of study
        - Discontinued from treatment
            - Reasons for discontinuation of treatment
    - Discontinued from study
        - Reason for study discontinuation

Listings will be provided for screen failures, as well as treatment and study discontinuations.

9.3. Protocol Deviations

A protocol deviation is any change, divergence, or departure from the study design or procedures defined in the clinical protocol. All protocol deviations will be reviewed and documented before database lock.

Protocol deviations will be classified as major or minor by medical review prior to the primary analysis. Major protocol deviations will be identified as those deviations from the study protocol that may significantly impact the completeness, accuracy, and/or reliability of the trial data; that may significantly affect a participant's rights, safety, or well-being. A subset of major protocol deviations referred to as IPDs will also be identified prior to the primary analysis. IPDs will be pre-specified in a separate document, they will consist of major protocol deviations resulting in the exclusion from the PP Set because they may affect efficacy. The number and percentage of participants with major deviations and IPDs will be tabulated for the ITT Set.

All protocol deviations will be listed for the ITT Set.

9.4. Demographics and Baseline Characteristics

Demographics and baseline characteristics will be summarized descriptively by randomized treatment assignment for the ITT Analysis Set. They will also be summarized by actual treatment received for the SAF and PP Set. The following variables will be included in the summary table:

- Age (years)
- Age category ($18 \leq age < 50$, $50 \leq age < 65$, $65 \leq age < 75$, $75 \leq age < 85$, $age \geq 85$)
- Sex (Male vs Female)
- Childbearing potential (Yes vs No)
- Race
- Ethnicity
- Baseline Weight (kg)
- Baseline Height (cm)
- Body Mass Index (BMI, kg/m^2)
- BMI Interpretation (underweight [<18.5], normal [18.5 – <25.0], overweight [25.0 – <30.0], obese [≥30.0])
- Tumor location (lower limb vs all other)
- Region (U.S. vs non-U.S.)
- Baseline narcotic use (Yes vs No)

Demographics and baseline characteristics will be listed for the ITT population.

9.5. Disease History

Participants' disease history will be summarized for ITT and SAF including:

- Time from original diagnosis to first dose date (years), calculated as

  (First dose date – original histopathological diagnosis date)/365.25
- Disease subtype
- Primary affected joint

Disease history listing will be provided for the SAF.

9.6. Prior TGCT Therapies and Procedures

Prior therapies and procedures will be tabulated for ITT and SAF including:

- Any prior surgery
- Any prior radiation therapy
- Any systemic therapy
    - Imatinib

  - Nilotinib
  - Other
- Best overall response to last therapy
- Reason for discontinuation of last therapy/procedure

Prior TGCT therapies and procedures will be listed for the SAF.

9.7. Prior and Concomitant Non-drug Therapy/Procedure

Prior non-drug therapies/procedures are defined as any non-drug therapies/procedures, with a stop date before the first dose of study drug.

Concomitant non-drug therapies/procedures are defined as any non-drug therapies/procedures, performed on or after the first dose of study drug and up to the end of study. Non-drug therapies/procedures that started before the first dose of study drug and were ongoing on the date of the first dose of study drug will be considered as both prior and concomitant non-drug therapies/procedures.

Prior and concomitant non-drug therapies/procedures will be coded using MedDRA version 26.0 (or higher) and will be summarized by system organ classifications (SOC) and preferred term (PT).

Prior and concomitant non-drug therapies/procedures will be listed for the SAF.

9.8. Prior and Concomitant Medications

Prior and concomitant medications will be coded using the latest World Health Organization (WHO) Drug Dictionary March 2023 Global B3 version or later.

Prior medications are defined as any medications, other than study drug, with a stop date before the first dose of study drug.

Concomitant medications are medications, other than study drug, being taken on or after the first study drug dose date and up to 30 days after the date of the last study drug dose. Medications that started before the first dose of study drug and were ongoing on the date of the first dose of study drug will be considered as both prior and concomitant medications. Prior and concomitant medications will be summarized by the WHO Drug Dictionary Anatomical Therapeutic Chemical 2nd level (ATC-2) and preferred name, for the SAF. If the ATC-2 level term is not available, the next available level (e.g., ATC-1) will be used. The numbers and percentages of participants who received the medications will be tabulated. A participant who takes more than one medication will be counted only once if these medications belong to the same extended ATC classification.

Prior and concomitant medications will be listed for the SAF.

Narcotic analgesic medication will be included in the prior and concomitant medications, but specific usage will be collected in participant diary. Diary data will be listed for the SAF.

9.9. Medical History

Medical history will be coded using the MedDRA version 26.0 or higher.

The frequency and percentage of participants experiencing any medical conditions will be tabulated by SOC and PT. If a PT or SOC was reported more than once for a participant, the participant will only be counted once in the incidence for that PT or SOC.

Medical history data will be listed for the SAF.

9.10. Efficacy Analyses

Results for endpoints based on assessment by the IRR (RECIST v1.1, mRECIST and TVS) will be obtained through review by two independent radiologists and adjudication as needed. Analyses of IRR results will be based on the assessments of reviewer number 1 (assigned independently by the IRR vendor), unless reviewer number 2 was selected through the adjudication process performed by the IRR vendor. This process is described in an Independent Review Charter.

9.10.1. Multiplicity Control

To control overall type I error, a hierarchical testing procedure will be utilized. Statistical testing will be performed for the analysis of primary and key secondary endpoints in the following order at a 2-sided 0.05 alpha level for each:

1. ORR per RECIST v1.1 at Week 25
2. ORR per TVS at Week 25
3. Mean change from baseline in active ROM at Week 25
4. Mean change from baseline in the PROMIS-physical function score at Week 25
5. Mean change from baseline in the worst stiffness NRS score at Week 25
6. Mean change from baseline in EQ-VAS at Week 25
7. Proportion of responders based on BPI-30 (worst pain) NRS score and narcotic analgesic use at Week 25

If the primary endpoint is not statistically significantly different between treatment groups, the secondary endpoints will not be formally tested and nominal p-values will be provided. Contingent upon significance in the primary endpoint, each of the secondary efficacy endpoints will be tested sequentially in the order above. If any of the endpoints are not statistically significantly different between treatment groups, all subsequent endpoints will not be formally tested and nominal p-values will be provided.

9.10.2. Primary Endpoint

Objective response rate (ORR) per RECIST v1.1 at Week 25 is defined as the proportion of participants with a CR or a PR as the Week 25 Tumor Response (as defined in Table 2) based on IRR per RECIST v1.1. ORR at Week 25 will be compared between the 2 treatment groups using a two-sided Cochran-Mantel-Haenszel (CMH) test stratified by the randomization stratification factors. The test will be performed at a 0.05 alpha level on the ITT Set. A 95% CI for the proportion in each arm using the Clopper-Pearson method as well as the difference in proportion and its associated Wald 95% CI will be presented. A Fisher's exact test and an exact 95% CI for the difference in proportion will be provided as a sensitivity analysis.

No confirmation (ie, CR or PR at the subsequent MRI assessment) will be required for a CR or PR per RECIST v1.1 as this is a randomized study. Complete and partial response will define response for the primary endpoint and additional efficacy analyses.

Determination of an overall response for each timepoint is based on the combination of responses for target lesions, and the presence or absence of one or more new lesions per RECIST v1.1. For this study, determination of the tumor response status for each participant in Part 1 with respect to the primary efficacy endpoint is shown in Table 2 below. All unscheduled assessments prior to Week 25 will be incorporated into the determination of ORR at Week 25 in a similar fashion as Week 13. End of Part 1 scans will be windowed according to Section 8.3. If a participant is missing a scan at Week 25, then they will be considered a non-responder.

Table 2: Definitions of Response for the Primary Efficacy Endpoint

Timepoint Response at Week 13	**Timepoint Response at Week 25**	**End of Part 1 (Week 25) Tumor Response Status (Primary Efficacy Endpoint)**
CR or PR	CR	Response (CR)
CR or PR	PD	Non-response (PD)
PR	Non-CR/non-PD/non-NE[a]	Response (PR)[b]
SD	CR or PR	Response (CR or PR)
SD	SD	Non-response (SD)
SD	PD	Non-response (PD)[c]
CR, PR, SD, and NE	NE	Non-response (NE)
PD	Any	Non-response (PD)
NE	CR or PR	Response (CR or PR)
NE	SD or PD	Non-response (SD or PD)

Abbreviations: CR=complete response; NE=not evaluated or unevaluable; PD=progressive disease; PR=partial response; SD=stable disease.

[a] Neither sufficient shrinkage to qualify for CR nor sufficient increase to qualify for PD, taking as reference the nadir at Week 13.

[b] A tumor that has achieved the criteria of PR will be considered an ongoing PR until progressive disease is objectively documented.

[c] To be considered SD, the tumor must achieve the criteria for SD at the Week 25 visit; shorter duration SD will not be considered SD at the End of Part 1.

A forest plot summarizing the difference in ORR at Week 25 between vimseltinib and placebo and its associated Wald 95% CI will be generated for the subgroups listed in Section 9.10.5 to evaluate the consistency of treatment effects in the subgroups.

A sensitivity analysis will compare the ORR *by* Week 25 between the 2 treatment groups using the same approach used for the primary endpoint analysis. In this analysis the assignment of response would differ from the approach described in Table 2 in that any participant who met the criteria for response at least once up to, and including, Week 25 would be counted as a responder.

The analysis may also be performed on the PP Set if the proportion of participants from the ITT Set who are excluded from PP Set exceeds 5%.

Response data by IRR, including the at Week 25 derived response, will be presented in by-participant listings.

9.10.3. Key Secondary Endpoints

As specified in the following sub-sections, multiple secondary endpoints will be analyzed using a mixed model for repeated measurements (MMRM) using the sandwich estimator to estimate the variance-covariance matrix. The dependent variable will be the change from baseline. Each of these models will include fixed effects for treatment group, timepoint, treatment group by timepoint interaction, stratification factor for region (U.S. versus non-U.S.), stratification factor for tumor location (lower limb vs. all other), and the baseline value of the corresponding endpoint. Statistical comparisons between treatment groups will be made at the specified timepoint. For the analysis of ROM only, tumor location will be replaced with joint type (knee, ankle, or other). An unstructured variance-covariance matrix will be used. If the unstructured variance-covariance matrix fails to converge, then alternative structures will be utilized. Statistical comparisons between treatment groups will be made based on a contrast statement at Week 25.

Analyses of the secondary endpoints described in the following sub-sections will be limited to data from Part 1 and will focus on comparisons between participants who received vimseltinib versus participants who received placebo during this period. The exception to this rule is for duration of response presented in Section 9.10.4.3.

These analyses may also be performed on the PP Set if the proportion of participants from the ITT Set who are excluded from PP Set exceeds 5%.

A separate analysis plan will describe the process for identifying the minimally clinical important difference (MCID) for PROMIS-PF, Worst Stiffness, and ROM. A responder analysis for each of these endpoints based on the MCID of each endpoint as determined by the [redacted] analysis will be performed.

9.10.3.1. Tumor Response by Tumor Volume Score (TVS)

In addition to being assessed using RECIST v1.1 (primary endpoint), tumor response will also be assessed using TVS. Objective response rate by TVS is defined as the proportion of participants with a CR or PR.

Tumor volume score ORR at Week 25 will be compared between the 2 treatment groups using the same approach used for the primary endpoint analysis. The tumor response status at Week 25 based on TVS will follow the same approach described in Table 2 for the primary endpoint.

Sensitivity analyses similar to the ones planned for the primary endpoint will be performed.

A forest plot summarizing the difference in TVS ORR at Week 25 between vimseltinib and placebo and its associated Wald 95% CI will be generated for the subgroups listed in Section 9.10.5 to evaluate the consistency of treatment effects in the subgroups.

All TVS data will be presented in by-participant listings.

9.10.3.2. Range of Motion (ROM)

Measurement of the affected and contralateral, non-affected joint will be assessed using a goniometer. The measurement (in degrees) of the affected joint will be used to derive a relative ROM obtained through normalization to the measurement from a reference standard value provided by the American Medical Association per motion and type (active or passive).

At screening, the motion with the smallest relative ROM value (worst) will be identified, and this motion will be used for evaluating the change in relative ROM subsequently. Only the motion with the most impaired ROM at screening will be selected for subsequent analyses. If there are ties, the multiple motions with the same relative ROM value at screening will be identified, and the average of relative ROM values will be calculated at each post screening visit and be used in the analysis as the single relative ROM value for that specific visit.

The main analysis of relative ROM will be based on the active measurement relative to a reference standard and will consist in a comparison between treatment groups of the mean change from baseline at Week 25 using a MMRM. The different timepoints that will be included in the model are Weeks 13 and 25, which respectively correspond to Day 1 of Cycles 4 and 7. The observed value and change from baseline in relative ROM will be summarized with descriptive statistics by timepoint and treatment group. The mean change from baseline will be calculated and presented in figure format at each timepoint by treatment group. In addition, such mean changes may be displayed by joint type and/or motion.

A forest plot summarizing the 95% CI of the difference in mean of change from baseline at Week 25 between vimseltinib and placebo without adjusting for the stratification factors will be generated for the subgroups listed in Section 9.10.5 to evaluate the consistency of treatment effects in the subgroups.

Additional analyses of these data will be performed on results based on the passive measurement and results normalized to the measurement obtained from the contralateral, non-affected joint.

All ROM data will be presented in by-participant listings.

9.10.3.3. Physical Function

The main analysis of PROMIS-PF will consist of a comparison between treatment groups of the mean change from baseline at Week 25 using a MMRM. The different timepoints that will be included in the model are Weeks 5, 9, 13, 17, 21, 25 which respectively correspond to Day 1 of Cycles 2, 3, 4, 5, 6, and 7. Values for analysis will be the earliest collected value in each cycle.

Fifteen questions from the PROMIS-PF item bank will be asked regardless of tumor location for the individual participant. The questions use one of two 5-point verbal rating scales: either 1 = “unable to do”, 2 = “with much difficulty”, 3 = “with some difficulty”, 4 = “with a little difficulty”, and 5 = “without any difficulty”; or 1 = “cannot do”, 2 = “quite a lot”, 3 = “somewhat”, 4 = “very little”, and 5 = “not at all." There is no specified recall period. The subset of questions used for scoring is based on the location of the tumor (upper or lower body) based on Gelhorn et al (6). The PROMIS-PF is scored using the Assessment Center API using the table below to generate a T-score which is used for analysis where “x” denotes the question is utilized for scoring. The T-score rescales the raw sum score into a standardized score with mean of 50 and a SD of 10 based on the US general population. If a participant answers at least half of the questions for their tumor location, then the T-score generated based on the API

scoring manual will be used for analysis and those questions with missing responses are ignored for scoring. If a participant answers less than half of the questions for their tumor location, then the T-score will be calculated and presented in listings but will not be used for analysis.

Table 3: PROMIS Scoring Based on Tumor Location

Lower	Upper	Question
x	x	Are you able to push open a heavy door?
x	x	Are you able to exercise for an hour?
x	x	Are you able to carry a heavy object (over 10 pounds/5kgs)?
x	x	Are you able to dress yourself, including tying shoelaces and buttoning your clothes?
x	x	Are you able to carry a laundry basket up a flight of stairs?
x		Are you able to stand for one hour?
x		Are you able to go up and down stairs at a normal pace?
x		Are you able to go for a walk for at least 15 minutes?
	x	Are you able to lift 10 pounds (5kg) above your shoulder?
	x	Are you able to change a light bulb overhead?
x	x	Does your health now limit you in doing heavy work around the house like scrubbing floors, or lifting or moving heavy furniture?
x	x	Does your health now limit you in lifting or carrying groceries?
x	x	Does your health now limit you in doing moderate work around the house like vacuuming, sweeping floors or carrying in groceries?
x	x	Does your health now limit you in going OUTSIDE the home, for example to shop or visit a doctor's office?
x		Does your health now limit you in bending, kneeling, or stooping?

The observed value and change from baseline in PROMIS-PF T-score will be summarized with descriptive statistics by timepoint and treatment group. The mean change from baseline will be calculated and presented in figure format at each timepoint by treatment group.

A forest plot summarizing the 95% CI of the difference in mean of change from baseline at Week 25 between vimseltinib and placebo without adjusting for the stratification factors will be generated for the subgroups listed in Section 9.10.5 to evaluate the consistency of treatment effects in the subgroups.

All PROMIS-PF data will be presented in by-participant listings.

9.10.3.4. Worst Stiffness Numeric Rating Scale (NRS)

The main analysis of worst stiffness NRS will consist of a comparison between treatment groups of the mean change from baseline at Week 25 using a MMRM. The different timepoints that will be included in the model are Weeks 5, 9, 13, 17, 21, 25 which respectively correspond to Day 1 of Cycles 2, 3, 4, 5, 6, and 7. For each of these timepoints, the value used for analysis will be the average of the results obtained at weeks 2 and 3 of the previous cycle. For example, the value used for a participant at Week 25 (Cycle 7 Day 1) will be the average of the observations collected during the 2nd and 3rd week of Cycle 6. At least 4 measurements are required to compute this average, otherwise it is set to missing.

The observed value and change from baseline in worst stiffness NRS will be summarized with descriptive statistics by timepoint and treatment group. The mean change from baseline will be calculated and presented in figure format at each timepoint by treatment group.

A forest plot summarizing the 95% CI of the difference in mean of change from baseline at Week 25 between vimseltinib and placebo without adjusting for the stratification factors will be generated for the subgroups listed in Section 9.10.5 to evaluate the consistency of treatment effects in the subgroups.

All worst stiffness NRS data will be presented in by-participant listings.

9.10.3.5. EQ-5D-5L

The main analysis of the EQ-VAS will consist of a comparison between treatment groups of the mean change from baseline at Week 25 using a MMRM. The different timepoints that will be included in the model are Weeks 5, 9, 13, 17, 21, 25 which respectively correspond to Day 1 of Cycles 2, 3, 4, 5, 6, and 7. Values for analysis will be the earliest collected value in each cycle.

The observed value and change from baseline in EQ-VAS will be summarized with descriptive statistics by timepoint and treatment group. The mean change from baseline will be calculated and presented in figure format at each timepoint by treatment group.

A forest plot summarizing the 95% CI of the difference in mean of change from baseline at Week 25 between vimseltinib and placebo without adjusting for the stratification factors will be generated for the subgroups listed in Section 9.10.5 to evaluate the consistency of treatment effects in the subgroups. The EQ-5D-5L will be summarized by frequency and percentage for each level of each dimension by timepoint and treatment group.

All EQ-5D-5L data will be presented in by-participant listings.

9.10.3.6. Brief Pain Inventory – Response

A responder analysis based on the BPI worst pain NRS item and analgesic use will be performed. A responder will be defined as a participant who: (i) experienced a decrease of at least 30% in the mean BPI worst pain NRS item and (ii) did not experience a 30% or greater increase in narcotic analgesic use. The change in BPI worst pain NRS for responder assessment will be assessed by comparing data collected during a 14-day period prior to the current visit with baseline values collected prior to the first dose of study drug. This will be referenced as BPI-30 response. A value for BPI worst pain NRS will be computed for each participant at Day 1 of individual cycles. This value will be the average of the measurements obtained during weeks 2 and 3 of the previous cycle. At least 4 measurements are required during weeks 2 and 3 of the previous cycle to compute this average, otherwise it is set to missing. If it is set to missing, BPI-30 response for that cycle will be assigned as a non-response. For derivation of BPI-30 response, the narcotic analgesic use will be evaluated for periods corresponding to the BPI worst pain NRS assessment periods. For participants who use different types of narcotic analgesic and/or dosage levels, analgesic use will be calculated by converting equianalgesic dosing to morphine-equivalent doses as described in Table 4 (3, 4, 5). A sensitivity analysis which defines a responder as experienced a decrease of at least 30% in the mean BPI worst pain NRS item and no increase in narcotic analgesic use.

Table 4: Narcotic Analgesic Drug (NAD) conversions to morphine-equivalent dose

Narcotic Analgesic Drug	Conversion Factor
Buprenorphine (SL)	80
Buprenorphine (TD)	100
Codeine	0.1
Dihydrocodeine	0.1
Fentanyl patch	100
Fentanyl (SL)	0.125
Hydrocodone	0.67
Hydromorphone	4
Levorphanol	11
Methadone 1-20 mg/day 21-40 mg/day 41-60 mg/day ≥61-80 mg/day	 4 8 10 12
Meperidine hcl	0.1
Morphine	1
Oxycodone	1.5
Oxymorphone	3
Pethidine	0.125
Pentazocine	0.37
Propoxyphene	0.23
Tapentadol	0.33
Tramadol	0.1

Conversion factors were sourced from WHO guidelines (4) when available and next from CDC guidelines (3). In cases where neither WHO nor CDC guidelines provided a morphine-equivalent conversion factor, alternate guideline such as CONSORT (5) were used.

BPI-30 response at Week 25 will be compared between the 2 treatment groups using the same approach used for the primary endpoint analysis. A summary of narcotic use over time will be provided.

A forest plot summarizing the unstratified 95% CI of the difference in BPI response rate at Week 25 between vimseltinib and placebo will be generated for the subgroups listed in Section 9.10.5 to evaluate the consistency of treatment effects in the subgroups.

All BPI data will be presented in by-participant listings.

9.10.4. Other Secondary Endpoints

9.10.4.1. Objective Response Rate per RECIST v1.1

Objective response rate per RECIST v 1.1 over the entire study, defined as the number of participants with a BOR of CR or PR at any time on study, and its corresponding 95% CI will be presented.

The best overall response (BOR) up to a specific timepoint for an individual participant will be determined as follows:

- Complete response (CR): at least one CR result before any Progressive Disease result, up to and including the specific timepoint of interest
- Partial response (PR): at least one PR result before any Progressive Disease result (and not qualifying for a CR), up to and including the specific timepoint of interest
- Stable disease (SD): at least one SD before any Progressive Disease result (and not qualifying for a CR or PR), up to and including the specific timepoint of interest
- Progressive disease: a result of Progressive Disease up to and including the specific timepoint of interest, not preceded by any result of CR, PR or SD
- Not Evaluable (NE): not meeting the above definitions, and no Progressive Disease before the specific timepoint of interest and the specific timepoint of interest assessment is NE or missing

The primary analysis of BOR will be conducted based on RECIST v1.1 based on IRR. Supportive analyses based on TVS and mRECIST from IRR and based on investigator assessment per RECIST v1.1 will be provided.

9.10.4.2. Tumor Response per mRECIST

Objective response rate by mRECIST at Week 25 will be compared between 2 treatment groups using a two-sided Cochran-Mantel-Haenszel test stratified by the randomization stratification factors at the 0.05 alpha level based on the ITT Set. The 95% CI in each arm and the difference, along with its 95% CI, will be presented. Although it is not part of the hierarchical testing, the analysis will use the same method used for the primary endpoint analysis. The tumor response status at Week 25 based on mRECIST will follow the same approach described in Table 2 for the primary endpoint. The criteria for mRECIST is identical to RECIST v1.1, except that measurements are taken along the short-axis dimension perpendicular to the longest dimension and ideally also to a reproducible adjacent landmark, such as the femoral bone or a tendon.

All tumor response per mRECIST data will be presented in by-participant listings.

9.10.4.3. Duration of Response

Duration of response (DOR) is defined as the time from the first documented objective response (CR or PR) until the time of disease progression or death by any cause, whichever occurs earlier. DOR will be summarized in two sets of participants, firstly for participants with an objective response at Week 25 and secondly for those who achieved CR or PR as BOR on study to the study treatment in the ITT Set. DOR in weeks is calculated as

- (Earlier of Date of Progressive disease or death, or censoring – date of first response + 1) / 7

Participants will be censored according to following rules:

- For participants who undergo surgical resection of target or non-target lesions, or have received anti-tumor treatments other than the study treatment before documented date of the first disease progression, DOR will be censored at the date of the latest evaluable progression-free radiologic assessment prior to start date of the anti-tumor therapy.

- For participants who have not progressed nor died, DOR will be censored at the last evaluable radiologic assessment.
- For participants who have first disease progression or die after two or more consecutive missed/non-evaluable assessments, DOR will be censored at the time of the evaluable radiologic assessment immediately prior to the two or more consecutive missed/non-evaluable radiologic assessments.

DOR will be summarized for responders in each of the following three groups:

1. DOR in vimseltinib arm: Participants randomized to receive vimseltinib in Part 1 using data from Part 1 and Part 2 (including extension period);
2. DOR in placebo arm: Participants randomized to receive placebo in Part 1, using all available data until initial dose of vimseltinib in Part 2;
3. DOR after crossover to vimseltinib: Participants randomized to receive placebo in Part 1 and crossed over to vimseltinib in Part 2.

For Group (1), tumor assessment data from both Part 1 and Part 2 (including extension period) will be combined, and the baseline assessment will be the one recorded during the Screening period. Censoring for Group (1) will follow the rules listed above. For Group (2), the duration of response will be censored based on the rules listed above with vimseltinib being considered as another anti-tumor treatment, and with the screening assessment serving as baseline. For Group (3), the tumor assessment data in Part 2 (including extension period) will be used, with the baseline assessment being the last one before the first dose of vimseltinib in Part 2. Censoring for Group (3) will follow the rules listed above.

DOR time will be summarized via Kaplan-Meier (KM) methodology including the total number and percentage of participants with events and by type of event (disease progression or death), total number and percentage of censored participants and by censoring reason, the DOR percentiles (25^{th}, 50^{th} [median], and 75^{th}), DOR probabilities at pre-specified timepoints (e.g., 25 weeks and 49 weeks if applicable) with associated 2-sided 95% confidence intervals, and KM plot of the estimated survival functions for DOR.

Duration of response will also be summarized for Groups (1) and (3) combined, reflecting all responses observed while receiving vimseltinib.

KM plots will be provided to graphically present the DOR by group.

The analysis described above for DOR for those participants with an objective response at Week 25 will be performed separately based on RECIST v1.1 and TVS per IRR.

The analysis described above for DOR for those who achieved CR or PR as BOR on study will be performed separately based on RECIST v1.1, TVS, and mRECIST data. Duration of response data will be presented in a listing.

9.10.5. Sub-group Analyses

The primary efficacy endpoint and key secondary efficacy endpoints will be analyzed in participant subgroups defined as follows:

- Tumor location (lower limb vs all other)

- Region (U.S. vs non-U.S.)
- Participants with disease located in large joints (shoulder, elbow, hip, or knee) vs small joints (joints other than the shoulder, elbow, hip, or knee)
- Participants with disease located in the knee
- Disease subtype (localized vs diffuse)
- Participants at sites in the EU region only
- Age ($18 \leq age < 50$, $50 \leq age < 65$, $age \geq 65$)
- Gender (female vs male)
- Prior surgery (yes vs no)
- Prior imatinib/nilotinib (yes vs no)

The applicable key secondary efficacy endpoints are TVS, ROM, PROMIS-PF, worst stiffness NRS, EQ-VAS and BPI-30 response.

9.10.6. Sensitivity Analyses for Missing Data

The preponderance of missing data for the key secondary efficacy endpoints of Active ROM, PROMIS-PF, Worst Stiffness NRS, and EQ-VAS will be evaluated. If there is more than 10% of missing data at Week 25 in any of the endpoints, then the impact of the missing data will be explored for that endpoint and will include, but may not be limited to, the following analyses:

- Multiple imputation based on randomized treatment to mimic missing at random
- A Pattern Mixture Model approach where participants who discontinued treatment due to an adverse event on vimseltinib will be imputed based on the missing not at random assumption and be imputed as if they received placebo.

In any exploration of missing data impact where a seed is required, the seed will be set to 20230428.

9.10.7. Exploratory Efficacy Analysis (Part 1 & 2)

There will be no formal testing for exploratory endpoints, descriptive statistics will be presented and nominal p-values may be provided.

As will be done for safety analyses, exploratory efficacy endpoints presented over time will be summarized following two different approaches. The first approach will be limited to data from Part 1 and will focus on comparisons between participants who received vimseltinib versus participants who received placebo during this period. The second approach will combine data from both Part 1 and Part 2 (including extension period) and will focus on the observations and assessments while participants are receiving vimseltinib. Therefore, in this second approach, participants randomized to Placebo in Part 1, will have as their reference day the first dose date of vimseltinib which will occur in Part 2, and their data from Part 1 while receiving placebo will be excluded.

For observed value and change from baseline analyses, summary statistics at each scheduled visit will be presented.

9.10.7.1. Objective Response Rate at Specific Timepoints

Response at a specific timepoint will be derived based on reported overall tumor responses at different evaluation timepoints from the date of first dose up to and including the specific timepoint assessment. Response at Week 25 will be determined based on Section 9.10.2 and responses at other timepoints will be determined based on:

- Complete response (CR): at least one CR reported up to and including the specific timepoint assessment and the specific timepoint assessment is not Progressive Disease or NE
- Partial response (PR): at least one PR but no CR reported up to and including the specific timepoint assessment and the specific timepoint assessment is not Progressive Disease or NE
- Stable disease (SD): at least one SD but no CR nor PR reported up to and including the specific timepoint assessment and the specific timepoint assessment is not Progressive Disease or NE
- Progressive disease: Progressive Disease at the specific timepoint assessment or Progressive Disease before the specific timepoint regardless of response status at the specific timepoint
- Not Evaluable (NE): no Progressive Disease before the specific timepoint and the specific timepoint assessment is NE or missing

ORR at Week 25, 49 or 73 is defined as the proportion of participants with Response at the specific time point being either CR or PR. The denominator for will include all participants in the ITT Set who have reached the Week 25, 49 or 73 timepoint or have discontinued from treatment or the study prior to that timepoint. An exact 95% confidence interval (CI) on the rates of response will also be presented.

These exploratory analyses will be conducted based on RECIST v1.1, TVS and mRECIST from IRR, and based on investigator assessment per RECIST v1.1.

Swimlane plots will present objective response and time to response duration for individual participants where time to response is defined as date of first CR or PR - treatment start date + 1.

9.10.7.2. Clinical Benefit Rate

Clinical benefit rate (CBR) is defined as the proportion of participants who have a response (CR or PR) or stable disease (SD). CBR will be tabulated at Week 25, 49, and 73. CBR will be calculated among participants in the ITT Set. The denominator will include all participants in the ITT Set who have reached the Week 25, 49 or 73 timepoint or have discontinued from the study prior to that timepoint.

Clinical benefit at Week 25, 49 or 73 will be derived based on reported overall tumor responses at different evaluation timepoints from the date of first dose up to and including the Week 25, 49 or 73 timepoint.
Clinical benefit status at Week 25, 49, or 73:

- Clinical Benefit (CB):

  - at least 1 CR or PR result, up to and including the specific timepoint of interest
  - only SD or NE results obtained up to the specific timepoint of interest and a SD result at the specific timepoint of interest
- No Clinical Benefit (NCB): Not meeting the above definition for CB

These exploratory analyses will be conducted based on RECIST v1.1, TVS and mRECIST from IRR, and based on investigator assessment per RECIST v1.1.

9.10.7.3. Tumor Size Over Time

Observed value, change from baseline and percent change from baseline in tumor size will be summarized over time.

The largest percentage change (in absolute value) in tumor size, up to week 25 and through all available follow-up, will be summarized in waterfall plots for RECIST v1.1.

9.10.7.4. Tissue Damage Score

Observed value and change from baseline in TDS components will be summarized over time.

9.10.7.5. Range of Motion Over Time

Observed value and change from baseline in active ROM of the affected joint, relative to a reference standard, will be summarized over time, including Week 49 as a timepoint of interest.

Observed value and change from baseline in active ROM of the affected joint, relative to the contralateral joint, will be summarized over time, including Week 25 and Week 49 as timepoints of interest.

Observed value and change from baseline in passive ROM of the affected joint, relative to a reference standard and to the contralateral joint, will be summarized over time, including Week 25 and Week 49 as timepoints of interest.

9.10.7.6. Clinician and Patient Global Impression of Change and Severity

Clinician Global Impression of Change (CGIC) and of Severity (CGIS) as well as Patient Global Impression of Change (PGIC) and of Severity (PGIS) will be summarized over time, including Week 13, Week 25 and Week 49 as timepoints of interest.

- Observed value will be summarized for CGIC and PGIC, as well as response in CGIC and PGIC, defined as an answer of “Very Much Improved” or “Much Improved”
- Observed value and change from baseline will be summarized for CGIS and PGIS.

9.10.7.7. Brief Pain Inventory

The analysis performed for BPI-30 response described in Section 9.10.3.6 will be replicated using BPI average pain NRS.

Observed value and change from baseline in BPI worst pain NRS and average pain NRS will be summarized over time.

9.10.7.8. Burden of Side Effects – GP5 from FACT-G

GP5 "burden-of-side-effects" question from the Functional Assessment of Cancer Therapy-General (FACT-G) will be summarized over time, including Week 25 as a timepoint of interest. This will include count and proportion of participants burdened by side effects defined as those with a response of 3 ("quite a bit") or 4 ("very much").

9.10.7.9. Worst Stiffness NRS Over Time

Observed value and change from baseline in worst stiffness NRS will be summarized over time.

9.10.7.10. EQ-VAS Over Time

Observed value and change from baseline in EQ-VAS will be summarized over time.

9.10.7.11. PROMIS-PF Over Time

Observed value and change from baseline in PROMIS-PF will be summarized over time.

9.10.8. Other Exploratory Analyses

Pharmacokinetic concentration data will be summarized over time and a listing will be provided.

Other than basic data summaries, the analyses for the following items will be described in documents separate from the CSR.

Pharmacokinetics

- Correlation of PK with efficacy and/or safety

Pharmacodynamics

- Effects of pharmacodynamics in relation to safety or efficacy

Pharmacogenomics

- Germline polymorphisms in genes involved in the metabolism or disposition of vimseltinib or related to safety and/or efficacy

9.11. Safety Analysis

All safety analyses will be conducted on the SAF. If there are both local and central assessments in the same nominal visit, then the central assessment will be used for analyses. The safety data will be summarized following two different approaches. The first approach will be limited to data from Part 1 and will focus on comparisons between participants who received vimseltinib versus participants who received placebo during this period. The second approach will combine data from both Part 1 and Part 2 (including extension period) and will focus on the observations and assessments while participants are receiving vimseltinib. Therefore, in this second approach, participants randomized to Placebo in Part 1 who crossed over to vimseltinib in Part 2, will have as their reference day the first dose date of vimseltinib in Part 2, and their data from Part 1 while receiving placebo will be excluded.

9.11.1. Extent of Treatment Exposure

Descriptive summaries will be provided by treatment group for the following variables for the SAF:

- Treatment duration (weeks), calculated as (date of last dose − date of first dose + 1)/7
- Number of cycles initiated
- Total dose received, defined as the sum of the actual doses (mg) administered
- Total planned dose (mg), defined as the sum of the prescribed doses (mg)
- Average weekly dose (mg/week), defined as total dose received (mg) / treatment duration (weeks)
- Relative dose intensity (%), defined as total dose received (mg) / total planned dose (mg) x 100
- Number of participants with drug withdrawn
- Number of participants with the following situations:
- Number of participants with dose increase
- Number of participants with dose reduction
- Number of participants with dose interruption
- Number of participants with missed dose

Treatment exposure will be listed for the SAF.

9.11.2. Adverse Event

Adverse events (AEs) will be coded using MedDRA version 26.0 (or higher). The intensity of AE will be documented using the National Cancer Institute (NCI) Common Terminology Criteria for Adverse Event (CTCAE) version 5.0.

A treatment-emergent AE (TEAE) is defined as any AE that occurs or worsens after administration of the first dose of study drug and through 30 days after the last dose of study drug or the day before the start of new antitumor therapy. Drug-related AEs reported after 30 days following the last dose of study drug will also be considered as treatment-emergent.

Drug related AEs include AEs reported as ‘POSSIBLY RELATED’ and ‘RELATED’.

Summary statistics (frequency and percentages of participants) will be presented by SOC and PT by treatment group for following:

- Overall summary of TEAEs
- Any TEAE
- Any nonserious TEAE
- Any Grade 3/4 TEAE
- Any serious TEAE

- Any drug-related TEAE
- Any Grade 3/4 drug-related TEAE
- Any drug-related serious TEAE
- Any TEAE leading to dose modification (including dose reduction and interruption)
- Any drug-related TEAE leading to dose modification (including dose reduction and interruption)
- Any TEAE leading to treatment discontinuation
- Any drug-related TEAEs leading to treatment discontinuation
- Any TEAE leading to death

The following will be tabulated by SOC, PT, and maximum severity grade:

- TEAEs
- Serious TEAEs
- Drug-related TEAEs
- Drug-related serious TEAEs

The following will be presented by PT:

- TEAEs
- TEAEs of Grade 3/4
- Serious TEAEs
- Drug-related TEAEs
- Drug-related TEAEs of Grade 3/4
- Drug-related serious TEAEs

If a SOC or PT was reported more than once for a participant, the participant would only be counted once in the incidence for that SOC or PT. For participants experiencing the same PT at multiple severity levels, the occurrence of the AEs with the maximum severity will be counted in the analysis of incidence by severity. For participants experiencing the same PT at multiple relationship levels, the occurrence of the AEs with the strongest relationship to study drug will be counted in the analysis of incidence by relationship to study drug.

The following listings will be provided for the SAF:

- All AEs (including AEs occurring prior and after the TEAE period)
- Serious Adverse Events
- Death
- AEs leading to study drug discontinuation
- AEs leading to dose modification (reduction or interruption)

9.11.3. Clinical Laboratory Tests

The following laboratory parameters will be summarized:

Hematology	Serum Chemistry	Coagulation	Urinalysis (dipstick)[a]
Hematocrit Hemoglobin Platelet count Red blood cell count WBC with differential	Alanine aminotransferase Albumin Alkaline phosphatase Aspartate aminotransferase Bicarbonate Calcium Cholesterol Chloride Creatinine Gamma-glutamyl transferase Glucose Potassium Sodium Magnesium Phosphorus Total bilirubin Conjugated (direct) bilirubin Unconjugated (indirect) bilirubin[b] Total protein Triglycerides Urea (blood urea nitrogen)	Activated partial thromboplastin time Prothrombin time International Normalized Ratio	Hemoglobin Glucose Ketone Protein Specific gravity

[a] An additional microscopic examination will be performed if any dipstick analytes are 2+ or higher. A participant should undergo further assessment (eg, a 24-hour urine collection) if the protein in analyte is >=2+
[b] Will be calculated by central laboratory.

Descriptive statistics (number of observations, mean, median, standard deviation, minimum, and maximum values) will be presented for clinical laboratory tests. Changes from baseline (including baseline value) by treatment group will be summarized at each scheduled visit. Graphical presentations may also be generated for specific laboratory parameters by visit and treatment group to investigate the trend over time and outliers in the data.

Abnormal laboratory results will be graded with NCI CTCAE version 5.0, if applicable. A shift table, presenting the 2-way frequency tabulation for baseline and the worst post-baseline values according to the NCI CTCAE grade, will be provided by treatment group for hematology and serum chemistry.

A Hy's law summary table and plot will be provided. A plot of alkaline phosphatase vs. total bilirubin will be provided.

A comprehensive listing of all hematology, chemistry (including CPK added in Protocol Amendment 3), coagulation, urinalysis, and pregnancy results will be provided for the SAF. The listing will include the test result, the normal range, change from baseline, and CTCAE grade. Separate listings for hematology and chemistry presenting clinically significant abnormalities or abnormalities of CTCAE grade 3 or 4 will be provided.

9.11.4. Electrocardiogram (ECG)

The following ECG measurements will be summarized:

- Heart rate
- RR interval (msec)
- QRS interval (msec)
- QT interval (msec)
- Fridericia corrected QT (QTcF) interval [$QTcF = QT/\sqrt[3]{RR/1000}$]

An overall investigator assessment of ECG will be provided (categories "normal", "abnormal, not clinically significant" and "abnormal, clinically significant"). The following categories will also be reported for QTcF:

- QTcF > 450 msec
- QTcF > 480 msec
- QTcF > 500 msec
- Change from baseline in QTcF > 30 msec
- Change from baseline in QTcF > 60 msec

The ECG measurements and changes from baseline in ECG will be summarized by treatment, visit with descriptive statistics for the SAF.

All ECG data will also be presented in listings.

9.11.5. Vital Signs

The following vital signs will be summarized:

- Systolic and diastolic blood pressure (mmHg)
- Heart rate (beats/minute)
- Respiration rate (breaths/minute)
- Body temperature (°C)
- Weight (kg)

Observed values and changes from baseline in vital signs will be summarized by treatment and visit using descriptive statistics for the SAF. By-participant listings will also be presented.

10. INTERIM ANALYSIS

No interim analysis will be performed for this study.

11. CHANGES TO PLANNED ANALYSIS IN PROTOCOL

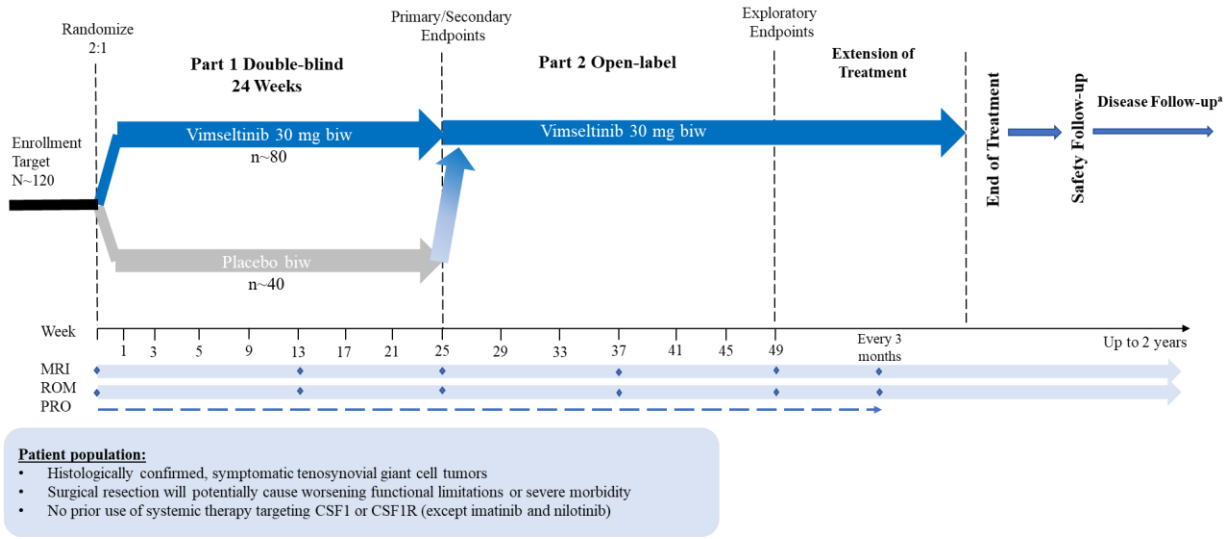
There are no major changes from the protocol planned analyses.

12. REFERENCES

1. Tap WD, Gelderblom H, Palmerini E, et al. Pexidartinib versus placebo for advanced tenosynovial giant cell tumour (ENLIVEN): a randomised phase 3 trial. Lancet. 2019; 394:478-487.
2. Cassier PA, Gelderblom H, Stacchiotti S, et al. Efficacy of imatinib mesylate for the treatment of locally advanced and/or metastatic tenosynovial giant cell tumor/pigmented villonodular synovitis. Cancer. 2012;118: 1649-1655.
3. Calculating total daily dose of opioids for safer dosage. US Centers for Disease Control and Prevention. Available at: https://www.cdc.gov/opioids/providers/prescribing/pdf/calculating-total-daily-dose.pdfhttps://www.cdc.gov/drugoverdose/pdf/calculating_total_daily_dose-a.pdf (Accessed on April 20, 2023).
4. WHO GUIDELINES FOR THE PHARMALOGICAL AND RADIOTHERAPEUTIC MANAGEMENT OF CANCER PAIN IN ADULTS AND ADOLESCENTS. Available at: https://apps.who.int/iris/bitstream/handle/10665/279700/9789241550390-eng.pdf
5. Defacto Long-term Opioid Therapy for Non-Cancer Pain (CONSORT). Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3286630/
6. FACIT.org. Overview. http://www.facit.org/FACITOrg/Overview.

APPENDIX 1.

Figure 1: Study Schema



Abbreviations: biw=twice weekly; CSF1=colony-stimulating factor 1; CSF1R=colony-stimulating factor 1 receptor; MRI=magnetic resonance imaging; PRO=patient-reported outcome; ROM=range of motion.

[a] Disease follow-up will be performed for up to 2 years.

Note: Participants will be eligible to receive study drug until radiological confirmation of disease progression (refer to protocol Section 10.5), unacceptable toxicity, withdrawal by participant, physician's decision, or commercial availability of vimseltinib and for as long as vimseltinib is being developed to support the indication, and continuation of treatment does not conflict with the Sponsor's right to terminate the study.

Signature Page for DCC-3014-03-001 SAP Final 2.0 09AUG2023

Approval	PPD Deciphera Pharmaceuticals, LLC 18-Aug-2023 16:22:49 GMT+0000

Approval	PPD Deciphera Pharmaceuticals, LLC 18-Aug-2023 16:23:36 GMT+0000

Approval	PPD Deciphera Pharmaceuticals, LLC 18-Aug-2023 16:47:15 GMT+0000

Veeva Vault RIM Document ID Number DCC-CLIN-001500