Statistical Analysis Plan

A Phase 2, Open-Label Extension, Efficacy and Safety Study of a RARγ-Specific Agonist (Palovarotene) in the Treatment of Preosseous Flare-ups in Subjects with Fibrodysplasia Ossificans Progressiva (FOP)

Study Number: PVO-1A-202 Part A

Original Protocol: 04 June 2014
Amendment 1: 30 September 2014
Amendment 2: 22 June 2015
Amendment 3: 10 March 2016

Statistical Analysis Plan Version: Final v2
Statistical Analysis Plan Date: 03 April 2017

Clementia Pharmaceuticals Inc.
4150 Saint Catherine Street West, Suite 550
Montreal, Quebec, H3Z 2Y5
Canada

Statistical Analysis Plan Signature Page

A Phase 2, Open-Label Extension, Efficacy and Safety Study of a RARγ-Specific Agonist (Palovarotene) in the Treatment of Preosseous Flare-ups in Subjects with Fibrodysplasia Ossificans Progressiva (FOP)

Protocol Number: PVO-1A-202 Part A

Signature of Approval for Statistical Analysis Plan (Version 2: 03 April 2017)

Prepared by:

MEDPACE INC.
PI, MS
Biostatistician

NAME: PI

SIGNATURE: _____ DATE: PI

Approved by:

CLEMENTIA PHARMACEUTICALS INC.
PI, MD
PI

NAME: PI

SIGNATURE: _____ DATE: PI

CLEMENTIA PHARMACEUTICALS INC.
PI
Project Statistician PI

NAME: PI

SIGNATURE: _____ DATE: Date: 2017.04.07 07:20:21 -04'00'

MEDPACE INC.
PI, MD
Medical Monitor

NAME: PI

SIGNATURE: PI DATE: PI

Table of Contents

List of Tables

List of Abbreviations

Abbreviation	Definition
AE	adverse event
CAJIS	Cumulative Analogue Joint Involvement Scale
C-SSRS	Columbia-Suicide Severity Rating Scale
CT	computed tomography
CTCAE	Common Terminology Criteria for Adverse Events
ECG	electrocardiogram
FAS	Full Analysis Set
FOP	Fibrodysplasia Ossificans Progressiva
FOP-PFQ	FOP-Physical Function Questionnaire
FPS-R	Faces Pain Scale-Revised
HO	heterotopic ossification
MedDRA	Medical Dictionary for Regulatory Activities
MRI	magnetic resonance imaging
NRS	numeric rating scales
PCS	potentially clinically significant
PROMIS	Patient-Reported Outcomes Measurement Information System
RARγ	retinoic acid receptor gamma
ROM	range of motion
SAE	serious adverse event
SAP	statistical analysis plan
ULN	upper limit of normal
US	ultrasound

1 Introduction

This Statistical Analysis Plan (SAP) describes the detailed analyses to be performed for Study PVO-1A-202 Part A. These analyses are based on Amendment 2 of the protocol dated 22 June 2015 and Amendment 3 of the protocol dated 10 March 2016. The analyses to be performed for Study PVO-1A-202 Part B will be described in a separate SAP. Note that analyses related to the Follow-up Component objectives as described in Section 2.1 will be addressed in the Part B SAP.

2 Overview

2.1 Study Objectives

2.1.1 Primary Objective

The primary objectives are:

- To evaluate the long-term safety and efficacy of prior palovarotene treatment in FOP subjects who completed Study PVO-1A-201 (**Follow-up Component**; up to 12 months).
- To evaluate the safety and efficacy of palovarotene in FOP subjects experiencing up to two, new, distinct flare-ups (**Flare-up Component**; up to 36 months). Efficacy will be based on the ability of palovarotene to prevent heterotopic ossification (HO) at the new, distinct flare-up site as assessed by low-dose computed tomography (CT) scan (or plain radiographs for subjects unable to undergo CT scan).

2.1.2 Secondary Objectives

The secondary objectives for the **Follow-up Component** are:

- To evaluate the amount of HO at the original flare-up site by plain radiograph.
- To evaluate active range of motion (ROM) by goniometer at the original flare-up site.
- To evaluate ROM as assessed by the Cumulative Analogue Joint Involvement Scale for FOP (CAJIS).
- To evaluate the subject and Investigator global assessment of movement at the original flare-up site.
- To evaluate pain and selling at the original flare-up site using numeric rating scales (NRS) or the Faces Pain Scale-Revised (FPS-R) in subjects under 8 years of age.
- To evaluate physical function using age-appropriate forms of the FOP-Physical Function Questionnaire (FOP-PFQ).
- To evaluate physical and mental health using age-appropriate forms of the Patient Reported Outcomes Measurement Information System (PROMIS) Global Health Scale.
- To evaluate the use of assistive devices and adaptations for daily living by FOP subjects.
- To evaluate bone, cartilage, angiogenesis, and inflammation biomarkers; and explore correlations between changes from baseline in biomarkers and clinical efficacy.

The secondary objectives for the **Flare-up Component** are:

- To evaluate the effect of palovarotene on active ROM at the new, distinct flare-up site as assessed by goniometer.

- To evaluate the effect of palovarotene on ROM as assessed by CAJIS.
- To evaluate the effect of palovarotene by the subject and Investigator global assessment of movement at the new, distinct flare-up site.
- To evaluate the effect of palovarotene on physical function using age-appropriate forms of the FOP-PFQ.
- To evaluate the effect of palovarotene on physical and mental health using age-appropriate forms of the PROMIS Global Health Scale.
- To evaluate the effect of palovarotene on the use of assistive devices and adaptations for daily living by FOP subjects.
- To evaluate the effects of palovarotene on soft tissue swelling and cartilage formation at the new, distinct flare-up site as assessed by magnetic resonance imaging (MRI); or the effect of palovarotene on soft tissue swelling as assessed by ultrasound (US) in subjects unable to undergo MRI.
- To evaluate bone, cartilage, angiogenesis, and inflammation biomarkers; and explore correlations between changes from baseline in biomarkers and clinical efficacy.

2.2 Study Population

Approximately 40 subjects who completed Study PVO-1A-201 will be enrolled in Study PVO-1A-202 Part A.

2.3 Study Design

Study PVO-1A-202 Part A is a Phase 2, multicenter, event-based, open-label extension of Study PVO-1A-201. This open-label extension will give subjects who completed Study PVO-1A-201 the opportunity to be followed for up to an additional 36 months, and receive open-label palovarotene should they experience up to two, new, distinct flare-ups.

All enrolled subjects will participate in the **Follow-up Component** of the study. The Follow-up Component will include visits at Study Screening and Study Months 6, 12, 18, 24, 30, and 36. Study Day 1 will be the date of informed consent. Visit windows are +/2 weeks. Assessments related to flare-ups during the Follow-up Component will be evaluated at the original flare-up site from Study PVO-1A-201.

Subjects who experience an eligible flare-up will participate in the **Flare-up Component** of the study. The Flare-up Component will include visits at Flare-up Screening and Flare-up Weeks 2, 4, 6, 9, and 12. Flare-up Day 1 will be the date of first dose of study drug for the flare-up. Visit windows are +/- 3 days. Assessments related to flare-ups during the Flare-up Component will be evaluated at new flare-up sites from Study PVO-1A-202 Part A. Note that subjects may have more than one unique new flare-up and that each flare-up will be assessed and analyzed independently.

2.4 Study Treatment

Subjects who experience an eligible flare-up will receive palovarotene 10 mg for 14 days followed by palovarotene 5 mg for 28 days or exposure equivalent doses based on subject

weight. Weight-adjusted doses of palovarotene estimated to achieve exposures equivalent to 10 mg and 5 mg are:

Weight Range Category	10-mg Equivalent	5-mg Equivalent	2.5-mg Equivalent *
20 to <40 kg	6 mg	3 mg	1.5 mg
40 to <60 kg	8 mg	4 mg	2 mg
≥60 kg	10 mg	5 mg	2.5 mg

* In the event of dose de-escalation.

In the event the subject requires dose de-escalation due to an intolerable side effect in Study PVO-1A-201, then the dose the subject will receive for a subsequent flare-up will be determined by the Investigator and the Medical Monitor. Dose reduction to one half of the palovarotene dose may also occur should the subject experience an adverse event that may otherwise result in study drug discontinuation; if a subject is already receiving the lowest possible dose, then study drug will be discontinued.

Subjects who do not experience an eligible flare-up will not be treated.

2.5 Randomization and Blinding

This is an open-label study and does not involve randomization or blinding.

3 General

Categorical data will be summarized with counts and percentages. Percentages will be calculated based on the number of non-missing values and will be reported to one decimal place.

Continuous data will be summarized with descriptive statistics including the number of non-missing values, mean, standard deviation, standard error, median, minimum, and maximum. Minimum and maximum will be reported to the same precision as the raw values, mean and median will be reported to one additional decimal place, and standard deviation and standard error will be reported to two additional decimal places. Some discretion may be applied.

All data will be listed. When listings will be provided instead of summaries or require additional description, they are described in this document.

3.1 Adjustments for Covariates

Baseline covariates may be added to statistical models as the data allow. The sample size of this study may preclude the inclusion of covariate effects in some instances.

3.2 Handling of Dropouts or Missing Data

As there is no effective treatment for these subjects, it is assumed they will be highly motivated to adhere to the protocol and treatment visits. Thus, it is expected that there will be limited missing data, and most subjects will complete the study. The primary analysis will be performed using observed data only. Missing data will not be replaced.

3.3 Interim Analyses

No interim analysis is planned.

3.4 Multicenter Studies

A review of by-center effects will be performed in the context of data listing review. However, the sample size precludes formal assessment of by-center effects.

3.5 Multiple Comparison/Multiplicity

As this is an extension of an initial proof-of-concept study, no correction for multiple testing is planned.

3.6 Examination of Subgroups

The sample size in this study is not sufficient to allow for rigorous assessment of outcomes by subgroups, and the study is not powered for subgroup assessment. However, exploratory assessments of the outcomes may be performed by a number of extrinsic or intrinsic factors, including age of the subject, location of the flare-up, number of flare-up treatments, or other variables as appropriate.

3.7 Study Day and Flare-up Study Day

Study Day will be determined from the date of informed consent (Study Day 1); Study Day -1 will be the day immediately prior to Study Day 1.

Flare-up Study Day will be determined from the date of first dose of a treated flare-up (Flare-up Study Day 1); Flare-up Study Day -1 will be the day immediately prior to Flare-up Study Day 1.

4 Analysis Populations

There will be five populations defined for the analysis.

4.1 Enrolled Population

The Enrolled Population will include all subjects enrolled in Study PVO-1A-202 Part A. Analyses based on the Enrolled Population will generally be summarized by treatment (Untreated, Palovarotene 10 mg/5 mg) and in total. The Enrolled Population will be the primary population for demographic and baseline summaries and for the analysis of adverse events.

4.2 Untreated Population

The Untreated Population will include all subjects who did not take any palovarotene in Study PVO-1A-202 Part A.

4.3 Treated Population

The Treated Population will include all subjects who took at least one dose of palovarotene in Study PVO-1A-202 Part A. Analyses based on the Treated Population will generally be summarized in total. Note that data collected for flare-ups may use each flare-up as the unit of

analysis rather than each subject. The Treated Population will be the primary population for analysis of safety assessments performed at Flare-up Component visits.

4.4 Efficacy Population

The Efficacy Population will include all subjects in the Treated Population who had an evaluable Week 6 or Week 12 image (CT scan or x-ray). Analyses based on the Efficacy Population will generally be summarized in total. Note that data collected for flare-ups may use each flare-up as the unit of analysis rather than each subject. The Efficacy Population will be the primary population for analysis of efficacy assessments performed at Flare-up Component visits.

4.5 Per Protocol Population

The Per Protocol (PP) Population will include all subjects in the Efficacy Population with no major protocol deviations that may impact the efficacy assessment. Exclusions from the PP Population will be identified prior to database lock. Analyses based on the PP Population will generally be summarized in total. Note that data collected for new unique flare-ups may use each flare-up as the unit of analysis rather than each subject. The PP Population will be a secondary population for analysis of selected efficacy assessments performed at Flare-up Component visits.

5 Treatment Period and Non-Treatment Period

The Treatment Period is defined as the period of days during the Flare-up Component(s) from Flare-up Study Day 1 through the date of completion of the Flare-up Component (ie, Flare-up Study Day 84). The Treatment Period includes all days during Flare-up Components for subjects in the Treated Population.

The Non-Treatment Period is defined as the period of days outside of the Treatment Period. The Non-Treatment Period includes all days in the study for subjects in the Untreated Population and days before, after, or in between Flare-up Components for subjects in the Treated Population.

6 Subject Data

6.1 Subject Disposition and Analysis Populations

The following will be summarized with counts and percentages of subjects in the Enrolled Population:

- Subjects who were enrolled (Enrolled Population).
- Subjects who did not take any palovarotene (Untreated Population).
- Subjects who took at least one dose of palovarotene (Treated Population).
- Subjects in the Treated Population who had an evaluable Week 6 or Week 12 image (CT scan or x-ray) (Efficacy Population).
- Subjects in the Efficacy Population who did not have major protocol deviations that may impact the efficacy assessment (PP Population).
- Subjects who discontinued from Part A early along with reasons for discontinuation.
- Subjects who completed Part A.

6.2 Eligibility Criteria

Eligibility criteria (inclusion/exclusion) will be listed.

6.3 Protocol Deviations

Major protocol deviations will be identified from the clinical data prior to freezing the database for Part A and may include, but are not limited to departures from the inclusion/exclusion criteria, non-compliance with investigational product, use of restricted concomitant medications, and non-compliance with study procedures. Major protocol deviations will be summarized for the Enrolled Population.

6.4 Demographic and Other Baseline Characteristics

Demographic and other baseline characteristics such as age, sex, race, ethnicity, height, and weight will be summarized for the Enrolled, Efficacy, and PP Populations.

6.5 Medical History

Medical history will be coded using the Medical Dictionary for Regulatory Activities (MedDRA Version 17.0). The incidence of medical history conditions will be summarized by MedDRA body system and preferred term for the Enrolled Population.

6.6 Flare-ups Starting During Part A

The number of treated flare-ups will be summarized with counts and percentages of subjects in the Treated, Efficacy, and PP Populations.

Treated flare-up data as reported by the subject and the Investigator including symptoms, probable causes, and location(s) will be summarized for the Treated, Efficacy, and PP Populations.

Data for untreated flare-ups will be listed.

6.7 Concomitant Medications

Concomitant medications will be coded using the WHO Drug Dictionary (March 2014). Glucocorticoids taken for a flare-up are defined as courses of glucocorticoids that start within 9 days after the start of the flare-up and are specified as taken for the current flare-up on the electronic case report form (eCRF). Courses of glucocorticoids are defined as days of continuous treatment allowing for up to a 2-day interruption.

The use of the following concomitant medications will be summarized by anatomic therapeutic class (ATC) term and preferred term for the Enrolled Population:

- Prior medications that were on-going at Study Day 1 excluding glucocorticoids. (Medications with start date before Study Day 1 and end date on or after Study Day 1.)
- New-onset medications during the Non-Treatment Period excluding glucocorticoids. (Medications with start date on or after Study Day 1 and not during the Treatment Period.)

- Glucocorticoids taken for flare-ups during the Non-Treatment Period.
- New-onset glucocorticoids during the Non-Treatment Period. (Courses of glucocorticoids with start date on or after Study Day 1 and not during the Treatment Period, and not taken for treated flare-up.)

The use of the following concomitant medications will be summarized by anatomic therapeutic class (ATC) term and preferred term for the Treated Population:

- Prior medications that were on-going at Flare-up Study Day 1 excluding glucocorticoids. (Medications with start date before Flare-up Study Day 1 and end date on or after Flare-up Study Day 1.)
- New-onset medications during the Treatment Period excluding glucocorticoids. (Medications with start date on or after Flare-up Study Day 1 and on or before the date of completion of the Flare-up Component, ie, Day 84.)
- Glucocorticoids taken for treated flare-ups.
- New-onset glucocorticoids during the Treatment Period. (Courses of glucocorticoids not taken for the treated flare-up with start date on or after Flare-up Study Day 1 and on or before the date of completion of the Flare-up Component.)

For the purposes of analysis, incomplete medication start and stop dates will be imputed. If a medication start date is incomplete, January will be imputed for missing month and the first day of the month will be imputed for missing day. If a medication stop date is incomplete, December will be imputed for missing month and the last day of the month will be imputed for missing day. If the imputed medication stop date is after the date of study completion, the date of study completion will be used instead.

A listing of courses of glucocorticoids will include additional derived data such as start and stop dates with associated Flare-up Day, start day of the glucocorticoid relative to the start of the flare-up, duration of use, and total amount taken.

6.8 Duration of Exposure

Duration of exposure to study drug for each treated flare-up will be calculated from the date of first dose for the flare-up to the date of last dose for the flare-up. Duration of exposure to study drug for each subject will be calculated as the number of days exposed to study drug, ie, the sum of the exposure for each flare-up.

Days of exposure to study drug for flare-ups and subjects will be summarized for the Treated Population with descriptive statistics. Weeks of exposure to study drug will be summarized with counts and percentages of flare-ups in the Treated Population according to the following categories:

- ≤ 2 weeks
- $>2 – 4$ weeks
- $>4 – 6$ weeks

Note that flare-ups with exposure over 42 days (due to windowing) will be included in the last category.

Dose reductions (no dose reductions, any dose reductions) will be summarized with counts and percentages of flare-ups and subjects in the Treated Population.

The number of days after flare-up onset to the date of first dose of study drug for the flare-up will be summarized for the Treated Population with descriptive statistics.

6.9 Duration of Participation in Part A

Duration of participation in Part A will be calculated from the date of informed consent to the date of completion of Part A/discontinuation. Months of participation will be calculated as days of participation / 30.5 and summarized for the Enrolled Population with descriptive statistics. Months of participation will also be summarized with counts and percentages of subjects in the Enrolled Population according to the following categories:

- ≤ 6 months
- >6 – 12 months
- >12 – 18 months
- >18 – 24 months
- >24 – 30 months
- >30 – 36 months

6.10 Compliance

Compliance to study drug for each treated flare-up will be calculated based on the number of actual capsules taken and the expected number of capsules as specified in the protocol for 42 days of dosing. The number of actual capsules will be calculated as the number of capsules dispensed – the number of capsules returned – the number of capsules lost. For example, a subject who took 10 capsules and was expected to take 42 capsules would be 10/42 = 23.8% compliant.

Percent compliance will be summarized for the Treated Population with descriptive statistics. Categories of percent compliance (<80%, ≥80%) will be summarized for the Treated Population with counts and percentages of flare-ups.

7 Efficacy Data

Efficacy assessments at Flare-up Component visits will generally be presented in total for the Efficacy Population. The primary efficacy analysis and selected other efficacy analyses will also be presented for the PP Population. Efficacy assessments at Follow-up Component visits will be listed.

Imaging Reads

Imaging data includes results from CT scan, x-ray, MRI, and US. Results will be read and interpreted using two different review processes (Primary Read and Global Read) as described in the study protocol.

The Primary Read review will be performed by two independent musculoskeletal radiologists during the study as images are made available. If there is sufficient agreement between the

independent reviews, both will be used for analysis. If there is insufficient agreement between the independent reviews, an adjudication review will be performed and used for analysis. Imaging procedures performed (yes, yes but not evaluable, no) will be summarized at scheduled visits with counts and percentages of subjects or flare-ups as appropriate.

The Global Read review will be performed at the end of the study by a global review team with access to all imaging modalities concurrently, as well as, selected clinical data for an individual subject. The main objective of the Global Read review is to reduce any potential variability created by the Primary Read process. More specifically, the Global Read review will resolve some of the structured independence of the Primary Read review where cases were evaluated on a time point-by-time point basis for each flare-up location, and each imaging modality was interpreted in isolation. The Global Read review will include assessments of the presence of HO including whether there is a new and/or enlarged lesion and the presence and severity of edema.

The outcomes from the pre-planned Primary Read review will be used to determine the primary conclusions for this study. However, outcomes from the Global Read review will be used to further explicate the palovarotene treatment effect, and in developing and designing future clinical trials for the FOP program.

7.1 Primary Efficacy

CT scan (or x-ray for subjects unable to undergo CT scan) will be performed at:

- Screening/Baseline, Week 6, and Week 12 (Flare-up Component).

The primary efficacy analysis is the proportion of flare-ups with no new HO at Week 12 as assessed by CT scan (or by x-ray for subjects unable to undergo CT scan). Note that "no new HO at Week 12" means no new HO at Week 12 compared to baseline.

The presence of HO at baseline and the incidence of no new HO and new HO at post-baseline visits will be summarized with counts and percentages of flare-ups.

Similar analyses will be performed for subgroups defined by baseline edema severity (no edema, any edema, mild edema, moderate edema, severe edema, and moderate or severe edema [higher-intensity flare-ups]). See Section 7.2.4 for a description of edema determination.

7.2 Secondary/Exploratory Efficacy

7.2.1 Volume of HO and HO Grades

Volume of HO and HO grades (defined below) will be determined from CT scan at visits as specified in Section 7.1 for primary efficacy.

The volume of HO at baseline will be calculated as the sum of the volumes for each lesion. The volume of *new* HO at post-baseline visits will be calculated as follows:

1. If there is no new HO, impute new volume = 0. (NVOL = 0)
2. If there is new HO

a. Subtract the volume of baseline HO from the volume of post-baseline HO for each lesion. (new volume at each lesion = $NVOL_1$, $NVOL_2$, … $NVOL_L$ where L is the number of lesions)
b. If the volume of post-baseline HO < the volume of baseline HO for a given lesion i, impute new volume = 0. ($NVOL_i = 0$)
c. Calculate the total volume of new HO. ($NVOL = NVOL_1 + NVOL_2 + \ldots + NVOL_L$)

If an adjudication review is performed, the volume from the adjudication review will be used for analysis. Otherwise, the average of the volumes from the independent reviews will be used for analysis.

Volume of HO at baseline and volume of new HO at post-baseline visits will be summarized with descriptive statistics. Similar summaries will be provided for the subgroup of flare-ups with new HO at any post-baseline visit.

HO lesions will be graded according to the following scale:

- Grade 0 – Imputed when there is no new HO
- Grade 1 – Fluid attenuation without evidence of calcification at CT
- Grade 2 – Calcification of soft tissues without evidence of bone formation
- Grade 3 – Immature bone formation
- Grade 4 – Mature bone with cortical differentiation

The HO grade at baseline will be calculated as the maximum HO grade across lesions. The *new* HO grade at post-baseline visits will be calculated as follows:

1. If there is no new HO, impute new HO grade = 0.
2. If there is new HO
    a. If there was no baseline HO, calculate the maximum HO grade across lesions.
    b. If there was baseline HO, calculate the maximum HO grade across lesions where there was no baseline HO.

HO grade at baseline and new HO grade at post-baseline visits will be summarized with counts and percentages of flare-ups.

7.2.2 Area of HO and HO Scores

Area of HO and HO scores will be determined from x-ray. Subjects unable to undergo CT scan will have x-ray performed at:

- Screening/Baseline, Week 6, and Week 12 (Flare-up Component).

Area of HO at baseline will be calculated as the sum of the areas for each lesion. The area of *new* HO at post-baseline visits will be calculated using both anterior/posterior (AP) and lateral views as follows:

1. If there is no new HO, impute new area = 0. (NAREA = 0)
2. If there is new HO

a. Subtract the area of baseline HO from the area of post-baseline HO for each lesion. ($NAREA_1$, $NAREA_2$, … $NAREA_L$ where L is the number of lesions)
b. If the area of post-baseline HO < the area of baseline HO for a given lesion i, impute new area = 0. ($NAREA_i = 0$)
c. Calculate the total area of new HO. ($NAREA = NAREA_1 + NAREA_2 + … + NAREA_i$)

For each review, the maximum area from the AP or lateral view will be used for analysis.

If an adjudication review is performed, the area from the adjudication review will be used for analysis. Otherwise, the average of the areas from the independent reviews will be used for analysis.

Area of HO at baseline and area of new HO at post-baseline visits will be listed.

HO lesions will be scored according to the following scale:

- 0 – No HO
- 1 – Single or multiple spicules (punctate) or islands (non-contiguous) of HO
- 2 – Coalescing islands or reticular complexes of bone
- 3 – Single contiguous HO having longest dimension ≤1/2 the diameter of the reference normotopic bone in any projection
- 4 – Single contiguous HO having longest dimension >1/2 but ≤1 diameter of the reference normotopic bone in any projection
- 5 – Single contiguous HO having longest dimension >1 but ≤2 diameter of the reference normotopic bone in any projection
- 6 – Single contiguous HO having longest dimension >2 diameters of the reference normotopic bone in any projection

The HO score at baseline will be calculated as the maximum HO score across lesions. The *new* HO score at post-baseline visits will be calculated as follows:

1. If there is no new HO, impute new HO score = 0.
2. If there is new HO
    a. If there was no baseline HO, calculate the maximum HO score across lesions.
    b. If there was baseline HO, calculate the maximum HO score across lesions where there was no baseline HO.

HO score at baseline and new HO score at post-baseline visits will be listed.

7.2.3 New or Enlarged HO and Severity of New HO

New HO lesions are those that did not exist at baseline. Enlarged HO lesions are lesions that existed at baseline but increased in volume (or area by x-ray) post-baseline. The incidence of new HO lesions only, enlarged HO lesions only, or both new HO and enlarged HO lesions will be summarized at each scheduled visit with counts and percentages of flare-ups.

The severity of new HO (mild, moderate, severe) will be summarized at each scheduled visit with counts and percentages of flare-ups.

7.2.4 Presence of Soft Tissue Swelling (Edema) and/or Cartilage Formation, Severity of Edema, and Volume of Edema

MRI (or US for subjects unable to undergo MRI) will be performed at:

- Screening/Baseline, Week 6, and Week 12 (Flare-up Component).

Note that US does not include an assessment of cartilage formation. If an adjudication review is performed, the severity and volume of edema from the adjudication review will be used for analysis. Otherwise, the maximum severity and the average volume of edema from the independent reviews will be used for analysis. Note also that when edema is assessed as mild with US, moderate will be imputed for analysis.

The presence of edema and the presence of cartilage formation will be summarized at each scheduled visit with counts and percentages of flare-ups.

The severity of edema (no edema, any edema, mild edema, moderate edema, severe edema, moderate or severe edema [higher intensity flare-ups]) will be summarized at each scheduled visit with counts and percentages of flare-ups.

The shift in severity of edema from baseline to each post-baseline visit will also be summarized with counts and percentages of flare-ups.

Volume of edema at each scheduled visit will be summarized with descriptive statistics.

7.2.5 Active Range of Motion

Range of motion (ROM) at the primary and secondary joints will be assessed using a goniometer at:

- Screening/Baseline, Week 6, and Week 12 (Flare-up Component).

Arcs of motion as defined in Table 1 will be calculated based on raw ROM measurements. If any of the contributing raw ROM measurements are missing, the arc of motion will not be calculated. Percent of normal arc of motion will be calculated as 100 x arc of motion / normal adult arc of motion.

Table 1. Arcs of Motion

Joint	Arc of Motion	Calculation Based on Raw ROM Measurements	Normal Adult Arc of Motion (degrees)
Hip	Extension/Flexion	Extension + Flexion	140
	Abduction/Adduction	Abduction + Adduction	60
	Total Arc of Motion	Extension + Flexion + Abduction + Adduction	200
Knee	Total Arc of Motion	Extension + Flexion	135
Shoulder	Total Arc of Motion	Flexion + Abduction-glenohumeral + Abduction-shoulder complex	490
Elbow	Extension/Flexion	Extension + Flexion	150
	Supination/Pronation	Supination + Pronation	160
	Total Arc of Motion	Extension + Flexion + Supination + Pronation	310
Wrist	Total Arc of Motion	Extension + Flexion	150
Ankle	Dorsiflexion/Plantarflexion	Dorsiflexion + Plantarflexion	70
	Subtalar Inversion/Subtalar Eversion	Subtalar Inversion + Subtalar Eversion	50
	Total Arc of Motion	Dorsiflexion + Plantarflexion + Subtalar Inversion + Subtalar Eversion	120
Cervical Spine	Extension/Flexion	Extension + Flexion	90
		Lateral	45
		Rotation	60
	Total Arc of Motion	Extension + Flexion + Lateral + Rotation	195
Jaw	Total Arc of Motion	Opening	35 mm
Lumbar Spine		Flexion	6 cm
		Extension	1 cm
	Total Arc of Motion	Flexion + Extension	7 cm

Source: Norken and White (2011)[1]

Percent of normal arc of motion values and changes from baseline will be summarized at each scheduled visit with descriptive statistics. Summaries will be provided separately for:

- Percent of normal total arc of motion at the primary joint (subjects have one total arc of motion at the primary joint)
- Percent of normal arc of motion at the primary joint by joint and arc of motion (subjects may have more than one arc of motion at the primary joint)
- Percent of normal arc of motion at the secondary joints by joint and arc of motion (subjects may have more than one arc of motion at one or more secondary joints)

Similar summaries for percent of normal total arc of motion will be provided for subgroups defined by presence of new HO at any post-baseline visit (new HO, no new HO).

7.2.6 Cumulative Analogue Joint Involvement Scale

Range of motion at the primary and secondary joints will also be assessed using the Cumulative Analogue Joint Involvement Scale (CAJIS) at:

- Screening/Baseline, Week 6, and Week 12 (Flare-up Component).

The assessments will be performed on 12 joints (shoulder, elbow, wrist, hip, knee, and ankle on both the right and left sides), and three body regions (jaw, cervical spine [neck], and thoracic/lumbar spine). Each joint/region will be scored as: 0 = uninvolved; 1 = partially involved; and 2 = ankylosed/completely involved.

The CAJIS Total score will be calculated as the sum of the scores of all joints/regions and ranges from 0 (no involvement) to 30 (maximally involved).

The CAJIS Upper Extremities subscore will be calculated as the sum of the scores from six joints (shoulder, elbow, and wrist on both the right and left sides) and one region (cervical spine [neck]) and ranges from 0 (no involvement) to 14 (maximally involved).

The CAJIS Mobility subscore will be calculated as the sum of scores from six joints (hip, knee, and ankle on both the right and left sides) and ranges from 0 (no involvement) to 12 (maximally involved).

The CAJIS Total score, CAJIS Upper Extremity subscore, and CAJIS Mobility subscore and changes from baseline will be listed.

7.2.7 Global Assessment of Movement

Flare-up movement compared to baseline will be assessed by the subject (or parent proxy) and Investigator at:

- Week 6 and Week 12 (Flare-up Component).

Assessment results will be listed.

7.2.8 FOP-Physical Function Questionnaire (FOP-PFQ)

Age-appropriate forms of the FOP-PFQ will be administered to subjects (or parent proxy) at:

- Screening/Baseline, Week 2, Week 4, Week 6, Week 9, and Week 12 (Flare-up Component).

The adult form will be administered to subjects 15 years and older. For subjects between 8 and 14 years of age, both the pediatric self-completed and the pediatric proxy-completed forms will be administered. The proxy-completed form will be used for analyses unless only the self-completed form is available.

The FOP-PFQ consists of 28 questions on the adult form and 26 questions on the pediatric form scored on a scale from 1 to 5, with lower scores indicating more difficulty. The questions are described in detail in Appendices 7A, 7B, and 7C of the protocol.

The Total score will be calculated as:

- The sum of the scores from each question and ranges from 28x1 = 28 to 28x5 = 140 for the adult form.
- The sum of the scores from each question and ranges from 26x1 = 26 to 26x5 = 130 for the pediatric form.

The Upper Extremities subscore will be calculated as:

- The sum of the scores from 15 questions (questions 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 25, and 26) and ranges from 15x1 = 15 to 15x5 = 75 for the adult form.
- The sum of the scores from 18 questions (questions 1, 2, 6, 7, 8, 9, 10, 11, 16, 17, 18, 19, 21, 22, 23, 24, 25, and 26) and ranges from 18x1 = 18 to 18x5 = 90 for the pediatric form.

The Mobility subscore will be calculated as:

- The sum of the scores from 13 questions (questions 13, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 27, and 28) and ranges from 13x1 = 13 to 13x5 = 65 for the adult form.
- The sum of the scores from 8 questions (questions 3, 4, 5, 12, 13, 14, 15, and 20) and ranges from 8x1 = 8 to 8x5 = 40 for the pediatric form.

The Total score, Upper Extremities subscore, and Mobility subscore are defined based on information from PROMIS and Hays et al (2013)[2] for the adult form and a medical review of the questions for the pediatric form.

If a subject is missing some (but not more than 20%) of the contributing question scores, the Total score, Upper Extremities subscore, or Mobility subscore will be calculated as the average observed score multiplied by the number of expected question scores. For example, the Total score would be calculated as the average of the non-missing scores x 28 for the adult form. The determination for sufficient non-missing scores will be made independently for the Total score, Upper Extremities subscore, and Mobility subscore.

As the analysis for FOP-PFQ will be performed across all subjects (adult and pediatric) and the number of contributing questions differs, the scores will be transformed to reflect a percentage of worst score. The percentage of worst score ranges from 0% to 100% with 0% indicating the best possible function and 100% indicating the worst possible function. Table 2 illustrates some sample derivations of the percentage of worst score.

Table 2. Sample Derivations for the Percentage of Worst Score for the FOP-PFQ

Sample Subject	Observed FOP-PFQ Score	Lowest Possible Score	Highest Possible Score	Range of Possible Scores	Distance = Highest - Observed	Distance/ Range	Percentage of Worst Score
1	45	15	75	60	30	0.500	50.0%
2	40	15	75	60	35	0.583	58.3%
3	35	15	75	60	40	0.667	66.7%
4	30	15	75	60	45	0.750	75.0%
5	25	15	75	60	50	0.833	83.3%

Percentage of worst scores and changes from baseline will be summarized at each scheduled visit with descriptive statistics. Similar summaries will be provided for the percentage of worst Mobility Subscore for the subgroup of subjects with flare-ups in lower extremities and for the percentage of worst Upper Extremity Subscore for the subgroup of subjects with flare-ups in the upper extremities.

Note that flare-ups in the lower extremities include the following flare-up locations reported by the Investigator: abdomen, ankle or foot, distal lower extremities, hip, knee, lower back, and lumbar spine. Flare-ups in the upper extremities include the following flare-up locations reported by the Investigator: chest, cervical spine, head, jaw, distal upper extremities, elbow, shoulder, wrist or hand, thoracic spine, and upper back.

7.2.9 PROMIS Global Health

An age-appropriate form of the PROMIS Global Health short form will be administered to subjects (or parent proxy) at:

- Screening/Baseline, Week 2, Week 4, Week 6, Week 9, and Week 12 (Flare-up Component).

The adult form will be administered to subjects 15 years and older. For subjects between 8 and 14 years of age, both the pediatric self-completed and the pediatric proxy-completed forms will be administered. The proxy-completed form will be used for analysis unless only the self-completed form is available.

The PROMIS Global Health short form consists of 10 questions on the adult form and nine questions on the pediatric form scored on varying scales. The questions are described in detail in Appendices 8A, 8B, and 8C of the protocol.

For the adult form, Global Physical Health and Global Mental Health scores will be calculated. The Global Physical Health score will be calculated as the sum of scores from questions 3, 6, 7, and 8, and will range from 4 (worse health) to 20 (better health). The Global Mental Health score will be calculated as the sum of scores from questions 2, 4, 5, and 10, and will range from 4 (worse health) to 20 (better health). In the calculation of the Global Physical Health and Global Mental Health scores, the following questions will be rescaled as shown in Table 3.

Table 3. Rescaled PROMIS Global Health Scale Scores

Question(s)	Score	Rescaled Score
7	0	5
	1-3	4
	4-6	3
	7-9	2
	10	1
8 and 10	1	5
	2	4
	3	3
	4	2
	5	1

If a subject is missing any of the contributing scores, the Global Physical Health or Global Mental Health score will not be calculated.

For the pediatric form, only a Total score will be calculated as the sum of scores from the first seven questions and will range from 7 (worse health) to 35 (better health). If a subject is missing some (but not more than three) of the contributing question scores, the Total score will be calculated as the average observed score multiplied by the number of expected question scores. For example, the Total score would be calculated as the average of the non-missing scores x 7.

Global Physical Health and Global Mental Health scores and Total scores will be converted to T-scores for analysis as described in detail in Appendix 1. A T-score of 50 is normal and increments of 10 +/- 1 standard deviation away from the norm. A T-score <50 indicates worse health, while a T-score >50 indicates better health. Note that higher values indicate better health.

T-scores and changes from baseline will be summarized at each scheduled visit with descriptive statistics. Data from adult and pediatric forms will be summarized separately.

7.2.10 Flare-up Pain and Swelling

Flare-up pain and swelling will be assessed at:

- Screening/Baseline, Week 2, Week 4, Week 6, Week 9, and Week 12 (Flare-up Component).

Adults and pediatric subjects 8 years and older will rate the pain associated with their flare-ups using a scale from 0 (no pain) to 10 (worst pain ever experienced). Pediatric subjects less than 9 years old will rate pain using the Faces Pain Scale – Revised (FPS-R), which ranges from 0 (no pain) to 10 (very much pain) in 2-point increments. Swelling associated with flare-ups will be rated on a scale from 0 (no swelling) to 10 (worst swelling ever experienced).

Ratings for pain and swelling and changes from baseline will be summarized at each scheduled visit with descriptive statistics.

7.2.11 Assistive Devices and Adaptations for Daily Living

Subjects will be given a list of FOP assistive devices and adaptations grouped into 12 categories and asked to select those they use for daily living. The assessment will occur at:

- Screening/Baseline, Week 6, and Week 12 (Flare-up Component).

The use of assistive devices and adaptations will be summarized overall and by category at each scheduled visit with counts and percentages of flare-ups. The number of devices and adaptations used will be summarized overall and by category at each scheduled visit with descriptive statistics.

The use of new-onset assistive devices and adaptations will be summarized overall and by category with counts and percentages of flare-ups. New-onset assistive devices and adaptations are defined as assistive devices and adaptations used at a post-baseline visit that were not used at baseline.

7.2.12 Cartilage, Bone, Angiogenesis, and Inflammation Biomarkers

Blood and urine samples for cartilage, bone, angiogenesis, and inflammation biomarkers will be collected at:

- Screening/Baseline, Week 2, Week 4, Week 6, and Week 12 (Flare-up Component).

Bone and cartilage biomarkers include osteocalcin, bone specific alkaline phosphatase (B-ALP), procollagen type I C-terminal propeptide (PICP), procollagen 1 N-terminal propeptide (P1NP), CD retinoic acid-sensitive protein (CD-RAP), and C-terminal telopeptide (CTX). The angiogenesis biomarker is the ratio of fibroblast growth factor and urine creatinine. Inflammation biomarkers include erythrocyte sedimentation rate (ESR), C reactive protein (CRP), interleukin 6 (IL-6), interleukin 1 beta (IL-1 beta), tumor necrosis factor (TNF), creatine kinase (CK), and lactate dehydrogenase (LDH).

Biomarkers values and changes from baseline will be summarized at each scheduled visit with descriptive statistics. Note that results obtained locally rather than from the central laboratory will not be used to calculate change from baseline and will be excluded from these summaries. Results captured as "<" or ">" will be rounded up or down accordingly for analysis (eg, "<0.61" will be rounded to 0.60).

Abnormal values at baseline and new-onset abnormal values post-baseline will be summarized with counts and percentages of flare-ups. New-onset abnormal values are defined as values from biomarkers that were normal at baseline and abnormal post-baseline or were abnormal at baseline and abnormal in the opposite direction post-baseline (eg, abnormal low at baseline and abnormal high post-baseline). Note that results obtained locally will be included in these summaries.

The following figures will be generated for the Flare-up Component (note that the unit of analysis is each flare-up):

- Mean biomarker values over time (one figure per biomarker, each figure with one line).

- Mean biomarker values over time by incidence of new HO (one figure per biomarker, each figure with one distinctive line per subgroup [new HO, no new HO]).
- Mean biomarker values over time by baseline edema status (one figure per biomarker, each figure with one distinctive line per subgroup [baseline edema present, baseline edema not present]).
- Biomarker values over time by flare-up (one figure per biomarker, each figure with one distinctive line per flare-up). Flare-ups will be identified in the legend with subject ID, flare-up identifier (eg, Flare A, Flare B), gender, and age.
- Mean percent of normal high biomarker values over time (one figure per biomarker, each figure with one line). Percent of normal high = 100 x value / upper limit of normal range.
- Mean percent of normal high biomarker values over time by incidence of new HO (one figure per biomarker, each figure with one distinctive line per subgroup [new HO, no new HO]).
- Mean percent of normal high biomarker values over time by baseline edema status (one figure per biomarker, each figure with one distinctive line per subgroup [baseline edema present, baseline edema not present]).
- Percent of normal biomarker values over time by flare-up (one figure per biomarker, each figure with one distinctive line per flare-up). Flare-ups will be identified in the legend with subject ID, flare-up identifier, gender, and age.

7.2.13 Duration of Active Symptomatic Flare-up

Subjects will keep a daily diary during the Flare-up Component where the question "Is your flare-up ongoing today?" is answered. Subjects with at least 80% of diary data (at least 67 days with non-missing data) will be used for analysis.

The number of symptomatic days will be calculated as the number of days where the diary question is answered "Yes". The number of symptomatic days will be summarized with descriptive statistics.

The status of the flare-up will be determined at Week 6 and Week 12. If there are no symptomatic days recorded in the diary, the flare-up will be considered "Resolved prior to the first dose of study drug". If the last symptomatic day recorded in the diary is on or before the specified visit and is followed by at least one asymptomatic day (ie, diary question is answered "No"), the flare-up will be considered "Resolved". If the last symptomatic day recorded is on the last diary entry, the flare-up will be considered "Not resolved". The status of the flare-up will be summarized at Week 6 and Week 12 with counts and percentages of subjects or flare-ups as appropriate.

The days to flare-up resolution will be calculated as follows:

- If the flare-up was considered "Resolved prior to the first dose of study drug", days to flare-up resolution = 0.
- If the flare-up was considered "Resolved", days to flare-up resolution = date of flare-up resolution – date of first dose + 1, where the date of flare-up resolution is the last symptomatic day.

- If the flare-up was considered "Not resolved", days to flare-up resolution = date of last diary entry – date of first dose + 1.

The days to flare-up resolution will be summarized with descriptive statistics.

8 Safety Data

Adverse events will be presented for the Enrolled Population as described in Section 8.1. Safety assessments at Flare-up Component visits will generally be presented in total for the Treated Population. Safety assessments at Follow-up Component visits will be listed.

8.1 Adverse Events

Adverse events (AEs) will be recorded from the time informed consent is signed through study completion. All AEs will be coded to body system and preferred term using the Medical Dictionary for Regulatory Affairs (MedDRA) Version 17.0. At least possibly related AEs are defined as AEs assessed by the Investigator as possibly, probably, or definitely related to study drug. Adverse events that require Common Terminology Criteria for Adverse Events (CTCAE) grading as indicated on the eCRF are considered to be retinoid-associated AEs.

Adverse events with start date on or after Flare-up Study Day 1 and on or before the date of completion for the Flare-up Component will be attributed to the Treatment Period. Adverse events with start date on or after Study Day 1 that are not attributed to the Treatment Period will be attributed to the Non-Treatment Period.

For the Untreated Population, AEs will be summarized in total (note, these are all Non-Treatment Period AEs). For the Treated Population, AEs will be summarized by period (Non-Treatment Period AEs and Treatment Period AEs) and in total. For the Enrolled Population, AEs will be summarized for the Non-Treatment Period and in total.

The incidence of the following AEs will be summarized with counts and percentages of subjects (overview and also by MedDRA body system and preferred term):

- Any AEs
- Any AEs by maximum severity
- Any AEs at least possibly related to study drug
- Any AEs at least possibly related to study drug by maximum severity
- Any retinoid-associated AEs
- Any retinoid-associated AEs by maximum CTCAE grading
- Any serious adverse events (SAEs)
- Any AEs leading to dose modification or interruption of study drug
- Any AEs leading to discontinuation of study drug
- Any AEs leading to discontinuation of study

Data listings will be provided for (1) all AEs, (2) SAEs, (3) AEs leading to dose modification or interruption of study drug, (4) AEs leading to discontinuation of study drug, and (5) AEs leading to discontinuation of study. Study Day and/or Flare-up Study Day for AE start and stop dates will be included on AE listings as appropriate.

8.2 Safety Laboratory Assessment

Samples for safety laboratory assessments (biochemistry, lipids, hematology, coagulation, and urinalysis) will be collected at:

- Screening/Baseline, Week 2, Week 4, Week 6, and Week 12 (Flare-up Component).

Values and changes from baseline will be summarized at each scheduled visit with descriptive statistics. Mean values (+/- standard error) will be plotted over time during the Flare-up Component for triglycerides, cholesterol, alanine aminotransferase, aspartate aminotransferase, amylase, and lipase. Note that results obtained locally rather than from the central laboratory will not be used to calculate change from baseline and will be excluded from these summaries.

Potentially clinically significant (PCS) values at baseline and new-onset PCS values post-baseline will be summarized with counts and percentages of subjects or flare-ups as appropriate. New-onset PCS values are defined as values that were not PCS at baseline and were PCS post-baseline. Note that results obtained locally will be included in these summaries. PCS values are defined in Table 4.

Table 4. Potentially Clinically Significant Safety Laboratory Values

	PCS Low	PCS High
Biochemistry		
Aspartate aminotransferase	N/A	3 x ULN
Alanine aminotransferase	N/A	3 x ULN
Amylase	N/A	3 x ULN
Lipase, pancreatic	N/A	3 x ULN
Bilirubin	N/A	>2 mg/dL
Thyroxine	<4.0 mcg/dL	>13.0 mcg/dL
Lipids		
Cholesterol	N/A	>300 mg/dL
Triglycerides	N/A	>400 mg/dL
Hematology		
Leukocytes	<2.8 x 10^9/L	>16.0 x 10^9/L
Hemoglobin (females)	<9.5 g/dL	>17.5 g/dL
Hemoglobin (males)	<11.5 g/dL	>19.0 g/dL
Hematocrit (females)	<32%	>54%
Hematocrit (males)	<37%	>60%
Platelets	<75 x 10^9/L	>700 x 10^9/L

A listing will be provided for all subjects with PCS lipase values that will include all lipase results during the study along with associated palovarotene and glucocorticoid dosing data (dose and start and stop dates).

8.3 Vital Signs

Vital signs (blood pressure, heart rate, respiratory rate, and temperature) will be collected at:

- Screening/Baseline, Week 2, Week 4, Week 6, and Week 12 (Flare-up Component).

Values and changes from baseline will be summarized at each scheduled visit with descriptive statistics. Mean values (+/- standard error) will be plotted over time during the Flare-up Component for blood pressure and heart rate.

Potentially clinically significant (PCS) values at baseline and new-onset PCS values post-baseline will be summarized with counts and percentages of subjects or flare-ups as appropriate. New-onset PCS values are defined as values that were not PCS at baseline and were PCS post-baseline. PCS values are defined in Table 5.

Table 5. Potentially Clinically Significant Vital Sign Values

	PCS Low	PCS High
Systolic blood pressure	<86 mmHg or a decrease of ≥25 mmHg from baseline	>180 mmHg or an increase of ≥25 mmHg from baseline
Diastolic blood pressure	<48 mmHg or a decrease of ≥20 mmHg from baseline	>110 mmHg or an increase of ≥20 mmHg from baseline
Heart rate	<45 bpm or a decrease of ≥20 bpm from baseline	(1) >105 bpm AND an increase of ≥20 bpm from baseline (2) >125 bpm

mmHg=millimeters of mercury, bpm=beats per minute.

8.4 Body Weight

Body weight will be collected at:

- Screening/Baseline, Week 2, Week 4, Week 6, and Week 12 (Flare-up Component).

Values and changes from baseline will be summarized at each scheduled visit with descriptive statistics. Mean values (+/- standard error) will be plotted over time during the Flare-up Component.

8.5 Knee and Hand/Wrist Radiograph

Subjects under the age of 18 years at the time of enrollment into Study PVO-1A-202 with open epiphyseal growth plates will undergo knee and hand/wrist radiographs (AP view) at:

- Screening/Baseline and Week 12 (Flare-up Component).

Knee and hand/wrist radiographs will be performed remotely (eg, at a local medical facility) for subjects who are unable or unwilling to attend site visits. Once a subject has achieved 100% skeletal maturity (as confirmed by radiography), knee and hand/wrist radiographs will no longer be required. The incidence of abnormalities may be summarized with counts and percentages of subjects.

8.6 Linear Growth Assessment

A linear growth assessment will be performed for subjects under the age of 18 years at the time of enrollment into Study PVO-1A-202 with open epiphyseal growth plates at:

- Screening/Baseline and Week 12 (Flare-up Component).

Height will be measured at Study Month 12 (Follow-up Component) and Screening/Baseline (Flare-up Component).

Based on work by Chumlea (1994)[3], knee height can be used to predict total height (stature) for children who are mobility-impaired. Separate equations were established for White and for Black children by gender, as follows:

- White boys: predicted stature = 40.54 + (2.22 x knee height)
- Black boys: predicted stature = 39.60 + (2.18 x knee height)
- White girls: predicted stature = 43.21 + (2.15 x knee height)
- Black girls: predicted stature = 46.59 + (2.02 x knee height)

When race is not reported as White or Black, the calculation for the White boy/girl will be used.

Values and changes from baseline will be summarized for height, knee to heel length, and predicted stature at each scheduled visit with descriptive statistics. Values for height and knee to heel length may be plotted over time with the actual palovarotene dose taken at each visit identified.

8.7 Electrocardiogram

A 12-lead electrocardiogram (ECG) will be performed at:

- Screening/Baseline and Week 6 (Flare-up Component).

Values and changes from baseline will be summarized for continuous ECG parameters at each scheduled visit with descriptive statistics.

Abnormal ECG categories will be assigned based on raw findings as shown in Table 6.

Table 6. Abnormal Electrocardiogram Categories

Abnormal ECG Category	Raw Findings
Rhythm	* Atrial premature complexes * Ventricular premature complexes * Sinus arrhythmia * Sinus tachycardia
Conduction	* Nonspecific intraventricular conduction delay * Right bundle branch block * Incomplete right bundle branch block * Left posterior fascicular block * Left bundle branch block * Incomplete left bundle branch block
Morphology/Chamber enlargement	* Left ventricular hypertrophy * Left ventricular hypertrophy with repolarization abnormality
Axis deviation	* Right axis deviation * Left axis deviation
Myocardial infarction	* Pathologic Q waves * Acute ST elevation MI * Acute non-ST elevation MI
ST Segment/T Waves/U Waves	* Early repolarization * Nonspecific ST and T wave abnormality * Nonspecific ST elevation * Nonspecific U wave abnormalities
Miscellaneous	* Long QT interval * Short PR interval

Overall interpretation and abnormal ECG categories will be summarized at each scheduled visit with counts and percentages of flare-ups. New-onset ECG abnormalities are defined as abnormalities that did not exist at baseline. New-onset ECG abnormalities will be summarized counts and percentages of flare-ups.

Incidence of QTc >450 msec, QTc >500 msec, or change from baseline in QTc >60 msec will be summarized with counts and percentages of flare-ups.

Potentially clinically significant (PCS) values at baseline and new-onset PCS values post-baseline will be summarized with counts and percentages of flare-ups. New-onset PCS values are defined as values that were not PCS at baseline and were PCS post-baseline. PCS values are defined in Table 7.

Table 7. Potentially Clinical Significant Electrocardiogram Values

	PCS Low	PCS High
PR interval	N/A	>200 msec or an increase of ≥20 msec
QRS interval	N/A	>100 msec or an increase of ≥10 msec
QT interval, QTcF, QTcB	N/A	>500 msec or an increase of ≥60 msec

msec = millisecond, QTcF = rate-corrected QT interval using Fridericia's formula (QT interval divided by the cube root of the RR interval), QTcB = rate-corrected QT interval using Bazett's formula (QT interval divided by the square root of the RR interval)

8.8 Physical Examination

A comprehensive physical examination will be performed at:

- Screening/Baseline, Week 6, and Week 12 (Flare-up Component).

Physical examination results will be summarized at each scheduled visit by body system with counts and percentages of flare-ups.

8.9 Columbia-Suicide Severity Rating Scale

Suicidal ideation and behavior will be assessed using the Columbia-Suicide Severity Rating Scale (C-SSRS) at:

- Screening/Baseline, Week 2, Week 4, Week 6, Week 9, and Week 12 (Flare-up Component).

Data listings will be provided for (1) all C-SSRS results; and (2) for subjects reporting any type 4 or type 5 suicidal ideations.

9 References

1. Norken CC, White DJ. *Measurement of Joint Motion: A Guide to Goniometry*. 4th ed. Philadelphia, PA: F.A. Davis Company; 2011.
2. Hays RD, Spritzer KL, Amtmann D, et al. Upper-extremity and mobility subdomains from the Patient-Reported Outcomes Measurement Information System (PROMIS) adult physical functioning item bank. *Arch Phys Med Rehabil.* 2013;94(11):2291-2296. doi:10.1016/j.apmr.2013.05.014.
3. Chumlea WC, Guo SS, Steinbaugh ML. Prediction of stature from knee height for black and white adults and children with application to mobility-impaired or handicapped persons. *J Am Diet Assoc*. 1994;94(12):1385-1388, 1391.

Appendices

Appendix 1. PROMIS T-Score Conversions

The following conversion tables allow a user to convert Global Physical Health, Global Mental Health, and Total scores into T-scores. T-score distributions are standardized such that a 50 represents the average (mean) for the US general population, and the standard deviation around that mean is 10 points. A high score always represents more of the concept being measured. Thus, a subject who has a T-score of 60 is one standard deviation better (more healthy) than the general population.

Adult Global Physical Health		
Raw Score	**T-Score**	**Standard Error**
4	16.2	4.8
5	19.9	4.7
6	23.5	4.5
7	26.7	4.3
8	29.6	4.2
9	32.4	4.2
10	34.9	4.1
11	37.4	4.1
12	39.8	4.1
13	42.3	4.2
14	44.9	4.3
15	47.7	4.4
16	50.8	4.6
17	54.1	4.7
18	57.7	4.9
19	61.9	5.2
20	67.7	5.9

Adult Global Mental Health		
Raw Score	**T-Score**	**Standard Error**
4	21.2	4.6
5	25.1	4.1
6	28.4	3.9
7	31.3	3.7
8	33.8	3.7
9	36.3	3.7
10	38.8	3.6
11	41.1	3.6
12	43.5	3.6
13	45.8	3.6
14	48.3	3.7
15	50.8	3.7
16	53.3	3.7
17	56.0	3.8
18	59.0	3.9
19	62.5	4.2
20	67.6	5.3

Pediatric Self-Completed Total		
Raw Score	**T-Score**	**Standard Error**
7	16.0	3.4
8	17.1	3.6
9	18.3	3.7
10	19.7	3.8
11	21.2	3.8
12	22.8	3.7
13	24.4	3.6
14	26.1	3.6

Pediatric Proxy-Completed Total		
Raw Score	**T-Score**	**Standard Error**
7	14.7	2.9
8	15.3	3.1
9	16.0	3.2
10	16.9	3.4
11	18.1	3.6
12	19.4	3.7
13	21.0	3.8
14	22.7	3.8

Pediatric Self-Completed Total		
Raw Score	**T-Score**	**Standard Error**
15	27.6	3.5
16	29.2	3.5
17	30.8	3.5
18	32.4	3.6
19	34.0	3.6
20	35.6	3.6
21	37.2	3.6
22	38.8	3.6
23	40.4	3.6
24	42.1	3.7
25	43.9	3.7
26	45.7	3.6
27	47.5	3.6
28	49.2	3.6
29	51.1	3.7
30	53.3	3.9
31	55.7	4.2
32	58.3	4.5
33	61.1	4.9
34	64.2	5.4
35	67.5	6.1

Pediatric Proxy-Completed Total		
Raw Score	**T-Score**	**Standard Error**
15	24.4	3.7
16	26.1	3.7
17	27.7	3.7
18	29.4	3.8
19	31.2	3.8
20	32.9	3.8
21	34.6	3.8
22	36.2	3.8
23	37.9	3.9
24	39.7	4.0
25	41.7	4.0
26	43.6	3.9
27	45.4	3.8
28	47.3	3.9
29	49.3	4.1
30	51.8	4.4
31	54.5	4.7
32	57.3	5.0
33	60.2	5.4
34	63.2	6.0
35	66.1	6.5

Statistical Analysis Plan

A Phase 2, Open-Label Extension, Efficacy and Safety Study of a RARγ-Specific Agonist (Palovarotene) in the Treatment of Preosseous Flare-ups in Subjects with Fibrodysplasia Ossificans Progressiva (FOP)

Study Number: PVO-1A-202 Part B

Original Protocol: 04 June 2014
Amendment 1: 30 September 2014
Amendment 2: 22 June 2015
Amendment 3: 10 March 2016

Statistical Analysis Plan Version: 1.0
Statistical Analysis Plan Date: 23 October 2018

Clementia Pharmaceuticals Inc.
4150 Saint Catherine Street West, Suite 550
Montreal, Quebec, H3Z 2Y5
Canada

Statistical Analysis Plan Signature Page

A Phase 2, Open-Label Extension, Efficacy and Safety Study of a RARγ-Specific Agonist (Palovarotene) in the Treatment of Preosseous Flare-ups in Subjects with Fibrodysplasia Ossificans Progressiva (FOP)

Protocol Number: PVO-1A-202 Part B

Signature of Approval for Statistical Analysis Plan (Version 1: 23 Oct 2018)

Prepared by:

MEDPACE INC.
PPD

NAME: PPD

SIGNATURE: _____ **DATE:** 24-Oct-2018

Approved by:

CLEMENTIA PHARMACEUTICALS INC.
PPD

NAME: _____

SIGNATURE: PPD _____ **DATE:** _____

CLEMENTIA PHARMACEUTICALS INC.
PPD

NAME: _____

SIGNATURE: PPD _____ **DATE:** _____

MEDPACE INC.
PPD

NAME: PPD

SIGNATURE: _____ **DATE:** 24-Oct-2018

Table of Contents

List of Tables

List of Abbreviations

Abbreviation	Definition
AE	adverse event
ATC	anatomic therapeutic class
AP	anterior/posterior
CAJIS	Cumulative Analogue Joint Involvement Scale
C-SSRS	Columbia-Suicide Severity Rating Scale
CT	computed tomography
CTCAE	Common Terminology Criteria for Adverse Events
ECG	electrocardiogram
eCRF	electronic case report form
FAS	Full Analysis Set
FOP	Fibrodysplasia Ossificans Progressiva
FOP-PFQ	FOP-Physical Function Questionnaire
FPS-R	Faces Pain Scale-Revised
GEE	generalized estimating equation
HO	heterotopic ossification
MedDRA	Medical Dictionary for Regulatory Activities
MRI	magnetic resonance imaging
NHS	Natural History Study
NRS	numeric rating scales
PCS	potentially clinically significant
PROMIS	Patient-Reported Outcomes Measurement Information System
RARγ	retinoic acid receptor gamma
ROM	range of motion
SAE	serious adverse event
SAP	statistical analysis plan
ULN	upper limit of normal
US	Ultrasound
WBCT	whole body CT

1 Introduction

This Statistical Analysis Plan (SAP) describes the analyses to be performed for Study PVO-1A-202 Part B based on Amendment 3 of the protocol dated 10 March 2016 and the France-specific Amendment 1 of the protocol for Study PVO-1A-204, dated 07 September 2016. Data collected for Study PVO-1A-202 Part B and Study PVO-1A-204 will be combined for the analyses. For the purposes of this SAP, "Study PVO-1A-202 Part B" (or simply "Part B") represents data from both studies.

The analyses performed for Study PVO-1A-202 Parts A and C are described in separate SAPs.

2 Overview

2.1 Study Objectives

2.1.1 Primary Objective

The primary objective is:

- To evaluate the safety and efficacy of different palovarotene dosing regimens in subjects with FOP. Efficacy will be assessed based on the ability of palovarotene to prevent the formation of new heterotopic ossification (HO) as assessed by low-dose computed tomography (CT) scan (or plain radiographs for subjects unable to undergo CT scan).

2.1.2 Secondary Objectives

- To evaluate the effect of palovarotene to prevent HO as assessed by low-dose, whole body computed tomography (WBCT) scan, excluding head.
- To evaluate the effect of palovarotene on active range of motion (ROM) at the flare-up site as assessed by goniometer.
- To evaluate the effect of palovarotene on ROM as assessed by the Cumulative Analogue Joint Involvement Scale (CAJIS) for FOP.
- To evaluate the effect of palovarotene by the subject and Investigator global assessment of movement at the flare-up site.
- To evaluate the effect of palovarotene on physical function using age-appropriate forms of the FOP Physical Function Questionnaire (PFQ).
- To evaluate the effect of palovarotene on physical and mental health using age-appropriate forms of the Patient Reported Outcomes Measurement Information System (PROMIS) Global Health Scale.
- To evaluate the effect of palovarotene on pain and swelling associated with the flare-up using numeric rating scales (NRS) or the Facial Pain Scale-Revised (FPS-R) in subjects under 8 years of age.
- To evaluate the effect of palovarotene on the use of assistive devices and adaptations for daily living by FOP subjects.
- To evaluate the effects of palovarotene on soft tissue swelling and cartilage formation at the flare-up site as assessed by magnetic resonance imaging (MRI); or the effect of

palovarotene on soft tissue swelling as assessed by ultrasound (US) in subjects unable to undergo MRI.

- To evaluate bone, cartilage, angiogenesis, and inflammation biomarkers and explore correlations between changes from baseline in biomarkers and clinical efficacy.
- To evaluate the pharmacokinetics of palovarotene during the first treated flare-up.
- To evaluate the ability of different dosing regimens of palovarotene to prevent new flare-ups.
- To evaluate the long-term safety and efficacy of prior palovarotene treatment at the original flare-up site in Study PVO-1A-201; efficacy will be based on the amount of new HO formed at the original flare-up site as assessed by low-dose CT scan (or plain radiographs in subjects unable to undergo CT scan) in the Pediatric Cohort or low-dose whole body CT scan, excluding head, in the Adult Cohort.

2.2 Study Population

A total of approximately 60 subjects will be enrolled in Part B including approximately 40 subjects who completed Study PVO-1A-201 and up to 20 new Adult Cohort subjects. The Adult Cohort will include subjects with at least 90% skeletal maturity, regardless of age; the Pediatric Cohort will include subjects with less than 90% skeletal maturity, regardless of age. Any Pediatric Cohort subject who achieves ≥90% skeletal maturity during Part B may be considered for enrollment into the Adult Cohort at the discretion of the Investigator.

Flare based analyses will include the treatment groups: PVO 20/10 mg, Prior Non-Flare-up/PVO 20/10 mg, and Combined PVO 20/10 mg (=PVO 20/10 mg + Prior Non-Flare-up/PVO 20/10 mg groups). PVO 20/10 mg is defined as flare-ups that did not have prior non-flare-up treatment. Prior Non-Flare-up/PVO 20/10 mg is defined as flare-ups that were treated with Non-Flare-up 5 mg treatment prior to the start of the flare-up treatment. It is important to mention that the flare-up analyses is based on flare-ups, thus subjects may be included in both treatment groups, depending on when non-flare-up dosing began.

WBCT analyses are based on the PVO Non-Flare-up/PVO 20/10 mg treatment group and will include the following subsets: Treated Flare-ups, Under/Untreated Flare-ups, No Flare-ups, and Combined Treated/No Flare-ups.

Only subjects with Non-Flare-up treatment are imaged using WBCT, thus only the PVO Non-Flare-Up/PVO 20/10 mg treatment group is present. This treatment group is subset into 3 different treatment groups, depending on whether or not the subjects had treated, untreated or no flares during the duration between the baseline and Month 12 WBCT scans. Treated Flare-ups are defined as subjects who received non-flare-up/flare-up regimen for all flare-ups. Under/Untreated Flare-ups are defined as subjects who did not receive palovarotene dosing for all flare-ups. No flare-ups are defined as subjects who received the non-flare-up regimen and did not report any flare-ups.

Safety treatment groups include the PVO 20/10 mg and Non-Flare-up 5 mg groups.

2.3 Study Design

PVO-1A-202 Part B is a Phase 2, multicenter, event-based, open-label extension of Study PVO-1A-201. This open-label extension will give subjects who completed Study PVO-1A-201 (including any subject who participated in Part A of Study PVO-1A-202) the opportunity to be followed for up to 2 years.

Subjects in the Adult Cohort will follow all procedures and undergo all assessments as specified in the Schedule of Assessments in Table 1 of the protocol. Site visits will occur at Part B Screening and at Study Months 12 and 24. The visit window for the visits at Study Months 12 and 24 is +/- 2 months where Study Day 1 is the date of screening in Part B. Remote visits (assessments performed at home or at a local medical facility) and telephone contact with the study site will occur every 3 months unless the Investigator deems that a site visit is necessary. The visit window for remote visits is +/- 2 weeks. Urine pregnancy tests will be performed each month for female subjects of childbearing potential.

Subjects in the Adult and Pediatric Cohorts receiving flare-up based treatment will follow all procedures and undergo all assessments as specified in the Schedule of Assessments in Table 2 of the protocol. These subjects will have site visits at Flare-Up Screening/Baseline (as soon as possible after the start of the flare-up) and at Flare-Up Day 84/end-of-treatment (EOT); and remote visits, where possible, at Flare-Up Days 14, 28, 42, 56, and 70. Should treatment be extended beyond Flare-up Day 84, 10 mg palovarotene (or the dose-equivalent) will be administered, in 4-week increments, until the flare-up resolves, with remote visits performed every 2 weeks while on treatment. The visit window is +/- 3 days where Flare-Up Day 1 is the date of first dose of flare-up based palovarotene except for Flare-up Day 84/EOT where the visit window is -3 days. Note that for analysis, visits will be displayed as Flare-Up Day 1, Flare-Up Week 2, Flare-Up Week 4, Flare-Up Week 6, Flare-Up Week 8, Flare-Up Week 10, Flare-Up Week 12, Flare-Up Week ## (for extended treatment), and EOT.

Pharmacokinetic blood samples will be collected within the first 7 days of starting flare-up based treatment and at Flare-up Day 84 at 0, 3, 6, 10, and 24 hours post-dose for the first treated flare-up only.

To assess an adequate number of flare-ups with the new dosing regimens, it is anticipated that each subject will participate in Part B of the study for approximately 18-24 months. Once a flare-up has resolved, the Adult Cohort subjects will resume non-flare-up based treatment of 5 mg once daily. Initiation of flare-up based treatment for a new flare-up can occur at any time after resolution of the prior flare-up.

Data regarding the original flare-up for those subjects from Study PVO-1A-201 will be collected at the Part B Screening visit (when the subject signs the informed consent) for the Adult Cohort according to the Schedule of Assessments in Table 1 of the Protocol; or at the Study Month 12 visit (starting from initial enrollment into Study PVO-1A-202) for the Pediatric Cohort according to the Schedule of Assessments in Table 3 of the Protocol.

Assessments for the Study Month 12 visit can be performed in conjunction with any of the site visits associated with flare-up based treatment (Flare-up Screening or Flare-up Day 84) as long

as those visits are within 5 months before or after the scheduled Study Month 12 visit. Subjects on Non-Flare-up dosing regimen will also be contacted by telephone at Study Months 6, 18, and 24 to assess concomitant medications and adverse events (AEs). The visit window is +/- 2 weeks where Day 1 is the date of informed consent for Part A.

2.4 Study Treatment

Subjects in the Adult Cohort were scheduled to be treated with 5 mg palovarotene daily for up to 24 months (non-flare-up based treatment); however, some adult subjects chose not to receive non-flare-up based treatment.

Flare-up based treatment will begin as soon as possible after a subject reports onset of a flare-up and the flare-up has been confirmed by the Investigator. In the event of an eligible flare-up, subjects in the Adult Cohort will receive open-label palovarotene treatment as follows:

- 20 mg for 28 days. The first dose will be taken upon flare-up confirmation by the Investigator.
- 10 mg for 56 days, for a total flare-up treatment duration of 84 days; treatment may be extended if the flare-up is ongoing and continue until the flare-up resolves. Dosing will be extended in 4-week intervals and be based on clinical signs and symptoms as assessed by the Investigator.

Subjects not receiving non-flare-up based dosing will only be treated during eligible flare-ups and will receive weight-adjusted doses of 20 and 10 mg palovarotene as follows:

Weight Range Category	20-mg Equivalent	15-mg Equivalent	10-mg Equivalent	7.5-mg Equivalent*	5-mg Equivalent*
20 to <40 kg	12.5 mg	10 mg	6 mg	4 mg	3 mg
40 to <60 kg	15 mg	12.5 mg	7.5 mg	5 mg	4 mg
≥60 kg	20 mg	15 mg	10 mg	7.5 mg	5 mg

* In the event of dose de-escalation.

If the subject experiences intolerable side effects during flare-up based dosing, the dose may be reduced to the next lower dose; if the subject is already receiving the lowest possible dose, then study drug will be discontinued. Treatment may be extended after the 84-day treatment period (flare-up based dosing) if the flare-up is ongoing and continue until the flare-up resolves. Dosing will be extended in 4-week intervals based on clinical signs and symptoms as assessed by the Investigator, with remote visits performed every 2 weeks while on treatment. In the event the subject requires dose de-escalation due to an intolerable side effect for a previous flare-up, then the dose the subject will receive for a subsequent flare-up will be determined by the Investigator and the Medical Monitor.

After the Flare-up Day 84/EOT visit has been conducted and the flare-up is considered resolved, subjects in the Adult Cohort will resume the non-flare-up based treatment.

Subjects not receiving non-flare-up based dosing who do not experience an eligible flare-up will not be treated.

2.5 Randomization and Blinding

This is an open-label study and does not involve randomization or blinding.

3 General

Categorical data will be summarized with counts and percentages. Percentages will be calculated based on the number of non-missing values and will be reported to one decimal place.

Continuous data will be summarized with descriptive statistics including the number of non-missing values, mean, standard deviation, standard error, median, minimum, and maximum. Minimum and maximum will be reported to the same precision as the raw values, mean and median will be reported to one additional decimal place, and standard deviation and standard error will be reported to two additional decimal places.

All data will be listed. When listings are provided instead of summaries or require additional description, they are described in this document.

3.1 Adjustments for Covariates

No statistical modeling will be performed so adjustment for covariates is not planned.

3.2 Handling of Dropouts or Missing Data

As there is no effective treatment for subjects with FOP, it is assumed they will be highly motivated to adhere to the protocol and treatment visits. Thus, it is expected that there will be minimal missing data, and most subjects will complete the study. The efficacy analysis will be performed using observed data only. Missing efficacy data will not be imputed.

3.3 Interim Analyses

No interim analysis is planned.

3.4 Multicenter Studies

A review of by-center effects will be performed in the context of data listing review. However, the sample size precludes formal assessment of by-center effects.

3.5 Multiple Comparison/Multiplicity

As this is an extension of an initial proof-of-concept study, no adjustment for multiple testing is planned.

3.6 Examination of Subgroups

The sample size in this study is not sufficient to allow rigorous assessment of outcomes by subgroups, and the study is not powered for subgroup assessment. However, exploratory assessments of the outcomes may be performed with a limited number of extrinsic or intrinsic factors, including age of the subject, location of the flare-up, amount of new HO, number of flare-up treatments, or other variables as appropriate.

3.7 Part B Study Day, Non-Flare-up Day, and Flare-up Day

For all subjects, Study Day will be determined from the date of informed consent (Part B Study Day 1); Part B Study Day -1 will be the day immediately prior to Part B Study Day 1.

For subjects receiving non-flare-up based treatment, Non-Flare-up Day will be determined from the date of first dose of non-flare-up treatment (Part B Non-Flare-up Day 1).

For subjects receiving flare-up treatment, Flare-up Day will be determined from the date of first dose of flare-up treatment (Flare-up Day 1).

4 Analysis Populations

There will be seven Part B populations defined for the analysis.

4.1 Enrolled Population

The Enrolled Population will include all subjects enrolled in Part B. The Enrolled Population will be the primary population for demographic and baseline summaries and for analysis of adverse events.

4.2 Treated Population

The Treated Population will include all subjects who took at least one dose of palovarotene in Part B (either flare-up treatment or non-flare-up treatment). The analysis will be stratified as follows: PVO 20/10 mg only (including weight-adjusted doses), PVO 20/10 mg (including weight-adjusted doses) with prior non-flare-up treatment, and non-flare-up treatment only.

4.3 Flare-up Population

The Flare-up Population will include all subjects in the Treated Population who took at least one dose of flare-up based treatment in Part B. The analysis will be stratified based on subjects who took PVO 20/10 mg only (including weight-adjusted doses) and subjects took PVO 20/10 mg (including weight-adjusted doses) with prior non-flare-up treatment.

4.4 Flare-up Per Protocol Population

The Flare-up Per Protocol (PP) Population will include all subjects in the Flare-up Population with no major protocol deviations that may impact the efficacy assessment and who were at least 80% compliant with flare-up dosing. Exclusions from the Flare-up PP Population will be identified prior to database lock. The analysis will be stratified as before for the Flare-up Population.

4.5 Pharmacokinetic Population

The Pharmacokinetic (PK) Population will include all subjects in the Flare-up Population who have sufficient blood samples collected for valid estimation of PK parameters. The analysis will be stratified as before for the Flare-up Population.

4.6 WBCT Population

The WBCT Population will include subjects in the Flare-up Population who had prior non-flare-up treatment and have Baseline and Month 12 WBCT scans.

4.7 WBCT Per-Protocol Population

The WBCT Per-Protocol Population will include all subjects in the WBCT Population with no major protocol deviations that may impact the efficacy assessment and who were at least 80% compliant with non-flare-up dosing. Exclusions from the WBCT PP Population will be identified prior to database lock.

4.8 Non-Flare-up Treated Population

The Non-Flare-up Treated Population will include all subjects in the Treated Population who took at least one dose of non-flare-up based treatment in Part B. The Non-Flare-up Treated Population will be used to assess safety data for non-flare-up based treatment assessments.

5 Subject Data

In general, subject data will be summarized for the Enrolled Population by cohort and in total.

5.1 Subject Disposition and Analysis Populations

The following will be summarized with counts and percentages using the number of subjects in the Enrolled Population as the denominator:

- Subjects who were enrolled (Enrolled Population).
- Subjects in the Enrolled Population who took at least one dose of palovarotene (Treated Population).
- Subjects in the Treated Population who took at least one dose of flare-up based palovarotene (Flare-up Population).
- Subjects in the Flare-up Population who did not have major protocol deviations that may impact the efficacy assessment and were at least 80% compliant with flare-up dosing (Flare-up PP Population).
- Subjects in the Flare-up Population who had sufficient blood samples collected for valid estimation of PK parameters (PK Population).
- Subjects in the Flare-up Population who took at least one dose of non-flare-up based palovarotene and also have Baseline and Month 12 WBCT scans (WBCT Population).
- Subjects in the WBCT Population who did not have major protocol deviations that may impact the efficacy assessment and were at least 80% compliant with non-flare-up dosing (WBCT PP Population).
- Subjects who discontinued from Part B early along with reasons for discontinuation.
- Subjects who completed Part B.
- Duration of participation in Part B. This will be calculated from the date of informed consent to the date of completion of Part B/discontinuation.

- Months of participation calculated as days of participation / 30.5. Months of participation will also be summarized with counts and percentages of subjects in the Enrolled Population according to the following categories:
  - ≤ 6 months
  - >6 – 12 months
  - >12 – 18 months
  - >18 – 24 months

Summaries will be presented for the following groups: subjects receiving non-flare-up based treatment only, subjects receiving flare-up based treatment only, and subjects receiving both non-flare-up based and flare-up based treatment.

5.2 Eligibility Criteria

Eligibility criteria (inclusion/exclusion) will be listed with any specific criteria that were not met shown in the by-subject listing for the Enrolled Population.

5.3 Protocol Deviations

Major protocol deviations will be identified prior to locking the database for Part B and may include, but are not limited to departures from the inclusion/exclusion criteria, non-compliance with investigational product, use of restricted concomitant medications, and non-compliance with study procedures. Major protocol deviations will be summarized for the Enrolled Population.

5.4 Demographic and Other Baseline Characteristics

Demographic and other baseline characteristics such as age, sex, race, ethnicity, height, and weight will be summarized for all analysis populations.

5.5 Medical History

Medical history will be coded using the Medical Dictionary for Regulatory Activities (MedDRA Version 17.0). The incidence of medical history conditions will be summarized by MedDRA body system and preferred term for the Enrolled Population.

5.6 Treated Flare-ups Starting During Part B

The number of treated flare-ups will be summarized with counts and percentages of subjects in the Flare-up Population and Flare-up PP Population. Similar summaries will be provided for the number of flare-ups with dosing extended beyond 12 weeks, the number of flare-ups for which an interruption of study drug occurred during flare-up dosing, and the number of flare-ups for which study drug was discontinued during flare-up dosing.

Treated flare-up data as reported by the subject and the Investigator including symptoms, probable causes, and location(s) will be summarized for the Flare-up and Flare-up PP Populations.

The number of days after flare-up onset to the date of first dose of study drug for the flare-up will be summarized for the Flare-up Population with descriptive statistics.

Data for untreated flare-ups will be listed.

5.7 Concomitant Medications

Concomitant medications will be coded using the WHO Drug Dictionary (March 2014). Glucocorticoids taken for a flare-up are defined as courses of glucocorticoids that start within 9 days after the start of the flare-up and are specified as taken for the current flare-up on the electronic case report form (eCRF). Courses of glucocorticoids are defined as days of continuous treatment allowing for up to a 2-day interruption.

The use of the following concomitant medications will be summarized by anatomic therapeutic class (ATC) term and preferred term for the Enrolled Population:

- Prior medications that were on-going at Study Day 1 excluding glucocorticoids. (Medications with start date before Study Day 1 and end date on or after Study Day 1.)
- New-onset medications outside flare-up based treatment periods excluding glucocorticoids. (Medications with start date on or after Study Day 1 and not during Flare-up Day 1 through Flare-up Day 84/EOT.)
- Glucocorticoids outside flare-up based treatment periods that were on-going at Study Day 1. (Courses of glucocorticoids with start date before Study Day 1 and end date on or after Study Day 1.)
- New-onset Glucocorticoids outside flare-up based treatment periods. (Courses of glucocorticoids with start date on or after Study Day 1, not during Flare-up Day 1 through Flare-up Day 84/EOT, and are specified as not being taken for a flare-up.)

The use of the following concomitant medications will be summarized by ATC term and preferred term for the Flare-up Population:

- Prior medications that were on-going at Flare-up Day 1 excluding glucocorticoids. (Medications with start date before Flare-up Day 1 and end date on or after Flare-up Day 1.)
- New-onset medications during flare-up based treatment periods excluding glucocorticoids. (Medications with start date during Flare-up Day 1 through Flare-up Day 84/EOT.)
- Glucocorticoids taken for treated flare-ups defined as courses of glucocorticoids that start within 9 days after the start of the flare-up and are specified as taken for the current flare-up on the eCRF.
- New-onset glucocorticoids during flare-up based treatment periods defined as courses of glucocorticoids that start more than 9 days after the start of the flare-up or are specified as not being taken for the current flare-up on the eCRF.

For the purposes of analysis, incomplete medication start and stop dates will be imputed. If a medication start date is incomplete, January will be imputed for missing month and/or the first day of the month will be imputed for missing day. If a medication stop date is incomplete,

December will be imputed for missing month and the last day of the month will be imputed for missing day. If the imputed medication stop date is after the date of study completion, the date of study completion will be used instead.

A listing of courses of glucocorticoids will include additional derived data such as start and stop dates, based on the first dose date of flare-up treatment, with associated Flare-up Day, start day of the glucocorticoid relative to the start of the flare-up onset date, duration of use, and total amount taken.

5.8 Exposure

Exposure to flare-up based treatment for each treated flare-up will be calculated from the date of first dose for the flare-up to the date of last dose for the flare-up. Days of exposure to flare-up based treatment will be summarized for the Flare-up Population with descriptive statistics. Similarly, exposure to high flare-up treatment (the 20-mg dose), low flare-up treatment (the 10-mg dose), and non-flare-up treatment (5 mg) will be summarized. Separate summaries will include and exclude weight-adjusted doses.

Weeks of exposure to flare-up based treatment will be summarized with counts and percentages of flare-ups in the Flare-up Population according to the following categories:

- ≤2 weeks
- >2 – 4 weeks
- >4 – 6 weeks
- >6 – 8 weeks
- >8 – 10 weeks
- >10 – 12 weeks
- >12 weeks

Months of exposure to non-flare-up based treatment will be summarized with counts and percentages in the Non-Flare-up Treated Population according to the following categories:

- >0 – 3 months
- >3 – 6 months
- >6 – 9 months
- >9 – 12 months
- >12 months

Total exposure to study drug for each subject will be calculated as the number of days exposed to study drug. For subjects who received only flare-up based treatment, total exposure is the sum of the exposure for each flare-up. Note that this calculation excludes days between flare-ups. For subjects who received non-flare-up based treatment, total exposure is calculated from the date of first dose of study drug to the date of last dose of study drug, adjusting for any days of study drug interruption. Days of exposure will be summarized with counts and percentages of subjects in the Treated Population.

Dose reductions (no dose reductions, any dose reductions) will be summarized with counts and percentages of flare-ups and subjects in the Treated Population. Similarly, dose reduction categories (ie, 20 to 15 mg, 20 to 10 mg, etc.) will be summarized with counts and percentages.

The number of subjects who experienced an interruption of study drug during non-flare-up dosing and the number of subjects who discontinued study drug during non-flare-up dosing will be summarized with counts and percentages.

5.9 Compliance

Overall compliance to study drug will be calculated for each subject based on actual study drug taken (mg) and the expected study drug taken (mg) over the course of Part B.

Actual study drug taken will be calculated as the sum of study drug taken from each bottle where study drug taken from each bottle is the number of capsules taken multiplied by capsule strength.

Expected study drug taken will be calculated based on the protocol-defined treatment regimen. Subjects experiencing eligible flare-ups are expected to take 20 mg for 28 days followed by 10 mg for 56 days (or longer if treatment is extended). If the number of days of treatment is extended, weight-adjusted doses, and dose reductions should be taken into account in this calculation. Subjects in the Adult Cohort who are taking non-flare-up treatment are expected to take 5 mg each day they are not taking flare-up based treatment.

For example, if a subject in the Adult Cohort took 90 5-mg capsules from bottle 1, 28 20-mg capsules from bottle 2, and 50 10-mg capsules from bottle 3, then the actual study drug taken is equal to 90 x 5 + 28 x 20 + 50 x 10 = 1510 mg. The subject started non-flare-up based treatment on Study Day 1, started flare-up based treatment on Study Day 11, and completed Part B on Study Day 174. Treatment was not extended for this flare-up, and doses were not weight-adjusted or reduced. Therefore, the subject was expected to take 20 mg for 28 days followed by 10 mg for 56 days for the flare-up and 5 mg for 90 days (174 total days in Part B – 84 days of flare-up based treatment). Expected study drug taken = 90 x 5 + 28 x 20 + 56 x 10 = 1570 mg. Compliance would be calculated as: 1510 mg / 1570 mg = 96.2%.

Percent compliance will be summarized for the Treated Population with descriptive statistics. Categories of percent compliance (<80%, ≥80%) will be summarized for the Treated Population with counts and percentages of subjects.

6 Efficacy Data

Efficacy data for flare-up based treatment assessments will generally be presented by the following treatment groups for the Flare-up Population: PVO 20/10 mg, Prior Non-Flare-up/PVO 20/10 mg, and Combined PVO 20/10 mg. The primary efficacy analysis and other selected efficacy analyses will also be performed using the Flare-up PP Population.

Imaging Reads

Imaging data includes results from CT scan, x-ray, MRI, US, and WBCT scan. Results will be read and interpreted using the Primary Read process only.

The Primary Read review will be performed by two independent musculoskeletal radiologists during the study as images are made available. If there is sufficient agreement between the independent reviews, both will be used for analysis. If there is insufficient agreement between the independent reviews, an adjudication review will be performed and used for analysis. The definition of what constitutes agreement between the independent reviews is described in the PVO-1A202 FOP Independent Review Charter, version 10, dated 16Feb2017. Imaging procedures performed (categorized as "Yes", "Yes but not evaluable", or "No") will be summarized at scheduled visits with counts and percentages of subjects or flare-ups as appropriate.

6.1 Primary Efficacy Endpoint

The primary efficacy endpoint is the proportion of flare-ups with no new HO at Week 12 as assessed by CT scan (or by x-ray for subjects unable to undergo CT scan). Note that "no new HO at Week 12" means no new HO at Week 12 compared to baseline where baseline is Flare-up Day 1 and Week 12 is Week 12 for subjects who completed treatment at Week 12.

Similar analyses will be performed for subjects who had extended treatment. This data will be referenced as the EOT time point.

The primary efficacy endpoint will be presented for the Flare-up Population and the Flare-up PP Population for the following treatment groups: PVO 20/10 mg, Prior Non-Flare-up/PVO 20/10 mg, and Combined PVO 20/10 mg.

The presence of "HO at baseline" and the incidence of "No new HO", "New HO at Week 12", and "New HO at EOT" will be summarized with counts and percentages of flare-ups. The asymptotic 95% confidence intervals for the proportions will be presented.

Similar analyses will be performed for subgroups defined by baseline edema (missing, no edema, and any edema) and baseline edema severity (no edema, any edema, mild edema, moderate edema, severe edema, and moderate or severe edema [higher-intensity flare-ups]). See Section 6.2.5 for a description of edema determination.

6.2 Secondary Efficacy Endpoints for Flare-up Based Treatment

The secondary efficacy endpoints for Flare-up based treatment will be presented by treatment group (PVO 20/10 mg, Prior Non-Flare-up/PVO 20/10 mg, and Combined PVO 20/10 mg) for the Flare-up Population and the Flare-up PP Population.

Secondary efficacy endpoints for flare-up based treatment include:

1. Change from baseline in amount of bone formation (volume) as assessed by low-dose CT scan at Flare-up Week 12 and EOT.
2. Presence of soft tissue swelling and/or cartilage by MRI at Flare-up Week 12 and EOT; or presence of soft tissue swelling by US at Flare-up Week 12 and EOT in subjects unable to undergo MRI.
3. Change from baseline in active ROM measured by goniometer at the flare-up site at Flare-up Week 12 and EOT.

4. Change from baseline in ROM as assessed by CAJIS at Flare-up Week 12 and EOT.
5. Subject and Investigator global assessment of movement at the flare-up site at Flare-Up Week 12 and EOT.
6. Change from baseline in physical function using age-appropriate forms of the FOP-PFQ at Flare-up Weeks 4, 8, 12, and EOT.
7. Change from baseline in physical and mental health using age-appropriate forms of the PROMIS Global Health Scale at Flare-up Weeks 4, 8, 12, and EOT.
8. Change from baseline in the use of assistive devices and adaptations for daily living by FOP subjects at Flare-up Week 12 and EOT.
9. Change from baseline in cartilage, bone, angiogenesis, and inflammation biomarkers at Flare-up Weeks 4, 8, 12, and EOT.
10. Duration of active, symptomatic flare-up (start date and end date), as assessed by the subject and the Investigator.
11. WBCT Volume at Baseline, Month 12, and Change from Baseline to Month 12. Similar analyses will be performed for subjects with a new HO at Month 12 and for subjects with no new HO at Month 12.

Baseline for flare-up based assessments is Flare-up Day 1.

6.2.1 Volume of HO and HO Grades

The volume of HO and HO grades (defined below) will be determined from CT imaging.

The volume of HO at baseline will be calculated as the sum of the volumes for each lesion. The volume of *new* HO at post-baseline visits will be calculated as follows:

1. If there is no new HO, assign new volume = 0. (NVOL = 0)
2. If there is new HO
    a. Subtract the volume of baseline HO from the volume of post-baseline HO for each lesion (new volume at each lesion = $NVOL_1$, $NVOL_2$, … $NVOL_L$ where L is the number of lesions).
    b. If the volume of post-baseline HO < the volume of baseline HO for a given lesion i, impute new volume = 0 ($NVOL_i = 0$).
    c. Calculate the total volume of new HO ($NVOL = NVOL_1 + NVOL_2 + \ldots + NVOL_L$).

If an adjudication review is performed, the volume from the adjudication review will be used for analysis. Otherwise, the average of the volumes from the independent reviews will be used for analysis.

Volume of HO at baseline and volume of new HO at post-baseline visits will be summarized with descriptive statistics. Similar summaries will be provided for those subjects in the Flare-up Population who had flare-ups with new HO at any post-baseline visit.

HO lesions will be graded according to the following scale:

- Grade 0 – Imputed when there is no new HO
- Grade 1 – Fluid attenuation without evidence of calcification at CT

- Grade 2 – Calcification of soft tissues without evidence of bone formation
- Grade 3 – Immature bone formation
- Grade 4 – Mature bone with cortical differentiation

The HO grade at baseline will be calculated as the maximum HO grade across lesions. The *new* HO grade at post-baseline visits will be calculated as follows:

1. If there is no new HO, assign new HO grade = 0.
2. If there is new HO
    a. If there was no baseline HO, calculate the maximum HO grade across lesions.
    b. If there was baseline HO, calculate the maximum HO grade across lesions where there was no baseline HO.

HO grade at baseline and new HO grade at post-baseline visits will be summarized with counts and percentages of flare-ups.

6.2.2 Area of HO and HO Scores

The area of HO and HO scores will be determined from x-ray for subjects unable to undergo CT scan.

Area of HO at baseline will be calculated as the sum of the areas for each lesion. The area of *new* HO at post-baseline visits will be calculated using both anterior/posterior (AP) and lateral views as follows:

1. If there is no new HO, assign new area = 0 (NAREA = 0).
2. If there is new HO
    a. Subtract the area of baseline HO from the area of post-baseline HO for each lesion ($NAREA_1$, $NAREA_2$, … $NAREA_L$ where L is the number of lesions).
    b. If the area of post-baseline HO < the area of baseline HO for a given lesion i, impute new area = 0 ($NAREA_i = 0$).
    c. Calculate the total area of new HO ($NAREA = NAREA_1 + NAREA_2 + \ldots + NAREA_i$).

For each review, the maximum area from the AP or lateral view will be used for analysis.

If an adjudication review is performed, the area from the adjudication review will be used for analysis. Otherwise, the average of the areas from the independent reviews will be used for analysis.

Area of HO at baseline and area of new HO at post-baseline visits will be listed.

Area of HO and volume of HO (CT scan) will not be pooled. They will be presented in separate listings.

6.2.3 HO Lesion Scoring

HO scores will be determined from x-ray for subjects unable to undergo CT scan.

HO lesions will be scored according to the following scale:

- 0 – No HO

- 1 – Single or multiple spicules (punctate) or islands (non-contiguous) of HO
- 2 – Coalescing islands or reticular complexes of bone
- 3 – Single contiguous HO having longest dimension ≤1/2 the diameter of the reference normotopic bone in any projection
- 4 – Single contiguous HO having longest dimension >1/2 but ≤1 diameter of the reference normotopic bone in any projection
- 5 – Single contiguous HO having longest dimension >1 but ≤2 diameter of the reference normotopic bone in any projection
- 6 – Single contiguous HO having longest dimension >2 diameters of the reference normotopic bone in any projection

The HO score at baseline will be calculated as the maximum HO score across lesions. The *new* HO score at post-baseline visits will be calculated as follows:

1. If there is no new HO, impute new HO score = 0.
2. If there is new HO
    a. If there was no baseline HO, calculate the maximum HO score across lesions.
    b. If there was baseline HO, calculate the maximum HO score across lesions where there was no baseline HO.

HO score at baseline and new HO score at post-baseline visits will be listed.

HO lesion scores will not be pooled with HO lesion grades (CT scan) above. They will be described separately.

6.2.4 New or Enlarged HO and Severity of New HO

New HO lesions are those that did not exist at baseline. Enlarged HO lesions are lesions that existed at baseline but increased in volume by CT (or area by x-ray) post-baseline. The incidence of new HO lesions only, enlarged HO lesions only, or both new HO and enlarged HO lesions will be summarized at each scheduled visit with counts and percentages of flare-ups.

The severity of new HO (mild, moderate, severe) will be summarized at each scheduled visit with counts and percentages of flare-ups.

6.2.5 Presence of Soft Tissue Swelling (Edema) and/or Cartilage Formation, Severity of Edema, and Volume of Edema

Presence of edema and cartilage formation will be determined from MRI (or US for subjects unable to undergo MRI). Note that US does not include an assessment of cartilage formation.

If an adjudication review is performed, the severity and volume of edema from the adjudication review will be used for analysis. Otherwise, the maximum severity and the average volume of edema from the independent reviews will be used for analysis. Note also that when edema is assessed as mild with US, moderate will be imputed for analysis.

The presence of edema and the presence of cartilage formation will be summarized at each scheduled visit with counts and percentages of flare-ups.

The severity of edema (no edema, any edema, mild edema, moderate edema, severe edema, moderate or severe edema [higher intensity flare-ups]) will be summarized at each scheduled visit with counts and percentages of flare-ups.

The shift in severity of edema from baseline to each post-baseline visit will also be summarized with counts and percentages of flare-ups.

Volume of edema at each scheduled visit will be summarized with descriptive statistics.

6.2.6 Active Range of Motion

Range of motion at the primary and secondary joints will be assessed using a goniometer. The primary joint will be indicated on the eCRF and all other joints will be considered secondary.

Arcs of motion as defined in Table 1 will be calculated based on raw ROM measurements. If any of the contributing raw ROM measurements are missing, the arc of motion will not be calculated. Percent of normal arc of motion will be calculated as 100 x arc of motion / normal adult arc of motion.

Table 1. Arcs of Motion

Joint	Arc of Motion	Calculation Based on Raw ROM Measurements	Normal Adult Arc of Motion (degrees)
Hip	Extension/Flexion	Extension + Flexion	140
	Abduction/Adduction	Abduction + Adduction	60
	Total Arc of Motion	Extension + Flexion + Abduction + Adduction	200
Knee	Total Arc of Motion	Extension + Flexion	135
Shoulder	Total Arc of Motion	Flexion + Abduction-glenohumeral + Abduction-shoulder complex	490
Elbow	Extension/Flexion	Extension + Flexion	150
	Supination/Pronation	Supination + Pronation	160
	Total Arc of Motion	Extension + Flexion + Supination + Pronation	310
Wrist	Total Arc of Motion	Extension + Flexion	150
Ankle	Dorsiflexion/Plantarflexion	Dorsiflexion + Plantarflexion	70
	Subtalar Inversion/Subtalar Eversion	Subtalar Inversion + Subtalar Eversion	50
	Total Arc of Motion	Dorsiflexion + Plantarflexion + Subtalar Inversion + Subtalar Eversion	120
Cervical Spine	Extension/Flexion	Extension + Flexion	90
		Lateral	45
		Rotation	60
	Total Arc of Motion	Extension + Flexion + Lateral + Rotation	195
Jaw	Total Arc of Motion	Opening	35 mm
Lumbar Spine		Flexion	6 cm
		Extension	1 cm
	Total Arc of Motion	Flexion + Extension	7 cm

Source: Norken and White (2011)[1]

Percent of normal arc of motion values and changes from baseline will be summarized at each scheduled visit with descriptive statistics. Summaries will be provided separately for:

- Percent of normal total arc of motion at the primary joint (subjects have one total arc of motion at the primary joint)
- Percent of normal arc of motion at the primary joint by joint and arc of motion (subjects may have more than one arc of motion at the primary joint)
- Percent of normal arc of motion at the secondary joints by joint and arc of motion (subjects may have more than one arc of motion at one or more secondary joints)

Similar summaries for percent of normal total arc of motion will be provided for subgroups defined by presence of new HO at any post-baseline visit (new HO, no new HO).

6.2.7 Cumulative Analogue Joint Involvement Scale (CAJIS)

Range of motion across the whole body will also be assessed using the CAJIS.

The assessments will be performed on 12 joints (shoulder, elbow, wrist, hip, knee, and ankle on both the right and left sides), and three body regions (jaw, cervical spine [neck], and thoracic/lumbar spine). Each joint/region will be scored as: 0 = uninvolved; 1 = partially involved; and 2 = ankylosed/completely involved.

The CAJIS Total score will be calculated as the sum of the scores of all joints/regions and ranges from 0 (no involvement) to 30 (maximally involved).

The CAJIS Upper Extremities subscore will be calculated as the sum of the scores from six joints (shoulder, elbow, and wrist on both the right and left sides) and one region (cervical spine [neck]) and ranges from 0 (no involvement) to 14 (maximally involved).

The CAJIS Mobility subscore will be calculated as the sum of scores from six joints (hip, knee, and ankle on both the right and left sides) and ranges from 0 (no involvement) to 12 (maximally involved).

The CAJIS Total score, CAJIS Upper Extremities subscore, and CAJIS Mobility subscore and changes from baseline will be summarized at each scheduled visit with descriptive statistics. Similar summaries will be provided for subgroups defined by presence of new HO at any post-baseline visit (new HO, no new HO).

6.2.8 Global Assessment of Movement

Flare-up movement compared to baseline will be assessed by the subject (or parent proxy) and Investigator. Assessment results will be summarized at each scheduled visit with counts and percentages of flare-ups. Similar summaries will be provided for subgroups defined by presence of new HO at any post-baseline visit (new HO, no new HO).

6.2.9 FOP-Physical Function Questionnaire (FOP-PFQ)

Age-appropriate forms of the FOP-PFQ will be administered to subjects (or parent proxy). The adult form will be administered to subjects 15 years and older. For subjects between 8 and 14 years of age, both the pediatric self-completed and the pediatric proxy-completed forms will be administered. The proxy-completed form will be used for analyses unless only the self-completed form is available.

The FOP-PFQ consists of 28 questions on the adult form and 26 questions on the pediatric form scored on a scale from 1 to 5, with lower scores indicating more difficulty. The questions are described in detail in Appendices 7A, 7B, and 7C of the protocol.

The Total score will be calculated as:

- The sum of the scores from each question and ranges from 28x1 = 28 to 28x5 = 140 for the adult form.

- The sum of the scores from each question and ranges from 26x1 = 26 to 26x5 = 130 for the pediatric form.

The Upper Extremities subscore will be calculated as:

- The sum of the scores from 15 questions (questions 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 25, and 26) and ranges from 15x1 = 15 to 15x5 = 75 for the adult form.
- The sum of the scores from 18 questions (questions 1, 2, 6, 7, 8, 9, 10, 11, 16, 17, 18, 19, 21, 22, 23, 24, 25, and 26) and ranges from 18x1 = 18 to 18x5 = 90 for the pediatric form.

The Mobility subscore will be calculated as:

- The sum of the scores from 13 questions (questions 13, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 27, and 28) and ranges from 13x1 = 13 to 13x5 = 65 for the adult form.
- The sum of the scores from 8 questions (questions 3, 4, 5, 12, 13, 14, 15, and 20) and ranges from 8x1 = 8 to 8x5 = 40 for the pediatric form.

The Total score, Upper Extremities subscore, and Mobility subscore are defined based on information from PROMIS and Hays et al (2013)[2] for the adult form and a medical review of the questions for the pediatric form.

If a subject is missing some (but not more than 20%) of the contributing question scores, the Total score, Upper Extremities subscore, or Mobility subscore will be calculated as the average observed score multiplied by the number of expected question scores. For example, the Total score would be calculated as the average of the non-missing scores x 28 for the adult form. The determination for sufficient non-missing scores will be made independently for the Total score, Upper Extremities subscore, and Mobility subscore.

As the analysis for FOP-PFQ will be performed across all subjects (adult and pediatric) and the number of contributing questions differs, the scores will be transformed to reflect a percentage of worst score. The percentage of worst score ranges from 0% to 100% with 0% indicating the best possible function and 100% indicating the worst possible function. Table 2 illustrates some sample derivations of the percentage of worst score.

Table 2. Sample Derivations for the Percentage of Worst Score for the FOP-PFQ

Sample Subject	Observed FOP-PFQ Score	Lowest Possible Score	Highest Possible Score	Range of Possible Scores	Distance = Highest - Observed	Distance/ Range	Percentage of Worst Score
1	45	15	75	60	30	0.500	50.0%
2	40	15	75	60	35	0.583	58.3%
3	35	15	75	60	40	0.667	66.7%
4	30	15	75	60	45	0.750	75.0%
5	25	15	75	60	50	0.833	83.3%

Percentage of worst scores and changes from baseline will be summarized at each scheduled visit with descriptive statistics. Similar summaries will be provided for the percentage of worst Mobility Subscore for the subgroup of subjects with flare-ups in lower extremities and for the percentage of worst Upper Extremity Subscore for the subgroup of subjects with flare-ups in the upper extremities.

Note that flare-ups in the lower extremities include the following flare-up locations reported by the Investigator: abdomen, ankle or foot, distal lower extremities, hip, knee, lower back, and lumbar spine. Flare-ups in the upper extremities include the following flare-up locations reported by the Investigator: chest, cervical spine, head, jaw, distal upper extremities, elbow, shoulder, wrist or hand, thoracic spine, and upper back.

Similar summaries will be provided for subgroups defined by presence of new HO at any post-baseline visit (new HO, no new HO).

6.2.10 PROMIS Global Health

An age-appropriate form of the PROMIS Global Health short form will be administered to subjects (or parent proxy). The adult form will be administered to subjects 15 years and older. For subjects between 8 and 14 years of age, both the pediatric self-completed and the pediatric proxy-completed forms will be administered. The proxy-completed form will be used for analysis unless only the self-completed form is available.

The PROMIS Global Health short form consists of 10 questions on the adult form and nine questions on the pediatric form scored on varying scales. The questions are described in detail in Appendices 8A, 8B, and 8C of the protocol.

For the adult form, Global Physical Health and Global Mental Health scores will be calculated. The Global Physical Health score will be calculated as the sum of scores from questions 3, 6, 7, and 8, and will range from 4 (worse health) to 20 (better health). The Global Mental Health score will be calculated as the sum of scores from questions 2, 4, 5, and 10, and will range from 4 (worse health) to 20 (better health). In the calculation of the Global Physical Health and Global Mental Health scores, the following questions will be rescaled as shown in Table 3.

Table 3. Rescaled PROMIS Global Health Scale Scores

Question(s)	Score	Rescaled Score
7	0	5
	1-3	4
	4-6	3
	7-9	2
	10	1
8 and 10	1	5
	2	4
	3	3
	4	2
	5	1

If a subject is missing any of the contributing scores, the Global Physical Health or Global Mental Health score will not be calculated.

For the pediatric form, only a Total score will be calculated as the sum of scores from the first seven questions and will range from 7 (worse health) to 35 (better health). If a subject is missing some (but not more than three) of the contributing question scores, the Total score will be calculated as the average observed score multiplied by the number of expected question scores. For example, the Total score would be calculated as the average of the non-missing scores x 7.

Global Physical Health and Global Mental Health scores and Total scores will be converted to T-scores for analysis as described in detail in Appendix 1. A T-score of 50 is normal and increments of 10 +/- 1 standard deviation away from the norm. A T-score <50 indicates worse health, while a T-score >50 indicates better health. Note that higher values indicate better health.

T-scores and changes from baseline will be summarized at each scheduled visit with descriptive statistics. Data from adult and pediatric forms will be summarized separately. Similar summaries will be provided for subgroups defined by presence of new HO at any post-baseline visit (new HO, no new HO).

6.2.11 Flare-up Pain and Swelling

Adults and pediatric subjects 8 years and older will rate the pain associated with their flare-ups using a numeric rating scale (NRS) from 0 (no pain) to 10 (worst pain ever experienced). Pediatric subjects less than 9 years old will rate pain using the Faces Pain Scale – Revised (FPS-R), which ranges from 0 (no pain) to 10 (very much pain) in 2-point increments. Swelling associated with flare-ups will be rated on a scale from 0 (no swelling) to 10 (worst swelling ever experienced).

Ratings for pain and swelling and changes from baseline will be summarized at each scheduled visit with descriptive statistics.

6.2.12 Assistive Devices and Adaptations for Daily Living

Subjects will be given a list of FOP assistive devices and adaptations grouped into 12 categories and asked to select those they use for daily living.

The use of assistive devices and adaptations will be summarized overall and by category at each scheduled visit with counts and percentages of flare-ups. The number of devices and adaptations used will be summarized overall and by category at each scheduled visit with descriptive statistics.

The use of new-onset assistive devices and adaptations will be summarized overall and by category with counts and percentages of flare-ups. New-onset assistive devices and adaptations are defined as assistive devices and adaptations used at a post-baseline visit that were not used at baseline.

6.2.13 Bone, Cartilage, Angiogenesis, and Inflammation Biomarkers

Blood and urine samples will be collected for cartilage, bone, angiogenesis, and inflammation. Bone and cartilage biomarkers include osteocalcin, bone specific alkaline phosphatase (B-ALP), procollagen type I C-terminal propeptide (PICP), procollagen 1 N-terminal propeptide (P1NP), CD retinoic acid-sensitive protein (CD-RAP), and C-terminal telopeptide (CTX). The angiogenesis biomarker is the ratio of fibroblast growth factor and urine creatinine. Inflammation biomarkers include erythrocyte sedimentation rate (ESR), C reactive protein (CRP), interleukin 6 (IL-6), interleukin 1 beta (IL-1 beta), tumor necrosis factor (TNF), creatine kinase (CK), and lactate dehydrogenase (LDH).

Biomarkers values and changes from baseline will be summarized at each scheduled visit with descriptive statistics. Note that results obtained locally rather than from the central laboratory will not be used to calculate change from baseline and will be excluded from these summaries. Results captured as "<" or ">" will be rounded up or down accordingly for analysis (eg, "<0.61" will be rounded to 0.60).

Abnormal values at baseline and new-onset abnormal values post-baseline will be summarized with counts and percentages of flare-ups. New-onset abnormal values are defined as values from biomarkers that were normal at baseline and abnormal post-baseline or were abnormal at baseline and abnormal in the opposite direction post-baseline (eg, abnormal low at baseline and abnormal high post-baseline). Note that results obtained locally will be included in these summaries.

The following figures will be generated (note that the unit of analysis is each flare-up):

- Mean biomarker values over time (one figure per biomarker, each figure with one line).
- Mean biomarker values over time by incidence of new HO (one figure per biomarker, each figure with one distinctive line per subgroup [new HO, no new HO]).
- Mean biomarker values over time by baseline edema status (one figure per biomarker, each figure with one distinctive line per subgroup [baseline edema present, baseline edema not present]).

- Biomarker values over time by flare-up (one figure per biomarker, each figure with one distinctive line per flare-up). Flare-ups will be identified in the legend with subject ID, flare-up identifier (eg, Flare A, Flare B), gender, and age.
- Mean percent of normal high biomarker values over time (one figure per biomarker, each figure with one line). Percent of normal high = 100 x value / upper limit of normal range.
- Mean percent of normal high biomarker values over time by incidence of new HO (one figure per biomarker, each figure with one distinctive line per subgroup [new HO, no new HO]).
- Mean percent of normal high biomarker values over time by baseline edema status (one figure per biomarker, each figure with one distinctive line per subgroup [baseline edema present, baseline edema not present]).
- Percent of normal biomarker values over time by flare-up (one figure per biomarker, each figure with one distinctive line per flare-up). Flare-ups will be identified in the legend with subject ID, flare-up identifier, gender, and age.

6.2.14 Duration of Active Symptomatic Flare-up

Subjects will keep a daily diary during the Flare-up Component where the question "Is your flare-up ongoing today?" is answered. Subjects with at least 80% of diary data (at least 67 days with non-missing data) will be used for analysis.

The number of symptomatic days will be calculated as the number of days where the diary question is answered "Yes". The number of symptomatic days will be summarized with descriptive statistics.

The status of the flare-up will be determined at Week 12 and EOT. If there are no symptomatic days recorded in the diary, the flare-up will be considered "Resolved prior to the first dose of study drug". If the last symptomatic day recorded in the diary is on or before the specified visit and is followed by at least one asymptomatic day (ie, diary question is answered "No"), the flare-up will be considered "Resolved". If the last symptomatic day recorded is on the last diary entry, the flare-up will be considered "Not resolved". The status of the flare-up will be summarized at Week 12 and EOT with counts and percentages of flare-ups.

The days to flare-up resolution will be calculated as follows:

- If the flare-up was considered "Resolved prior to the first dose of study drug", days to flare-up resolution = 0.
- If the flare-up was considered "Resolved", days to flare-up resolution = date of flare-up resolution – date of first dose + 1, where the date of flare-up resolution is the last symptomatic day.
- If the flare-up was considered "Not resolved", days to flare-up resolution = date of last diary entry – date of first dose + 1.

The days to flare-up resolution will be summarized with descriptive statistics.

6.3 Secondary Efficacy Endpoints for Non-Flare-up Based Treatment

Efficacy data for non-flare-up based treatment assessments will generally be presented in total for the WBCT Population.

Secondary efficacy endpoints for non-flare-up based treatment include:

1. Change from baseline in whole body burden of HO as assessed by low-dose WBCT scan, excluding head, at Study Months 12 and 24.
2. Change from baseline in ROM as assessed by CAJIS at Study Months 12 and 24.
3. Change from baseline in physical function using age-appropriate forms of the FOP-PFQ at Study Months 12 and 24.
4. Change from baseline in physical and mental health using age-appropriate forms of the PROMIS Global Health Scale at Study Months 12 and 24.
5. Change from baseline in the use of assistive devices and adaptations for daily living by FOP subjects at Study Months 12 and 24.
6. Change from baseline in biomarkers at Study Months 12 and 24.

Baseline is Part B Screening.

6.3.1 Whole Body Burden of HO

Whole body burden (volume) of HO will be determined from the WBCT scan. HO will be assessed at 15 locations including head and neck, left and right elbow, left and right hip, left and right knee, left and right shoulder, left and right distal lower extremities, left and right distal upper extremities, lower spine/abdomen, and upper spine/chest.

Whole body burden of HO will be calculated as the sum of the volumes for each location. If there is no HO, 0 will be assigned for analysis.

Volume of HO and change from baseline (total and by region) will be summarized for subjects with any new HO at Month 12 and no new HO at Month 12, with descriptive statistics by the following subsets: Treated Flare-ups, Under/Untreated Flare-ups, and No Flare-ups.

6.3.2 Volume of HO

The change from baseline in the amount of new HO formed at the original flare-up site will be summarized and provided in a listing.

7 Safety Data

Adverse events will generally be presented by treatment group (PVO 20/10 mg only (including weight-adjusted doses); PVO 20/10 mg (including weight-adjusted doses) with prior non-flare-up treatment; and non-flare-up treatment only) and in total for the Enrolled and Treated Populations. The treatment groups will be based on when the AE occurred and will corresponding to the following groups: No treatment (prior to any treatment or between flares if not on non-flare-up dosing), Non-flare-up treatment, and Flare-up treatment PVO 20/10 mg (including weight-adjusted doses).

Safety data for flare-up based treatment assessments will generally be presented by cohort and in total for the Flare-up Population. Baseline is Flare-up Day 1 (the date of first dose of flare-up treatment).

Safety data for non-flare-up based treatment assessments will generally be presented in total for the Non-Flare-up Treated Population (the date of first dose of non-flare-up treatment; Non-flare-up Day 1). Baseline is Part B Screening/Baseline.

7.1 Adverse Events

Adverse events will be recorded from the time informed consent (Part B Study Day 1) is signed through study completion. All AEs will be coded to body system and preferred term using MedDRA Version 17.0. At least possibly related AEs are defined as AEs assessed by the Investigator as possibly, probably, or definitely related to study drug. Adverse events that require Common Terminology Criteria for Adverse Events (CTCAE) grading (Version 4.03, 14 June 2010) as indicated on the eCRF are considered to be retinoid-associated AEs.

Adverse events with start date on or after Flare-up Day 1 and on or before the date of completion for the Flare-up treatment will be attributed to the Flare-up Treatment Period. Adverse events with start date on or after Non-Flare-up Day 1 for which the patient is not undergoing flare-up treatment will be attributed to the Non-Flare Treatment Period. Adverse events with start date on or after Study Day 1 that are not attributed to the flare-up or non-flare-up treatments will be attributed to the Non-Treatment Period.

For the both the Enrolled Population and the Treated Population, AEs will be summarized by period (Non-Treatment Period, Non-Flare Treatment Period, Flare-up Treatment Period) and in total.

The incidence of the following AEs will be summarized with counts and percentages of subjects (overall and by MedDRA body system and preferred term):

- Any AEs
- Any AEs by maximum severity
- Any AEs at least possibly related to study drug
- Any AEs at least possibly related to study drug by maximum severity
- Any retinoid-associated AEs
- Any retinoid-associated AEs at least possibly related to study drug
- Any retinoid-associated AEs by maximum CTCAE grading
- Any retinoid-associated AEs by maximum severity
- Any retinoid-associated AEs by seriousness
- Any serious adverse events (SAEs)
- Any severe/life threatening retinoid-associated AE by maximum CTCAE grading
- Any AEs leading to dose modification or interruption of study drug
- Any AEs leading to discontinuation of study drug
- Any AEs leading to discontinuation of study

Data listings will be provided for (1) all AEs, (2) SAEs, (3) AEs leading to dose modification or interruption of study drug, (4) AEs leading to discontinuation of study drug, and (5) AEs leading to discontinuation of study. Part B Study Day and/or Flare-up Day for AE start and stop dates will be included on AE listings as appropriate.

7.1.1 Number of Flare-ups Per Subject-Month

Flare-ups are counted using the number of subject/Investigator-reported flare-ups and AEs with the 'condition aggravated" preferred term that are medically reviewed and classified as unique flare-ups (not treated with palovarotene and not recorded as a subject/Investigator- reported flare-up). Rates will be calculated by dividing the total number of flare-ups by the total subject months of follow-up.

7.2 Safety Laboratory Assessments

Values and changes from baseline will be summarized at each scheduled visit with descriptive statistics. Both the change from Screening/Baseline (patient based) and the change from Flare-up Day 1 (flare-up based) will be presented. Mean values (+/- standard error) will be plotted over time for all laboratory parameters and biomarkers. If multiple records occur within a post-baseline visit, the first non-missing value will be used. Note that results obtained locally rather than from the central laboratory will not be used to calculate change from baseline and will be excluded from these summaries.

Potentially clinically significant values at baseline and new-onset potentially clinically significant (PCS) values post-baseline will be summarized with counts and percentages of subjects or flare-ups as appropriate. Separate analyses will be performed for the non-flare-up treatment visits and the flare-up treatment visits. New-onset PCS values are defined as values that were not PCS at baseline and were PCS post-baseline. Note that results obtained locally will be included in these summaries. PCS values are defined in Table 4.

Table 4. Potentially Clinically Significant Safety Laboratory Values

	PCS Low	PCS High
Biochemistry		
Aspartate aminotransferase	N/A	3 x ULN
Alanine aminotransferase	N/A	3 x ULN
Amylase	N/A	3 x ULN
Lipase, pancreatic	N/A	3 x ULN
Bilirubin	N/A	>2 mg/dL
Thyroxine	<4.0 mcg/dL	>13.0 mcg/dL
Lipids		
Cholesterol	N/A	>300 mg/dL
Triglycerides	N/A	>400 mg/dL
Hematology		
Leukocytes	<2.8 x 10^9/L	>16.0 x 10^9/L
Hemoglobin (females)	<9.5 g/dL	>17.5 g/dL
Hemoglobin (males)	<11.5 g/dL	>19.0 g/dL
Hematocrit (females)	<32%	>54%
Hematocrit (males)	<37%	>60%
Platelets	<75 x 10^9/L	>700 x 10^9/L

A listing will be provided for all subjects with PCS lipase values that will include all lipase results during the study along with associated palovarotene and glucocorticoid dosing data (dose and start and stop dates). A by-subject listing of all laboratory parameters will be generated.

7.3 Vital Signs

Vital signs include blood pressure, heart rate, respiratory rate, and temperature.

Values and changes from baseline will be summarized at each scheduled visit with descriptive statistics. Both the change from Screening/Baseline (patient base) and the change from Flare-up Day 1 (flare-up based) will be presented. Mean values (+/- standard error) will be plotted over time for blood pressure and heart rate. If multiple records occur within a post-baseline visit, the first non-missing value will be used.

Potentially clinically significant values at baseline and new-onset PCS values post-baseline will be summarized with counts and percentages of subjects or flare-ups as appropriate. Separate analyses will be performed for the non-flare-up treatment visits and the flare-up treatment visits. New-onset PCS values are defined as values that were not PCS at baseline and were PCS post-baseline. PCS values are defined in Table 5.

Table 5. Potentially Clinically Significant Vital Sign Values

	PCS Low	PCS High
Systolic blood pressure	<86 mmHg or a decrease of ≥25 mmHg from baseline	>180 mmHg or an increase of ≥25 mmHg from baseline
Diastolic blood pressure	<48 mmHg or a decrease of ≥20 mmHg from baseline	>110 mmHg or an increase of ≥20 mmHg from baseline
Heart rate	<45 bpm or a decrease of ≥20 bpm from baseline	(1) >105 bpm AND an increase of ≥20 bpm from baseline (2) >125 bpm

mmHg = millimeters of mercury, bpm = beats per minute.

7.4 Body Weight

Values and changes from baseline will be summarized at each scheduled visit with descriptive statistics. Mean values (+/- standard error) will be plotted over time.

7.5 Knee and Hand/Wrist Radiograph

Subjects under the age of 18 years at the time of enrollment into Study PVO-1A-202 with open epiphyseal growth plates will undergo knee and hand/wrist radiographs (AP view) for assessment of epiphyseal growth plate.

Knee and hand/wrist radiographs will be performed remotely (eg, at a local medical facility) for subjects who are unable or unwilling to attend site visits. Once a subject has achieved 100% skeletal maturity (as confirmed by radiography), knee and hand/wrist radiographs will no longer be required. Summary statistics of bone age and chronological age (values and changes from baseline) will be calculated at each scheduled visit. Changes in bone age will be plotted over time.

Summary tables of growth plate abnormalities identified in radiographs will include the following:

- Number and percent of subject with epiphyseal growth plates open (by hand/wrist and/or knee).
- Number and proportion of subjects with any abnormality at each visit, for each specific abnormality (e.g. growth recovery lines, Sclerosis, Under mineralization, etc.).
- Number and proportion of subjects with any change from the prior timepoint for hand/wrist and knee by each abnormality for each visit.
- Number and proportion of subjects with any change from the prior timepoint for hand/wrist and knee by each abnormality across all post-screening visits. For summaries across all timepoints, the worst case for a given subject will be used in the analysis.
- Number and proportion of subjects with any change from the prior timepoint for hand/wrist and knee across all abnormalities and post-screening visits. For summaries across all abnormalities and timepoints, the worst case for a given subject will be used in the analysis.

7.6 Linear Growth Assessment

A linear growth assessment will be performed for subjects under the age of 18 years at the time of enrollment into Study PVO-1A-202 with open epiphyseal growth plates using stadiometry and knee height assessments.

7.7 Electrocardiogram

Values and changes from baseline will be summarized for continuous ECG parameters at each scheduled visit with descriptive statistics. Both the change from Screening/Baseline (patient based) and the change from Flare-up Day 1 (flare-up based) will be presented. If multiple records occur within a post baseline visit, the first non-missing value will be used.

Abnormal ECG categories will be assigned based on raw findings as shown in Table 6.

Table 6. Abnormal Electrocardiogram Categories

Abnormal ECG Category	Raw Findings
Rhythm	* Atrial premature complexes * Ventricular premature complexes * Sinus arrhythmia * Sinus tachycardia
Conduction	* Nonspecific intraventricular conduction delay * Right bundle branch block * Incomplete right bundle branch block * Left posterior fascicular block * Left bundle branch block * Incomplete left bundle branch block
Morphology/Chamber enlargement	* Left ventricular hypertrophy * Left ventricular hypertrophy with repolarization abnormality
Axis deviation	* Right axis deviation * Left axis deviation
Myocardial infarction	* Pathologic Q waves * Acute ST elevation MI * Acute non-ST elevation MI
ST Segment/T Waves/U Waves	* Early repolarization * Nonspecific ST and T wave abnormality * Nonspecific ST elevation * Nonspecific U wave abnormalities
Miscellaneous	* Long QT interval * Short PR interval

Overall interpretation and abnormal ECG categories will be summarized at each scheduled visit with counts and percentages of flare-ups. New-onset ECG abnormalities are defined as

abnormalities that did not exist at baseline. New-onset ECG abnormalities will be summarized counts and percentages of subjects or flare-ups as appropriate.

Incidence of QTc >450 msec, QTc >500 msec, or change from baseline in QTc >60 msec will be summarized with counts and percentages of subjects or flare-ups as appropriate.

Potentially clinically significant values at baseline and new-onset PCS values post-baseline will be summarized with counts and percentages of subjects or flare-ups as appropriate. New-onset PCS values are defined as values that were not PCS at baseline and were PCS post-baseline. PCS values are defined in Table 7.

Table 7. Potentially Clinical Significant Electrocardiogram Values

	PCS Low	PCS High
PR interval	N/A	>200 msec or an increase from baseline of ≥20 msec
QRS interval	N/A	>100 msec or an increase from baseline of ≥10 msec
QT interval, QTcF, QTcB	N/A	>500 msec or an increase from baseline of ≥60 msec

msec = millisecond, QTcF = rate-corrected QT interval using Fridericia's formula (QT interval divided by the cube root of the RR interval), QTcB = rate-corrected QT interval using Bazett's formula (QT interval divided by the square root of the RR interval)

7.8 Physical Examination

Physical examination results will be summarized at each scheduled visit by body system with counts and percentages of subjects or flare-ups as appropriate.

7.9 Columbia-Suicide Severity Rating Scale

Suicidal ideation and behavior will be assessed using the Columbia-Suicide Severity Rating Scale (C-SSRS).

Data listings will be provided for (1) all C-SSRS results; and (2) for subjects reporting any type 4 or type 5 suicidal ideations.

8 Pharmacokinetics Data

Pharmacokinetics will be assessed within 7 days of flare-up based treatment initiation and at the end of treatment for the first treated flare-up. Blood samples will be taken at 0, 3, 6, 10, and 24 hours post-dose.

Plasma palovarotene concentrations will be summarized with descriptive statistics (number of non-missing values, arithmetic mean, standard deviation, standard error, coefficient of variation, geometric mean, median, minimum, and maximum). All plasma concentrations below the limit of quantification (LOQ) will be set to zero for the purpose of calculating descriptive statistics. If

at any time point, 1/3 or more of subjects have <LOQ values, descriptive statistics will not be calculated at that time point. PK concentrations will be imputed using the following:

- If 0-hour concentration missing, then 24-hour concentration used to replace the missing concentration.
- If 24-hour concentration missing, then 0-hour concentration used to replace the missing concentration.
- Invalid 24-hour concentrations were replaced with the 0-hour concentrations. Twenty-four-hour concentrations were determined to be invalid if concentration is greater than the 10-hour value (i.e. assumed to have been obtained after next day dose instead of before next day dose).

If any other blood sample collections are missing, PK parameters will not be assessed. Actual sampling times will be used for the PK analysis, if available. Nominal sampling times will be used for descriptive statistics.

Individual concentration-time profiles for palovarotene will be presented (linear and semi-logarithmic scales). Arithmetic means (+/- standard deviation) concentration-time profiles will also be presented (linear and semi-logarithmic scales).

An exploration of the effects of body weight on PK will be performed as the data allow.

The PK analysis of palovarotene will be conducted using model-independent methods as implemented in WinNonlinTM. The PK variables analyzed will include $AUC_{0\text{-}24(ss)}$, $C_{max(ss)}$, $C_{min(ss)}$, $T_{max(ss)}$, CL/F, λ_z, and $t_{1/2z}$ at steady-state. The PK variables are defined as follows:

$AUC_{0\text{-}24(ss)}$	Area under the concentration versus time curve over the 24-hour dosing interval; calculated using linear trapezoid rule.
$C_{max(ss)}$	Maximum or peak measured plasma concentration at steady-state.
$C_{min(ss)}$	Minimum or trough measured plasma concentration at steady-state.
$T_{max(ss)}$	Time of maximum or peak measured plasma concentration at steady-state, obtained by inspection.
CL/F	Clearance defined as: Dose/$AUC_{0\text{-}24(ss)}$.
λ_z	Terminal rate constant
$t_{1/2z}$	Apparent terminal elimination half-life; calculated as $\ln(2)/\lambda_z$. The number of data points included in the regression will be determined by visual inspection, but a minimum of three data points in the terminal phase, excluding C_{max}, will be required to estimate λ_z.

9 Changes from Protocol

For the purposes of clarity and consistency of nomenclature, the Flare-up Rate Population described in the protocol was renamed as the Non-Flare Treated Population.

The initial secondary objective listed in Section 2.1.2 was not included in Amendment 3 of the protocol for Study PVO-1A-202 Part B or Amendment 1 of the protocol for Study PVO-1A-204.

The protocol describes the primary comparison of outcomes to be from data collected from the NHS and Phase 2 study. These comparisons will be performed in a future analysis and the analysis described here will instead focus on comparisons solely from Study PVO-1A-202 Part B between treatment groups of PVO 20/10 mg and Prior Non-Flare-up/PVO 20/10 mg, for flare based analyses; and treatment groups of treated, untreated, or no-flares, for WBCT based analyses.

10 References

1. Norken CC, White DJ. *Measurement of Joint Motion: A Guide to Goniometry*. 4th ed. Philadelphia, PA: F.A. Davis Company; 2011.
2. Hays RD, Spritzer KL, Amtmann D, et al. Upper-extremity and mobility subdomains from the Patient-Reported Outcomes Measurement Information System (PROMIS) adult physical functioning item bank. *Arch Phys Med Rehabil.* 2013;94(11):2291-2296. doi:10.1016/j.apmr.2013.05.014.
3. D'Agostino, RB., Jr., Tutorial in Biostatistics: Propensity Score Methods for Bias Reduction in the Comparison of a Treatment to a Non-Randomized Control Group. Statistics in Medicine, 1998, 17, 2265-2281.
4. Chumlea WC, Guo SS, Steinbaugh ML. Prediction of stature from knee height for black and white adults and children with application to mobility-impaired or handicapped persons. *J Am Diet Assoc*. 1994;94(12):1385-1388, 1391.

Appendices

Appendix 1. PROMIS T-Score Conversions

The following conversion tables allow a user to convert Global Physical Health, Global Mental Health, and Total scores into T-scores. T-score distributions are standardized such that a 50 represents the average (mean) for the US general population, and the standard deviation around that mean is 10 points. A high score always represents more of the concept being measured. Thus, a subject who has a T-score of 60 is one standard deviation better (more healthy) than the general population.

Adult Global Physical Health		
Raw Score	**T-Score**	**Standard Error**
4	16.2	4.8
5	19.9	4.7
6	23.5	4.5
7	26.7	4.3
8	29.6	4.2
9	32.4	4.2
10	34.9	4.1
11	37.4	4.1
12	39.8	4.1
13	42.3	4.2
14	44.9	4.3
15	47.7	4.4
16	50.8	4.6
17	54.1	4.7
18	57.7	4.9
19	61.9	5.2
20	67.7	5.9

Adult Global Mental Health		
Raw Score	**T-Score**	**Standard Error**
4	21.2	4.6
5	25.1	4.1
6	28.4	3.9
7	31.3	3.7
8	33.8	3.7
9	36.3	3.7
10	38.8	3.6
11	41.1	3.6
12	43.5	3.6
13	45.8	3.6
14	48.3	3.7
15	50.8	3.7
16	53.3	3.7
17	56.0	3.8
18	59.0	3.9
19	62.5	4.2
20	67.6	5.3

Pediatric Self-Completed Total		
Raw Score	**T-Score**	**Standard Error**
7	16.0	3.4
8	17.1	3.6
9	18.3	3.7
10	19.7	3.8
11	21.2	3.8
12	22.8	3.7
13	24.4	3.6
14	26.1	3.6

Pediatric Proxy-Completed Total		
Raw Score	**T-Score**	**Standard Error**
7	14.7	2.9
8	15.3	3.1
9	16.0	3.2
10	16.9	3.4
11	18.1	3.6
12	19.4	3.7
13	21.0	3.8
14	22.7	3.8

Pediatric Self-Completed Total		
Raw Score	**T-Score**	**Standard Error**
15	27.6	3.5
16	29.2	3.5
17	30.8	3.5
18	32.4	3.6
19	34.0	3.6
20	35.6	3.6
21	37.2	3.6
22	38.8	3.6
23	40.4	3.6
24	42.1	3.7
25	43.9	3.7
26	45.7	3.6
27	47.5	3.6
28	49.2	3.6
29	51.1	3.7
30	53.3	3.9
31	55.7	4.2
32	58.3	4.5
33	61.1	4.9
34	64.2	5.4
35	67.5	6.1

Pediatric Proxy-Completed Total		
Raw Score	**T-Score**	**Standard Error**
15	24.4	3.7
16	26.1	3.7
17	27.7	3.7
18	29.4	3.8
19	31.2	3.8
20	32.9	3.8
21	34.6	3.8
22	36.2	3.8
23	37.9	3.9
24	39.7	4.0
25	41.7	4.0
26	43.6	3.9
27	45.4	3.8
28	47.3	3.9
29	49.3	4.1
30	51.8	4.4
31	54.5	4.7
32	57.3	5.0
33	60.2	5.4
34	63.2	6.0
35	66.1	6.5

Statistical Analysis Plan

A Phase 2, Open-Label Extension, Efficacy and Safety Study of a RARγ-Specific Agonist (Palovarotene) in the Treatment of Preosseous Flare-ups in Subjects with Fibrodysplasia Ossificans Progressiva (FOP)

Study Number: PVO-1A-202 Part C

Original Protocol:	04 June 2014
Amendment 1:	30 September 2014
Amendment 2:	22 June 2015
Amendment 3:	10 March 2016
Amendment 4:	01 September 2017
Amendment 5:	06 June 2018
Amendment 6:	08 March 2019
Amendment 7:	01 November 2019
Amendment 8:	30 November 2020

Statistical Analysis Plan Version:	1.1
Statistical Analysis Plan Date:	16 December 2022

Clementia Pharmaceuticals Inc.
1000 de la Gauchetiere, suite 1200
Montreal, QC H3B 4W5 Canada

Statistical Analysis Plan Signature Page

A Phase 2, Open-Label Extension, Efficacy and Safety Study of a RARγ-Specific Agonist (Palovarotene) in the Treatment of Preosseous Flare-ups in Subjects with Fibrodysplasia Ossificans Progressiva (FOP)

Protocol Number: PVO-1A-202 Part C

Signature of Approval for Statistical Analysis Plan

Approved by:

Ipsen Pharmaceuticals, Inc. PPD	
NAME: PPD	
SIGNATURE:	**DATE:** Dec 17, 2022
Ipsen Pharmaceuticals, Inc. PPD	
NAME: PPD	
SIGNATURE:	**DATE:** Dec 16, 2022

Table of Contents

List of Tables

List of Abbreviations

Abbreviation	Definition
AE	adverse event
ATC	anatomic therapeutic class
AP	anterior/posterior
CAJIS	Cumulative Analogue Joint Involvement Scale
C-SSRS	Columbia-Suicide Severity Rating Scale
CSR	Clinical Study Report
CT	computed tomography
CTCAE	Common Terminology Criteria for Adverse Events
ECG	electrocardiogram
eCRF	electronic case report form
FOCBP	Female subjects of childbearing potential
FOP	Fibrodysplasia Ossificans Progressiva
FOP-PFQ	FOP-Physical Function Questionnaire
GEE	generalized estimating equation
HO	heterotopic ossification
LPLV	last patient last visit
MedDRA	Medical Dictionary for Regulatory Activities
NDA	New Drug Application
NHS	Natural History Study
PCS	potentially clinically significant
PROMIS	Patient-Reported Outcomes Measurement Information System
RARγ	retinoic acid receptor gamma
ROM	range of motion
SAE	serious adverse event
SAP	statistical analysis plan
ULN	upper limit of normal
US	United States
WBCT	whole body computed tomography
wLME	weighted linear mixed effect model

1 Introduction

1.1 Context for Version 1.1

This statistical analysis plan amendment is introduced in preparation for the writing of the final CSR, which will include summaries of data collected after the start of the prolonged palovarotene dosing interruption that started in late 2019/early 2020 described below and up to LPLV in late 2022. This includes data from subjects who restarted palovarotene after the interruption and subjects who did not restart. The final CSR is not intended to include a rerun of tables and content already included in the interim CSRs from 2021 and 2022.

1.2 Context for Version 0.3 (Final)

This Statistical Analysis Plan (SAP) describes the analyses to be performed for Study PVO-1A-202 Part C based on Amendment 4 of the protocol dated 01 September 2017, and all subsequent amendments; along with the France-specific Amendment 2 of the protocol for Study PVO-1A-204 dated 02 November 2017, and all subsequent amendments. Data collected for Study PVO-1A-202 Part C and Study PVO-1A-204 will be combined for the analyses. For the purposes of this SAP, "Study PVO-1A-202 Part C" (or simply "Part C") represents data from both studies.

The analyses performed for Study PVO-1A-202 Parts A and B are described in separate SAPs. PK analysis is described in this SAP. However, the PK analysis will be performed separately.

The focus of this document is the analyses to be included in an interim CSR for the palovarotene NDA, which will not include summaries of all efficacy endpoints. This SAP may be amended to include additional details of analyses to be performed for the final CSR based on the locked Part C data.

As WBCT and X-Ray bone safety data are collected sequentially from Study PVO-1A-201 through PVO-1A-202C, the combination of bone safety analysis will be provided in the integrated safety analysis. However, listing of bone safety for Part C will be provided in this SAP.

This amendment is introduced after the second interim analysis in Study PVO-1A-301 ("Study 301") but before the planned third interim analysis (anticipated to be held in mid-May 2020). The futility boundary for Study 301 was crossed at the second interim analysis, signifying that the prespecified Bayesian compound Poisson distribution model with square-root transformation indicated that the trial was unlikely to show sufficient evidence of benefit. The independent data monitoring committee (DMC) informed the sponsor of this outcome on January 15, 2020, and the efficacy assessments were subsequently unblinded to the sponsor.

Based on additional post-hoc analyses presented to the DMC on January 21, 2020, the DMC noted that using the square-root transformation of the data in the primary analysis appears to have moved the statistical conclusion from significant therapeutic benefit to showing futility of the treatment. The DMC also noted that the dilemma created by these highly disparate results precludes a confident conclusion about futility. As such, the DMC recommended that additional analyses be conducted and provided to them for consideration at the next interim analysis.

US FDA instituted a partial clinical hold on dosing of palovarotene due to premature physeal closure (PPC) for subjects <14 years of age on December 4, 2019; as of the finalization of this SAP amendment, dosing

has not yet resumed for these subjects. Dosing on all other subjects in Study 301 and Study 202 Part C not subject to US FDA's partial clinical hold was interrupted following the January 21, 2020 DMC meeting while the protocol amendment allowing the continuation of the study after the futility outcome and incorporating the additional analyses requested by the DMC was reviewed by regulatory agencies, IRBs, and investigational sites. These subjects had palovarotene dosing interrupted on or around January 24, 2020. The sponsor allowed dosing to resume as of March 26, 2020, provided each individual site had obtained Ethics Committee approval to do so and the site was able to fulfil regulatory and operational requirements. Subjects were expected to restart dosing over a span of weeks due to these logistical considerations.

2 Overview

2.1 Study Objectives

2.1.1 Primary Objective

The primary objective is:

- To evaluate the safety and efficacy of different palovarotene dosing regimens in subjects with FOP. Efficacy will be based on the ability of palovarotene to prevent heterotopic ossification (HO) as assessed by low-dose whole body computed tomography (WBCT) scan, excluding head.

2.1.2 Secondary Objectives

- To evaluate the effect of palovarotene on range of motion (ROM) as assessed by the Cumulative Analogue Joint Involvement Scale for FOP (CAJIS).
- To evaluate the effect of palovarotene on physical function using age-appropriate forms of the FOP-Physical Function Questionnaire (PFQ).
- To evaluate the effect of palovarotene on physical and mental health using age-appropriate forms of the Patient Reported Outcomes Measurement Information System (PROMIS) Global Health Scale.
- To evaluate the pharmacokinetics of palovarotene.
- Part D: To implement safety measures recommendations in order to ensure that assessments of safety continue for up to 2 years post last dose of study treatment for skeletally immature subjects.

2.2 Study Design

PVO-1A-202 Part C is a Phase 2, multicenter, open-label extension of Study PVO-1A-202 Part B. This continuation of the open-label extension study will give subjects who participated in Part B the opportunity to be followed for up to an additional 36 months. There will be no new subjects in Part C.

Part D annual post last dose of study treatment assessments for up to 2 years will be obtained in subjects who were skeletally immature at the time of study treatment discontinuation in order to obtain longer-term safety data. No new subjects will be enrolled in Part D. Subjects will follow assessments outlined in Table 4 of the protocol as long as they are not 100% skeletally mature. Part C plus Part D total duration will not exceed 48 months.

Added assessments for spinal health will be carried out on low dose WBCT scans collected in the study.

Non Flare-up Based Treatment

Subjects will follow all procedures and undergo all assessments as specified in the Schedule of Assessments in Table 1 and Table 2 of the protocol. Subjects who cannot receive non-flare-up based treatment will only undergo annual assessments (clinical laboratory tests will not be performed). In the event of a flare-up or traumatic event, these subjects will receive flare-up based treatment and undergo all flare-up based assessments, including clinical laboratory tests and radiographs if they have not reached 100% skeletal maturity.

- **Subjects from Part B Continuing Non-Flare-Up Based Treatment into Part C**

Subjects who began non-flare-up based treatment during Part B will continue this visit schedule into Part C and will receive non-flare-up based treatment for up to an additional 48 months. Therefore, these subjects may undergo non-flare-up based treatment for up to 72 months over the entire study. These subjects will follow all procedures and undergo all assessments as specified in the Schedule of Assessments in Table 1 of the protocol, including low dose WBCT scan (excluding head) at all annual site visits (Months 12, 24, 36, 48, 60 and 72). Non-flare-up Day 1 is the first day that non-flare-up based treatment was initiated during Part B, and total duration of treatment will continue into Part C.

Remote visits (e.g., at home or at a local medical facility; and via telephone contact from clinical site personnel) will occur every 3 months unless the Investigator deems that a site visit is necessary; urine pregnancy tests will be performed each month for female subjects of childbearing potential (FOCBP).

- **Subjects from Part B Starting Non-Flare-Up Based Treatment During Part C**

Subjects who will start non-flare-up based treatment during Part C will receive non-flare-up based treatment for up to 48 months. These subjects will follow all procedures and undergo all assessments as specified in the Schedule of Assessments in Table 2 of the protocol, including low-dose WBCT scan (excluding head) at Screening and at all annual site visits (Part C Screening and Months 12, 24, 36 and 48.). Non-flare-up Day 1 is the first day that non-flare-up based treatment is initiated during Part C for these subjects.

Remote visits (e.g., at home or at a local medical facility; and via telephone contact from clinical site personnel) will occur every 3 months unless the Investigator deems that a site visit is necessary; urine pregnancy tests will be performed each month for female subjects of childbearing potential (FOCBP).

Blood samples for PK assessment of non-flare-up dosing will be collected at the first 3-month safety assessment at pre-dose and 3, 6, 10, and 24 hours post-dose; if samples cannot be or were not obtained at the 3-month safety assessment, or if a subject is on flare-up based treatment, then the PK blood sample for non-flare-up based treatment can be obtained during any subsequent 3-month safety visit.

Flare-up Based Treatment

Subjects receiving flare-up based treatment will follow all procedures and undergo all assessments as specified in the Schedule of Assessments in Table 3 of the protocol. Assessments will occur at Flare-up Cycle Safety Day 1 at 4 weeks, and every 8 weeks thereafter until treatment of the last flare-up or traumatic event in the flare-up cycle is completed. If after 12 weeks all flare-ups have not resolved then additional treatment and assessments will be performed every 4 weeks until all flare-ups have resolved. All assessments will occur remotely unless the Investigator deems it necessary to evaluate subjects at the clinical site. Once all flare-ups or traumatic events in a cycle have resolved and flare-up based treatment has been completed, subjects will resume non-flare-up based treatment with 5 mg palovarotene once daily

(weight-adjusted doses for skeletally immature subjects). Additional knee and hand/wrist radiograph assessments will also be performed every 3 months (±2 weeks) in skeletally immature subjects who (1) received flare up dosing in the period of time since their last assessment; and (2) had not achieved 100% skeletally maturity on their last assessment.

Off Treatment Part D:

No study drug will be administered in Part D. For skeletally immature subjects in Part D, added Year 1 (Y1) and Year 2 (Y2), post last dose of study treatment assessments that include linear height, knee height, physical exam, vital signs, radiographic assessments of the knee and hand/wrist, low-dose WBCT imaging, adverse events, and concomitant medications. Once subjects reach skeletal maturity their participation in Part D will end. The total duration of participation in Part C and Part D combined is a maximum of 4 years (±1 month).

Pharmacokinetics

The pharmacokinetics (PK) of palovarotene dosing will be assessed at the first 3-month safety assessment during non-flare-up based treatment; if samples cannot be or were not obtained during the first 3-month safety assessment, or if subjects are on flare-up based treatment, then PK blood samples for non-flare-up treatment can be obtained during any subsequent 3-month safety visit.

Pharmacokinetics (PK) will also be assessed twice during flare-up based treatment: once during the 20-mg regimen at any time between Study Days 4 to 28, and once during the 10-mg regimen at any time between Study Days 32 to 84, for the first treated flare-up only. If this is not possible for the first treated flare-up, PK blood samples can be obtained during any subsequent flare-up dosing cycle. Pharmacokinetic blood samples will be collected at pre-dose and 3, 6, 10, and 24 hours post-dose. Subjects who underwent PK assessment under PVO-1A-202 Protocol Amendment 3 (Part B) will not have PK assessed again; however, these subjects will require a non-flare-up treatment PK assessed at a 3-month safety visit.

2.3 Study Treatment

As of 04 December 2019, all subjects <14 years of age were required to interrupt study drug due to a partial clinical hold placed on the palovarotene clinical development program by the US FDA. Treatment will subsequently not be restarted in children < 14 years of age.

Subjects and/or their parents/caregivers will report potential flare-up symptoms to site personnel; such symptoms include, but are not limited to, pain, swelling, redness, decreased range of motion, stiffness, and warmth. Only one symptom is required to define a flare-up. If these symptoms are consistent with previous flare-ups, include a subject reported onset date, and are confirmed by the Investigator as associated with a flare-up, subjects will receive open-label palovarotene treatment as follows:

- 20 mg for 4 weeks (28 days) once daily. The first dose will be taken upon flare-up confirmation by the Investigator. To be followed by:
- 10 mg for 8 weeks (56 days) once daily, for a total flare-up treatment duration of 12 weeks (84 days); 10 mg treatment may be extended if the flare-up is ongoing and continue until the flare-up resolves. Dosing will be extended in 4-week intervals and be based on clinical signs and symptoms as assessed by the Investigator.

Flare-up based dosing should also be initiated if the Investigator confirms the presence of a substantial high-risk traumatic event likely to lead to a flare-up.

All dosing, regardless of non-flare-up or flare-up based, will be weight-adjusted in subjects under the age of 18 years with less than 90% skeletal maturity on hand/wrist radiography at Screening. (Note: all weight-based dosing will cease when subjects achieve ≥90% skeletal maturity based on hand/wrist radiography, but radiographic assessment of the growth plate will continue until these subjects achieve 100% skeletal maturity at both knee and hand/wrist locations.)

Table 1: Weight-Adjusted Palovarotene Doses and Dose De-Escalation Doses

Weight Range Category	20-mg Equivalent	15-mg Equivalent	10-mg Equivalent	7.5-mg Equivalent*	5-mg Equivalent*	2.5-mg Equivalent*
20 to <40 kg	12.5 mg	10 mg	6 mg	4 mg	3 mg	1.5 mg
40 to <60 kg	15 mg	12.5 mg	7.5 mg	5 mg	4 mg	2 mg
≥60 kg	20 mg	15 mg	10 mg	7.5 mg	5 mg	2.5 mg

* In the event of dose de-escalation from 20-mg, 10-mg, or 5-mg equivalent, respectively.

If the subject experiences intolerable side effects, the dose may be reduced to the next lower dose; if the subject is already receiving the lowest possible dose as shown in the Table 1, then study drug will be discontinued. Treatment may be extended after the 84-day treatment period (flare-up based dosing) if the flare-up is ongoing and continue until the flare-up resolves. Dosing will be extended in 4-week intervals based on clinical signs and symptoms as assessed by the Investigator, with remote visits performed every 4 weeks while on treatment. In the event the subject requires dose de-escalation due to an intolerable side effect for a previous flare-up, then the dose the subject will receive for a subsequent flare-up will be determined by the Investigator and the Medical Monitor.

Should a subject experience an intercurrent flare-up (defined as a new flare-up or marked worsening of the original flare-up), or if the Investigator confirms the presence of a substantial high-risk traumatic event likely to lead to a flare-up, at any time during flare-up based treatment, the 12-week dosing regimen will restart upon new intercurrent flare-up confirmation by the Investigator (ie, 4 weeks of 20 mg palovarotene followed by 8 weeks of 10 mg palovarotene [or weight-based equivalent]). A Flare-up Cycle will include the first flare-up or traumatic event and any subsequent intercurrent flare-ups or traumatic event during the same dosing period. Flare-Up Cycle Safety Day 1 is defined as the first day that study drug is administered for the first flare-up or traumatic event during that cycle. Safety assessments will be performed on Flare-up Cycle Safety Day 1, at 4 weeks, and every 8 weeks thereafter until treatment of the last flare-up or traumatic event in the Flare-up Cycle is completed. If any flare-up in a cycle has not resolved after 12 weeks, treatment and safety assessments will be extended and 10 mg palovarotene (or the weight-based equivalent) will be administered in 4-week intervals until all the flare-ups resolve and flare-up based treatment has been completed. It is possible that subjects may experience more than one Flare-up Cycle during the study.

2.4 Study Populations

A total of approximately 60 subjects will be enrolled in Part C including up to 40 subjects enrolled from Study PVO-1A-201 and up to 20 new adult subjects enrolled during Part B. All subjects will have participated in Part B and no new subjects will be enrolled in Part C and D.

Flare-up Based Treatment Group:

Flare-up based analyses will be based on flare-up cycles in Part C and will include the treatment group combined PVO 20/10 mg.

Combined PVO 20/10 mg is defined as flare-up cycles that flare-ups were treated with PVO 20/10 mg (or weight-based equivalent) during Part C, regardless of whether non-flare-up PVO 5 mg (or weight-based equivalent) treatment had been administered prior to the flare-up dosing.

Non-Flare-up Based Treatment Group:

Non-flare-up based analyses is based on subjects and will include the treatment group with PVO 5 mg (or weight-based equivalent).

PVO 5 mg group is defined as subjects who received at least one dose of non-flare-up regimen (regardless of whether they received the flare-up regimen) during Part C.

Untreated (Never Received PVO Treatment) Group:

Untreated is defined as subjects who were not on the non-flare-up regimen and did not receive flare-up based treatment in Part C.

2.5 Randomization and Blinding

This is an open-label study and does not involve randomization or blinding.

3 General Statistical Consideration

Categorical data will be summarized with counts and percentages. Percentages will be calculated based on the number of non-missing values and will be reported to one decimal place.

Continuous data will be summarized with descriptive statistics including the number of non-missing values, mean, standard deviation, standard error, median, minimum, and maximum. Minimum and maximum will be reported to the same precision as the raw values, mean and median will be reported to one additional decimal place, and standard deviation and standard error will be reported to two additional decimal places.

Non-flare-up based summaries will be performed based on visits in Table 1 and Table 2 in the protocol.

Flare-up based summaries will be performed based on flare-ups at flare-up visits: Flare-up Day 1, Flare-up Week 4 (28 days), Flare-up Week 12 (Day 84) etc.

To distinguish flare-up based analysis from subject based analysis, m will be used for number of flare-ups and n will be used for number of subjects.

In summary of change from baseline analysis for Enrolled Population, baseline from non-flare-up based treatment will be used for subjects in group PVO 20/10 mg only.

3.1 Definition of Baseline

Unless otherwise stated, baseline will be defined as Section 3.1.1 to 3.1.3.

3.1.1 Baseline for Non-Flare-up Based Treatment

- For subjects from Part B continuing non-flare-up based treatment into Part C (i.e., subjects who ever received at least one dose of non-flare-up based treatment in both Part B and Part C), baseline

will be based on assessment prior to first non-flare-up treatment in Part B. For subjects who continued from Part B into Part C a Part C baseline is defined as the first assessment in Part C.

- For subjects who did not receive any non-flare-up based treatment in Part B but start non-flare-up based treatment during Part C, baseline will be based on assessment prior to first non-flare-up treatment in Part C.

3.1.2 Baseline for Flare-up Based Treatment

For subject receiving flare-up based treatment in Part C, baseline values will be the last available value prior to or on the date of first dose of flare-up treatment with the cycle.

3.1.3 Baseline for Subjects Never Receive PVO Treatment in Part C

For subjects who never received any PVO Treatment in Part C, baseline is defined as the most recent non-missing measurement on Part C Study Day 1.

If baseline is missing in Part C, then baseline will be based on assessment prior to first non-flare-up treatment in Part B.

3.1.4 Baseline for Subjects who Restarted Palovarotene After Dosing interruption in Part C

For subjects who paused dosing either because of the partial clinical hold for subjects <14 years on December 4, 2019 or because of interruption after Study PVO-1A-301 crossed futility on January 24th 2020 and subsequently re-started palovarotene their re-start baseline is defined as the first assessments after re-start.

3.2 Part C Study Day, Flare-up Cycle Days and Non-Flare-up Days

3.2.1 Part C Study Day 1

For all subjects, Part C Study Day 1 is defined as the date of informed consent in Part C. Part C Study Day -1 will be the day immediately prior to Part C Study Day 1.

3.2.2 Flare-up Cycle and Flare-Up Cycle Day 1

- A Flare-up Cycle will include the first flare-up and any subsequent intercurrent flare-ups during the same dosing period.
- Flare-Up Cycle Day 1 is defined as the first day that study drug is administered for the first flare-up during that cycle.

3.2.3 Duration of Flare-up Treatment

For subjects receiving flare-up treatment, duration of flare-up treatment (days) is defined as total flare-up cycle dosing periods (days) across all flare-up cycles.

For a flare-up cycle, the flare-up cycle dosing period (days) is defined as days that subjects received flare-up based treatment and is calculated as (last dose date of the flare-up cycle – first date of the flare-up cycle – days where flare-up meds were not taken/interrupted during the flare-up cycle + 1).

3.2.4 Duration of Non-Flare-up Treatment

Based on study design, subjects are assumed to be on non-flare-up based treatment from the date they received first non-flare-up treatment unless they receive flare-up based treatment. Thus, duration of non-flare-up treatment (days) will be calculated as (last non-flare-up treatment date - first non-flare-up treatment

date – duration of flare-up treatment (days) - days where non-flare-up meds were not taken/interrupted + 1). Duration of flare-up treatment (days) is defined in 3.2.3.

3.3 Definition of New-Onset Potentially Clinically Significant (PCS) Value

New-onset potentially clinically significant (PCS) values are defined as values that were not PCS at baseline and were PCS at post-baseline.

3.4 Adjustments for Covariates

As requested by US FDA during review of the US NDA, Ipsen is conducting multiple analyses comparing annualized new HO between Study 202 Part C and NHS, including

- GEE for new HO from all scans in Study 202 Part C, such that baseline is more than 30 days after the end of the previous flare-up, versus NHS. (Subjects without 2 WBCTs in Study 202 Part C where the first is at least 30 days after the end of the previous flare-up are not included.)
- GEE for new HO from all scans post-restart in Study 202 Part C, such that baseline is more than 30 days after the end of the previous flare-up, versus NHS. (Subjects without 2 WBCTs post-restart in Study 202 Part C where the first is at least 30 days after the end of the previous flare-up are not included.)

These models are run with 3 covariates in addition to treatment – baseline age, sex, total HO – and with 5 covariates in addition to treatment – baseline age, sex, total HO, CAJIS, time since last flare up.

In addition to these regression-based techniques, FDA has requested that propensity-score-based methods, including matched pairs analyses and weighted analyses also be performed to compare annualized new HO in Study 202 versus NHS. Propensity scores are being generated using the set of 5 covariates listed above.

Additional analyses are also being performed, including those that pool Study 202 Part B and Part C as well as several in which Study 202 and Study 301 data are pooled.

3.5 Handling of Dropouts or Missing Data

As there is no effective treatment for subjects with FOP, it is assumed they will be highly motivated to adhere to the protocol and treatment visits. Thus, it is expected that there will be minimal missing data, and most subjects will complete the study. The efficacy analysis will be performed using observed data only. Missing efficacy data will not be imputed.

3.6 Unscheduled Assessments for Clinical Laboratory Test, Vital Sign and ECG

- For unscheduled visits, continuous data from laboratories results, vital signs and ECG will not be used for statistical summary (such as mean with standard deviation). However, the unscheduled results will also not be used for assessment of potentially clinically significance (PCS) or abnormality.
- If multiple records occur within a post-baseline visit, the first non-missing value will be used.
- Lab results beyond the detectable limits will be reported as detectable limits for calculating descriptive statistics.
- Results for lab obtained locally rather than from the central laboratory will also be used for summaries.
- All the laboratory test results, vital sign and ECG will be included in the data listings as reported.

3.7 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.

For the purposes of analysis, incomplete medication start and stop dates will be imputed.

- If the start date and end date are completely missing, then the start date is imputed as the date of informed consent, the end date is the date of last assessment for the subject, and the medication would be considered as both prior and concomitant.
- If a medication start date is incomplete, January will be imputed for missing month and/or the first day of the month will be imputed for missing day.
- If a medication stop date is incomplete, December will be imputed for missing month and the last day of the month will be imputed for missing day.
- If the imputed medication stop date is after the date of study completion, the date of study completion will be used instead.

3.8 Handling of Partial Dates for Adverse Events

- If the start date is completely missing, the start date is imputed as the date of first dose and the AE is considered treatment emergent and assigned to the non-flare-up dosing period.
- If the day and month of the start date are missing and the start year is the same as the year of first dose, the start date is imputed as the date of first dose and the AE is considered treatment emergent and assigned to the non-flare-up dosing period.
- If the start year is prior to the year of first dose, the start date is imputed as January 1 of the year and the AE is not considered treatment emergent.
- If the day of the start date is missing and the start month and year is within a period of time within a flare-up cycle, the start date is imputed as the start date of the flare-up cycle in the same month as the AE, and the AE is considered treatment emergent and assigned to the flare-up dosing period.
- Otherwise, if the day of the start date is missing and the start month and year is on or after the date of first dose, the start date is imputed as the date of first dose or the first day of the month of the AE, whichever is later, and the AE is considered treatment emergent and assigned to the non-flare-up dosing period. Otherwise, if the start month and year is before the month and year of first dose, then the start date is imputed as the first day of the month of the AE, and the AE is not treatment emergent.
- Partial AE end dates will be imputed as the last day of the month or year for missing end dates. Completely missing end dates will be imputed as the date of last known treatment exposure for the subject.

3.9 Interim Analyses

No formal interim analyses were planned for Part C.

This document describes analyses to be included in the final CSR.

3.10 Multicenter Studies

A review of by-center effects will be performed in the context of data listing review. However, the sample size precludes formal assessment of by-center effects.

3.11 Multiple Comparison/Multiplicity

As this is an extension of an initial proof-of-concept study, no adjustment for multiple testing is planned.

3.12 Examination of Subgroups

The sample size in this study is not sufficient to allow rigorous assessment of outcomes by subgroups, and the study is not powered for subgroup assessment. However, exploratory assessments of the outcomes may be performed with a limited number of extrinsic or intrinsic factors, including age of the subject, location of the flare-up, amount of new HO, number of flare-up treatments, patients with extended flare-up dosing, or other variables as appropriate.

4 Analysis Populations

4.1 Enrolled Population

The Enrolled Population will include all subjects enrolled in Part C.

Unless stated, analysis for Enrolled Population will be presented by groups: PVO 20/10 mg Only, nonflare-up PVO 5 mg, PVO total, and Never Received PVO Treatment. The PVO 20/10 mg Only group only includes subjects who didn't take PVO 5 mg.

4.2 Treated Population

The Treated Population will include all subjects in the Enrolled Population who took at least one dose of palovarotene in Part C (either flare-up treatment or non-flare-up treatment).

The analysis for Treated Population will be presented by group combined PVO 20/10 mg in Part C, non-flare-up treatment PVO 5 mg and PVO Total.

The Treated Population will be the primary population for safety assessments.

4.3 Flare-up Treated Population

The Flare-up Treated Population will include all subjects in the Treated Population who took at least one dose of flare-up based treatment in Part C.

The analysis for Flare-up Treated Population will be presented by group: combined PVO 20/10 mg.

4.4 Flare-up Per Protocol Population

The Flare-up Per Protocol Population was to include all subjects in the Flare-up Treated Population with no major protocol deviations that may impact the efficacy assessment and who were at least 80% compliant with flare-up dosing.

Due to the small sample size, particularly when the additional restriction that baseline WBCT must be taken more than 30 days after the end of the previous flare-up (as was the case for Study 301 and NHS), formal per-protocol analyses are not planned to be performed.

4.5 Non-Flare-up Treated Population

The Non-Flare-up Treated Population includes all subjects in the Treated Population who took at least one dose of non-flare-up based treatment in Part C.

The analysis will be presented for subjects who took PVO non-flare-up based dosing with 5 mg (including weight-adjusted doses).

4.6 Restart Full Analysis Set

Restart Full Analysis Set (RFAS) includes all enrolled subjects in the Principal Enrolled Population (EP) who restarted palovarotene therapy after January 2020 and have at least two HO volume measurements in Part C trial on or after restart, with the first such WBCT imaged at least 30 days after the end of the previous flare-up.

4.7 Part D Analysis Set

Part D Population includes all skeletally immature subjects who stopped taking study medication for any reason before completion and had at least one follow-up visit following last dose. All analyses involving this population will be descriptive.

4.8 Pharmacokinetic Population

The Pharmacokinetic (PK) Population will include all subjects in the Treated Population who have at least one post-dose measurable concentration. The analysis will be presented for subjects who took non-flare-up based dosing of PVO 5 mg (including weight-adjusted) and/or flare-up based dosing of PVO 20/10 mg (including weight-adjusted doses).

5 Subject Data

5.1 Subject Disposition

Subjects disposition will be summarized for Enrolled Population:

- Number of subjects in each analysis population
- Number of subjects who discontinued treatment from Part C early and reasons for discontinuation
- Number of subjects who prematurely withdrew from Part C and reason for study discontinuation
- Number of subjects who completed Part C.
- Duration (months) of participation in Part C. This will be calculated as (date of completion of Part C/discontinuation - date of informed consent for Part C)/30.5.
- Number of subjects by following categories:
  - ≤6 months
  - >6 – 12 months
  - >12 – 18 months
  - >18 – 24 months
  - >24 months

Note: The total duration of participation in Part C and Part D, is a maximum of 4 years.
Subjects disposition will be summarized for Enrolled Population in Part D by listings.

5.2 Eligibility Criteria

Eligibility criteria (inclusion/exclusion) will be listed with any specific criteria that were not met shown in the by-subject listing for the Enrolled Population.

5.3 Protocol Deviations

Major protocol deviations will be identified prior to locking the database for Part C and may include, but are not limited to, departures from the inclusion/exclusion criteria, non-compliance with investigational product, use of restricted concomitant medications, and non-compliance with study procedures. Major protocol deviations will be summarized for the Enrolled Population.

For the NDA interim report, protocol deviations were not summarized, due to the study being ongoing at that time.

5.4 Demographic and Other Baseline Characteristics

Demographic and other baseline characteristics such as age, age group (6-15, 15+; 6-18, 18+), sex, race, ethnicity, height (cm), and weight (kg) will be summarized for all analysis populations.

5.5 Medical History

Medical history will not be collected as there are no new subjects in Part C. Subjects medical history was collected and summarized previously in Part A and Part B.

5.6 Prior and Concomitant Medications

Prior and concomitant medications will be coded using the WHO Drug Dictionary (March 2014).

- Prior medications are those the subject used prior to first dose of study drug. Prior medications can be discontinued before first dose of study drug or can be ongoing during the treatment phase.
- Concomitant medications are any treatments received by the subject after first dose of study drug, during Part C. Concomitant medications first received after first dose of study drug in Part C will also be referred to as new onset medications.
- New Onset medications are concomitant medications first received after the first dose of study drug

A given medication can be classified as a prior medication, or as a concomitant medication. Note that these categories are not mutually exclusive. A medication initiated prior to first dose of study drug and continued during the treatment-emergent period will be prior and concomitant.

The prior and new onset medications will be summarized based on the Enrolled Populations as follows:

- Prior medications discontinued prior to the first dose of study drug
- Prior medications ongoing at the first dose of study drug
- New onset medications during study
- Concomitant medications

Systemic glucocorticoids will be summarized including:

- Glucocorticoids taken for palovarotene treated flare-ups for Flare-up Treated population
- Glucocorticoids taken for condition aggravated adverse event for Enrolled Population
- Glucocorticoids taken for all other adverse event for Enrolled Population

The medications will be summarized according to the WHODrug Global March 2014 by Anatomic Therapeutic Class (ATC) and preferred term.. The ATC Class 3 name will be presented if available, and if not available the ATC class 2 name will be presented.

Systemic glucocorticoids used for flare-ups is defined as initiated 9 days prior to start of palovarotene for a flare-up cycle or within 9 days of flare-up cycle initiation which is categorized as a steroid used for treated flare-up or taken initially for a flare-up treated with palovarotene during a flare-up cycle.

For the purposes of analysis, incomplete medication start and stop dates will be imputed (See Section 3.7).

5.7 Treated Flare-ups Starting During Part C

Treated flare-ups for flare-up based treatment during Part C will be summarized for Flare-up Treated Population and Flare-up Per Protocol Population including following variables:

- Total number of flare-up cycles
- Number of flare-up cycles with dosing extended beyond 12 weeks
- Number of flare-up cycles for which an interruption of study drug occurred during flare-up dosing overall and by high or low flare-up based doing. Where an interruption includes ≥ 14 days without dosing.
- Number of flare-up cycles with dosing reductions overall and by high or low flare-up based doing
- Number of flare-up cycles for which study drug was discontinued during flare-up dosing overall and by high or low flare-up based doing
- Total number of treated flare-ups across all flare-up cycles
- Number of flare-ups treated per cycle
- Number of flare-ups treated per cycle categorized
  - 1 flare-up
  - 2 flare-ups
  - 3 flare-ups
  - 4 flare-ups
  - >5 flare-ups
- Number of intercurrent flare-ups treated per cycle
- Number of palovarotene treated trauma flare-ups

A dose interruption is defined as dosing interrupted for at least 14 days without dosing.

Number of subjects with treated flare-ups will also be summarized for Flare-up Treated Population including following variables:

- Number of subjects with treated flare-up cycles
- Number of subjects with treated flare-up cycles by categories
  - 1 flare-up cycle
  - 2 flare-up cycles
  - 3 flare-up cycles
  - 4 flare-up cycles
  - ≥5 flare-up cycles
- Number of subjects with flare-up cycle with dosing extended beyond 12 weeks
- Number of subjects for which an interruption of study drug occurred during flare-up cycle dosing overall and by high or low flare-up based doing
- Number of subjects for which study drug was discontinued during flare-up cycle dosing overall and by high or low flare-up based doing

- Number of subjects with dose reduction during flare-up cycle dosing overall and by high or low flare-up based doing

Treated flare-up data as reported by the subject and by Investigator including symptoms, probable causes, and location(s) will be summarized for the Flare-up Treated Population.

Flare-up location mapping is described in Table 11.

5.8 Annualized Flare-up Rate

Flare-ups are counted using the number of subject/Investigator-reported flare-ups (including intercurrent flare-ups). Rates will be calculated as the total number of flare-ups divided by duration of follow-up (year). The duration of follow-up (years) is defined as (last date on study – Part C Study Day 1 + 1)/365.25. Flare-up rate will be calculated separately for pre-interruption, during interruption, and after re-start of palovarotene dosing.

5.9 Exposure in Part C

Extent of study drug exposure will be summarized for flare-up based treatment and non-flare-up based treatment.

5.9.1 Exposure to Flare-up Based Treatment

The summary of exposure by subject across all flare-up cycles will be included:

- Duration of flare-up treatment by subject across all flare-up cycles, defined as sum of days of exposure across all flare-up cycles (see Section 3.2.3 for calculation)
- Duration of flare-up treatment by subject according to the following categories:
  - ≤3 months
  - >3 – <=6 months
  - >6 – <=9 months
  - >9 – <=12 months
  - >12 – <=15 months
  - >15 – <=18 months
  - >18 – <=21 months
  - >21 – <=24 months
  - >24 months
- Duration of high flare-up treatment by subject (20-mg dose)
- Duration of low flare-up treatment by subject (10-mg dose)

The summary of exposure by flare-up cycle will be included:

- Duration of flare-up treatment by flare-up cycle
- Duration of flare-up treatment by flare-up cycle, according to the following categories:
  - ≤2 weeks
  - >2 – 4 weeks
  - >4 – 6 weeks
  - >6 – 8 weeks
  - >8 – 10 weeks

  - >10 – 12 weeks
  - >12 – 16 weeks
  - >16 – 20 weeks
  - >20 weeks
- Duration of high flare-up treatment by flare-up cycle (20-mg dose)
- Duration of low flare-up treatment by flare-up cycle (10-mg dose)

If a subject has dose reduction, the exposure is counted by planned treatment.

The following summaries will also be presented for Flare-up Treated Population:

- Number of days after flare-up onset to first dose of study drug for flare-up cycles, calculated as sum of (date of first dose of flare-up cycle – date of flare-up onset + 1) across all flare-up cycles

5.9.2 Exposure to Non Flare-up Based Treatment

The following summary will be included:

- Number of subjects with non-flare-up based treatment
- Duration of non-flare-up based treatment (months), defined as days subjects received non-flare-up treatment/30.5 (see Section 3.2.4 for calculation)
- Number of subjects by following exposure categories in months:
  - >0 – 6 months
  - >6 – 12 months
  - >12 – 18 months
  - >18 – 24 months
  - >24 months
- Number of subjects with at least one non-flare-up dose reduction
- Number of dose reductions by type (e.g. 5 to 2.5 mg)
- Number of subjects with study drug interruption, where a dose interruption is defined as dosing interrupted for at least 14 days without dosing
- Number of subjects with study drug discontinuation

5.9.3 Total Exposure to Study Drug for Each Subject

Total exposure to study drug for each subject for Treated Population will be calculated as duration of non-flare-up treatment and flare-up treatment (see Section 3.2.3 and 3.2.4 for calculation).

Total actual dose (mg) will be summarized by Enrolled Population.

- Total dose is calculated as the cumulative dose (mg) across all dosing regimens (non-flare-up and flare-up) for each subject.
- Total non-flare-up dose is calculated as the cumulative dose across all the non-flare-up treatment periods.
- Total flare-up dose is calculated as the cumulative dose taken across all flare-up cycles.
- All total dose calculations account for weight-adjusted doses, dose reductions

5.9.4 Duration of Subject Participation in Part C

Duration (months) of participation in Part C is calculated as (the earlier date of completion of Part C or discontinuation - the date of informed consent + 1)/30.5. For subjects ongoing, it is calculated as (data cutoff date – date of informed consent + 1)/30.5.

Number and percentage of subjects will be summarized for following duration categories for Enrolled Population:

- <= 6 months
- >6 to 12 months
- >12 to 18 months
- >18 to 24 months
- >24 months

5.9.5 Duration of Subject Participation in Part D

- up to 24 months

5.10 Compliance

Compliance is defined as (duration of treatment/expected duration of treatment). Calculation of duration of treatment is described in Section 3.2.3 and 3.2.4.

The expected duration of treatment is calculated as following:

- For non-flare-up treatment, expected duration of treatment is calculated as (last non-flare-up dose date – first non-flare-up dose date) – (sum of (all Flare-up cycle dosing periods)) + 1.
- For flare-up treatment, expected duration of treatment is calculated as total of expected flare-up cycle dosing periods across all flare-up cycles. Each expected flare-up cycle dosing period is defined as (last flare-up cycle dose date – first flare-up cycle dose date + 1).

Compliance will be summarized for the Treated Population with descriptive statistics. Compliance in category (<80%, ≥80%) will also be summarized for counts and percentages of subjects.

6 Efficacy Data

Unless otherwise stated, efficacy data will be summarized using Enrolled Population presented by groups: PVO 20/10 mg Only, PVO 5 mg, PVO Total and Never Received PVO Treatment. The PVO 201/10 mg Only group only includes subjects who did not take PVO 5 mg.

Imaging Reads

Imaging data includes results from low-dose WBCT scans (excluding head) and radiographs.

WBCT HO status and total HO volume or new HO volume will be summarized by visit. Those analysis will be included in the interim CSR. All other analysis will be presented in the final CSR.

6.1 Primary Efficacy Endpoint

The primary efficacy endpoint is the annualized change in new HO volume as assessed by low-dose WBCT scan (excluding head) during Part C.

The change in new HO total volume will be compared to baseline where the baseline value was performed prior to the initiation of non-flare-up based dosing (in Part B or Part C).

The WBCT HO total volume will be calculated as the sum of the volumes for each location of the WBCT assessment. If there is no HO, 0 will be assigned for analysis.

The annualized change in new HO volume will not be summarized in the interim CSR for the palovarotene NDA, due to the ongoing nature of the reads. This analysis will be included in the final CSR. However, all WBCT data (total and by region) HO volume, HO status, New HO volume by visit will be listed.

Annualized HO will be summarized before the pause in palovarotene dosing, i.e. 4 December 2019 for patients younger than 14 years old or 24 January 2020 due to the crossing of the futility boundary in Study PVO-1A-301; and after restarting palovarotene.

Inferential analyses are described in Section 3.4 Adjustment for Covariates.

6.2 Secondary Efficacy Endpoints

Secondary efficacy endpoints for include:

1. Percent of subjects with new HO at Months 12, 24, 36, and overall.
2. Change from baseline in ROM as assessed by CAJIS at Months 6, 12, 18, 24, 30, and 36.
3. Change from baseline in physical function using age-appropriate forms of the FOP-PFQ at Months 6, 12, 18, 24, 30, and 36.
4. Change from baseline in physical and mental function for subjects ≥15 years old and mental function for subjects <15 years old using age-appropriate forms of the PROMIS Global Health Scale at Months 6, 12, 18, 24, 30, and 36.

Secondary endpoints will also be summarized before the pause in palovarotene dosing and after restarting palovarotene.

6.2.1 New HO

The presence of new HO at each scheduled visit will be determined by low-dose WBCT scan, excluding head. Note that "New HO" means new HO compared to baseline. Baseline will be based on assessment from WBCT prior to the start of Non-flare-up dosing in Part B or Part C.

The presence of "HO at baseline" and the incidence of "No new HO", and "New HO" will be summarized with counts and percentages overall and by regions.

6.2.2 Cumulative Analogue Joint Involvement Scale (CAJIS)

Range of motion across the whole body will also be assessed using the CAJIS.

The assessments will be performed on 12 joints (shoulder, elbow, wrist, hip, knee, and ankle on both the right and left sides), and three body regions (jaw, cervical spine [neck], and thoracic/lumbar spine). Each joint/region will be scored as: 0 = uninvolved; 1 = partially involved; and 2 = ankylosed/completely involved.

The CAJIS Total score will be calculated as the sum of the scores of all joints/regions and ranges from 0 (no involvement) to 30 (maximally involved).

The CAJIS Upper Extremities subscore will be calculated as the sum of the scores from six joints (shoulder, elbow, and wrist on both the right and left sides) and one region (cervical spine [neck]) and ranges from 0 (no involvement) to 14 (maximally involved).

The CAJIS Mobility subscore will be calculated as the sum of scores from six joints (hip, knee, and ankle on both the right and left sides) and ranges from 0 (no involvement) to 12 (maximally involved).

The CAJIS Total score, CAJIS Upper Extremities subscore, and CAJIS Mobility subscore and changes from baseline (defined in Section 3.1) will be summarized at each scheduled visit with descriptive statistics.

6.2.3 FOP-Physical Function Questionnaire (FOP-PFQ)

Age-appropriate forms of the FOP-PFQ will be administered to subjects (or parent proxy). The adult form will be administered to subjects 15 years and older. For subjects between 8 and 14 years of age, both the pediatric self-completed and the pediatric proxy-completed forms will be administered. The proxy-completed form will be used for analyses unless only the self-completed form is available.

The FOP-PFQ consists of 28 questions on the adult form and 26 questions on the pediatric form scored on a scale from 1 to 5, with lower scores indicating more difficulty. The questions are described in detail in Appendices 2A, 2B, and 2C of the protocol.

The Total score will be calculated as:

- The sum of the scores from each question and ranges from 28x1 = 28 to 28x5 = 140 for the adult form.
- The sum of the scores from each question and ranges from 26x1 = 26 to 26x5 = 130 for the pediatric form.

The Upper Extremities subscore will be calculated as:

- The sum of the scores from 15 questions (questions 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 25, and 26) and ranges from 15x1 = 15 to 15x5 = 75 for the adult form.
- The sum of the scores from 18 questions (questions 1, 2, 6, 7, 8, 9, 10, 11, 16, 17, 18, 19, 21, 22, 23, 24, 25, and 26) and ranges from 18x1 = 18 to 18x5 = 90 for the pediatric form.

The Mobility subscore will be calculated as:

- The sum of the scores from 13 questions (questions 13, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 27, and 28) and ranges from 13x1 = 13 to 13x5 = 65 for the adult form.
- The sum of the scores from 8 questions (questions 3, 4, 5, 12, 13, 14, 15, and 20) and ranges from 8x1 = 8 to 8x5 = 40 for the pediatric form.

The Total score, Upper Extremities subscore, and Mobility subscore are defined based on information from PROMIS and Hays et al (2013)[1] for the adult form and a medical review of the questions for the pediatric form.

If a subject is missing some (but not more than 20%) of the contributing question scores, the Total score, Upper Extremities subscore, or Mobility subscore will be calculated as the average observed score multiplied by the number of expected question scores. For example, the Total score would be calculated as the average of the non-missing scores x 28 for the adult form. The determination for sufficient non-missing scores will be made independently for the Total score, Upper Extremities subscore, and Mobility subscore.

As the analysis for FOP-PFQ will be performed across all subjects (adult and pediatric) and the number of contributing questions differs, the scores will be transformed to reflect a percentage of worst score. The percentage of worst score ranges from 0% to 100% with 0% indicating the best possible function and 100% indicating the worst possible function. Table 2 illustrates some sample derivations of the percentage of worst score.

Table 2: Sample Derivations for the Percentage of Worst Score for the FOP-PFQ

Sample Subject	Observed FOP-PFQ Score	Lowest Possible Score	Highest Possible Score	Range of Possible Scores	Distance = Highest - Observed	Distance/ Range	Percentage of Worst Score
1	45	15	75	60	30	0.500	50.0%
2	40	15	75	60	35	0.583	58.3%
3	35	15	75	60	40	0.667	66.7%
4	30	15	75	60	45	0.750	75.0%
5	25	15	75	60	50	0.833	83.3%

Percentage of worst scores and changes from baseline (defined in Section 3.1) of will be summarized at each scheduled visit with descriptive statistics.

6.2.4 PROMIS Global Health

An age-appropriate form of the PROMIS Global Health short form will be administered to subjects (or parent proxy). The adult form will be administered to subjects 15 years and older. For subjects between 8 and 14 years of age, both the pediatric self-completed and the pediatric proxy-completed forms will be administered. The proxy-completed form will be used for analysis unless only the self-completed form is available.

The PROMIS Global Health short form consists of 10 questions on the adult form and nine questions on the pediatric form scored on varying scales. The questions are described in detail in Appendices 8A, 8B, and 8C of the protocol.

For the adult form, Global Physical Health and Global Mental Health scores will be calculated. The Global Physical Health score will be calculated as the sum of scores from questions 3, 6, 7, and 8, and will range from 4 (worse health) to 20 (better health). The Global Mental Health score will be calculated as the sum of scores from questions 2, 4, 5, and 10, and will range from 4 (worse health) to 20 (better health). In the calculation of the Global Physical Health and Global Mental Health scores, the following questions will be rescaled as shown in Table 3.

Table 3: Rescaled PROMIS Global Health Scale Scores

Question(s)	Score	Rescaled Score
7	0	5
	1-3	4
	4-6	3
	7-9	2
	10	1
8 and 10	1	5
	2	4
	3	3
	4	2
	5	1

If a subject is missing any of the contributing scores, the Global Physical Health or Global Mental Health score will not be calculated.

For the pediatric form, only a Total score will be calculated as the sum of scores from the first seven questions and will range from 7 (worse health) to 35 (better health). If a subject is missing some (but not more than three) of the contributing question scores, the Total score will be calculated as the average observed score multiplied by the number of expected question scores. For example, the Total score would be calculated as the average of the non-missing scores x 7. The proxy-completed PROMIS form will be used for analyses unless only the self-completed form is available.

Global Physical Health scores and Global Mental Health scores will also be converted to T-scores as described in detail in Appendix 1. A T-score of 50 is normal and increments of 10 +/- 1 standard deviation away from the norm. A T-score <50 indicates worse health, while a T-score >50 indicates better health. Note that higher values indicate better health.

PROMIS Global Heath T-score and changes from baseline (defined in Section 3.1) will be summarized at each scheduled visit with descriptive statistics. Data from adult (both physical health and mental health) and pediatric forms will be summarized separately.

7 Safety Data

Safety data summaries in the final CSR will be based on the safety data presentations in recent documents, including the Safety Update included in the NDA and the Day 90 Safety Update; the primary difference will be that those summaries included data from all palovarotene studies, while the Study 202 presentations will only include data collected in Study 202.

While the majority of content listed below corresponds to the data presentation in the interim CSR, it is also relevant for pooled summaries that will be the basis for the final CSR safety discussion.

Adverse events will generally be presented by treatment period (See Section 7.1 for classification of treatment periods) and in total for the Enrolled Population.

The Principal Safety Set (Principal SS) includes all enrolled subjects with the R206H ACVR1 mutation receiving at least one dose of palovarotene in the current study.

Safety data for flare-up based treatment assessments will generally be presented by group combined PVO 20/10 mg for the Flare-up Treated Population.

For vitals, labs, ECGs, physical examination, linear growth assessment and CSSR, summary for non-flare-up based treatment assessments (at regular scheduled visits) will be presented by groups: PVO 20/10 mg only, PVO 5 mg, PVO Total and Never Received PVO Treatment using Enrolled Population. The PVO 201/10 mg only group only includes subjects who did not take PVO 5 mg.

Assessments obtained in Part D include yearly linear height, knee height, weight, physical exam, vital signs, radiographic assessments of the knee and hand/wrist, low-dose WBCT imaging excluding head, adverse events, and concomitant medications.

Pregnancy and Teratogenicity Assessment

Pregnancy and pregnancy outcomes will be monitored throughout the study

Pregnancy test results for females of child-bearing potential were provided as subject listings.

A listing of pregnancy and pregnancy outcomes will summarize subject results

7.1 Adverse Events

Adverse events will be recorded from the time informed consent (Part C Study Day 1) is signed through study completion. All AEs will be coded to system organ class and preferred term using will be coded using a current version of MedDRA **Study Drug Related Adverse Events:**

At least possibly related AEs are defined as AEs assessed by the Investigator as possibly, probably, or definitely related to study drug.

Adverse events that require Common Terminology Criteria for Adverse Events (CTCAE) grading (Version 4.03, 14 June 2010) as indicated on the eCRF are considered to be retinoid-associated AEs.

Adverse Events Analysis Periods:

All AEs collected during Part C from the time informed consent is signed are considered as treatment emergent adverse events (TEAEs).

TEAEs will be summarized by period (Non-Treatment Period, Non-Flare-up Treatment Period, and Flare-up Treatment Period, Total Treatment Period) and in total.

a) Flare-up treatment period

Adverse events with start date on or after Flare-up Cycle Day 1 and on or before the date of completion for the Flare-up treatment will be attributed to the Flare-up treatment period.

b) Non-Flare-up Treatment Period

Adverse events with start date on or after first Non-Flare-up dose date in Part C for which the subjects are not undergoing flare-up treatment will be attributed to the Non-Flare-up Treatment Period.

Non-Flare-up Treatment Period is period that subjects taking non-flare-up dose including non-flare-up dose interruption periods. All flare-up cycles are excluded.

c) Non-Treatment Period

Adverse events with start date on or after Study Day 1 that are not attributed to the flare-up or non-flare-up treatments will be attributed to the Non-Treatment Period.

Non-Treatment Period are period that subjects are receiving neither non-flare-up and nor flare-up treatment. In Part C, it is the period starting from Part C Study 1 to the day before first study drug dosing date (regardless of whether non-flare-up or flare-up based treatment), and the period from date of study drug discontinuation to end of study or study discontinuation date.

Adverse Events Summary:

The incidence of the following TEAEs will be summarized with counts and percentages of subjects (overall and by MedDRA system organ class and preferred term):

Adverse Event Variables

Adverse Event Observation Periods

- TEAEs are AEs with (first dose date of study drug during 202 C) ≤ (AE start date) ≤ (last dose date of study drug +7 days for all AEs, i.e., non-serious and serious AEs).
    - A chronic TEAE is a TEAE with an onset date during chronic treatment
    - A flare-up TEAE is a TEAE with an onset date during flare-up treatment
- Post-treatment AEs are AEs with a start date > last dose date in Part C + 7 days for All AEs (non-serious and serious AEs).

- Any AEs
- Any post treatment AEs
- Any TEAEs
- Any TEAEs by maximum severity
- Any TEAEs at least possibly related to study drug
- Any TEAEs at least possibly related to study drug by maximum severity
- Any serious adverse events (SAEs)
- Any TEAEs leading to dose modification
- Any TEAEs leading to interruption of study drug
- Any TEAEs leading to discontinuation of study drug
- Any TEAEs leading to discontinuation from study

Data listings will be provided for

- All TEAEs
- All post treatment AEs

- SAEs
- SAEs for premature physeal closure and bone safety
- Deaths
- TEAEs leading to dose modification of study drug
- TEAEs leading to interruption of study drug
- TEAEs leading to discontinuation of study drug
- TEAEs leading to discontinuation of study
- AEs ongoing at date of informed consent signed in Part C

Part C Study Day and/or Flare-up Day for AE start and stop dates will be included on AE listings as appropriate.

Adverse event severity will be assessed and reported according to criteria defined in the protocol (mild, moderate, severe). Adverse event causality will also be assessed in terms of relationship to study drug and reported according to criteria described in the protocol (not related, possibly related, probably related, definitely related).

Deaths

Deaths after the first dose of study drug are either on-treatment or post-treatment:

- Deaths on-treatment: deaths occurring after the first dose of study drug and up to 7 days after the last dose of study drug

- Death post-treatment: deaths occurring more than 7 days after the last dose of study drug until 30 days after last dose of study drug

The SAE, including death, reporting period for the non-flare-up-based treatment and the flare-up-based treatment begins at the time of informed consent and continues through study completion (including Part D) + 30 days.

7.2 Safety Laboratory Assessments

Blood and urine samples will be collected for hematology, biochemistry (includes lipids), and urinalysis testing.
Females of child-bearing potential must have a negative blood or urine pregnancy test (with sensitivity of at least 50 mIU/mL) prior to administration of palovarotene. A blood or urine pregnancy test will be conducted monthly in Part C.
For females of childbearing status: if pregnancy testing did not continue monthly per protocol post the study drug interruption then at a minimum a self-administered urine pregnancy test is to be done within 4 weeks before re-starting palovarotene treatment.

Values and changes from baseline (defined in Section 3.1) for hematology, biochemistry (includes lipids) parameters will be summarized at each scheduled visit with descriptive statistics. Both the change from baseline (subject based) and the change from Flare-up Cycle Day 1 (flare-up based) will be presented.

The following plots will be provided:

- Boxplots for lab parameter over time by scheduled visits and by flare-up visits.
- Mean percent of normal high for serum chemistry and lipid lab analytes including:
  - Aspartate Aminotransferase (AST)
  - Alanine Aminotransferase (ALT)
  - Gamma Glutamyl Transferase Gamma Glutamyl Transferase (GGT)
  - Creatinine
  - Lipase
  - Bilirubin
  - Triglycerides (lipid)
- Mean percent of normal low for hematology lab analytes including:
  - Hematocrit (HCT)
  - Hemoglobin (HGB)
  - Bicarbonate (HCO3)

If multiple records occur within a post-baseline visit, the first non-missing value will be used. Note that results obtained locally rather than from the central laboratory will also be used for summaries.

Potentially clinically significant (PCS) values are predefined criteria/thresholds presented in Table 4. PCS values at baseline and new-onset PCS values post-baseline will be summarized with counts and percentages of subjects or flare-ups as appropriate. Separate analyses will be performed for the non-flare-up treatment visits and the flare-up treatment visits. Percentage of incidence of new-onset PCS laboratory values over time will be plotted.

New-onset PCS values are defined as values that were not PCS at baseline (defined in Section 3.1) and were PCS post-baseline. Note that results obtained locally will be included in these summaries.

Table 4: Potentially Clinically Significant Safety Laboratory Values

	PCS Low	PCS High
Biochemistry		
Aspartate aminotransferase	N/A	3 x ULN
Alanine aminotransferase	N/A	3 x ULN
Amylase	N/A	3 x ULN
Lipase, pancreatic	N/A	3 x ULN
Bilirubin	N/A	>2 mg/dL
AST, ALT and total bilirubin	N/A	AST or ALT > 3 x ULN and total bilirubin > 2 x ULN
Thyroxine	<4.0 mcg/dL	>13.0 mcg/dL
Lipids		
Cholesterol	N/A	>300 mg/dL
Triglycerides	N/A	>400 mg/dL
Hematology		
Leukocytes	<2.8 x 109/L	>16.0 x 109/L
Hemoglobin (females)	<9.5 g/dL	>17.5 g/dL

	PCS Low	PCS High
Hemoglobin (males)	<11.5 g/dL	>19.0 g/dL
Hematocrit (females)	<32%	>54%
Hematocrit (males)	<37%	>60%
Platelets	<75 x 109/L	>700 x 109/L

Urinalysis parameters will be tabulated for Non-Flare-up Treated Population and Flare-up Treated Population.

A listing will be provided for all subjects meeting PCS laboratory criteria that will include all PCS results during the study along with associated palovarotene and glucocorticoid dosing data (dose and start and stop dates). A by-subject listing of all laboratory parameters will be generated. Listing for subjects with PCS lipase values will also be listed separately.

7.3 Vital Signs

Vital signs include blood pressure, heart rate, respiratory rate, and temperature.

Values and changes from baseline will be summarized at each scheduled visit with descriptive statistics. Both the change from baseline (defined Section 3.1) and the change from Flare-up Cycle Day 1 (flare-up based) will be presented. If multiple records occur within a post-baseline visit, the first non-missing value will be used.

Potentially clinically significant (PCS) values are predefined criteria/thresholds presented in Table 5. PCS values at baseline and new-onset PCS values post-baseline will be summarized with counts and percentages of subjects or flare-ups as appropriate. Separate analyses will be performed for the non-flare-up treatment visits and the flare-up treatment visits. New-onset PCS values are defined as values that were not PCS at baseline (defined in Section 3.1) and were PCS post-baseline.

Table 5: Potentially Clinically Significant Vital Sign Values

	PCS Low	PCS High
Systolic blood pressure	<86 mmHg or a decrease of ≥25 mmHg from baseline	>180 mmHg or an increase of ≥25 mmHg from baseline
Diastolic blood pressure	<48 mmHg or a decrease of ≥20 mmHg from baseline	>110 mmHg or an increase of ≥20 mmHg from baseline
Heart rate	<45 bpm or a decrease of ≥20 bpm from baseline	(1) >105 bpm AND an increase of ≥20 bpm from baseline (2) >125 bpm

mmHg = millimeters of mercury, bpm = beats per minute.

7.4 Body Weight

Values and changes from baseline for body weight and body weight z-score will be summarized by visit with descriptive statistics for Non-flare-up Treated Population and Flare-up Treated Population. Mean values will be plotted over time.

7.5 Linear Growth Assessment

A linear growth assessment will be performed for subjects under the age of 18 years at the time of enrollment into study with open epiphyseal growth plates using stadiometer and knee height assessments.

Values and changes from baseline (defined as assessment at screening visit in Part C) will be summarized for linear height and knee to heel length by visit with descriptive statistics.

The values and change from baseline for linear height z-score will also be calculated for each scheduled visit.

Z-scores will be obtained from https://www.cdc.gov/nccdphp/dnpao/growthcharts/resources/sas.htm.

The following spaghetti plots by chronological age at assessment will be provided:

- linear height and linear height z-scores over time
- Knee height over time

Growth velocity will also be summarized in tables and figures. The calculation of growth velocity will be performed as follows: ((Current Height – Previous Height)/(duration (days)=(current date – previous date))) * 365.25 = Annualized Growth Velocity (cm/year). In addition to providing the calculated growth velocities per time period for each subject, the data will also be categorized according to worst growth velocity and last growth velocity by the following three categories: <4 cm/year; 4 to 5 cm/year; and >5 cm/year.

Growth assessments (linear height, knee height, and femur and tibia lengths and growth velocity over 12 months) will also be assessed by age category and premature growth plate closure (PPC) status.

-

7.6 Electrocardiogram

Values and changes from baseline (defined Section 3.1) will be summarized for continuous ECG parameters at each scheduled visit with descriptive statistics. The change from baseline (subject based) will be presented. If multiple records occur within a post baseline visit, the first non-missing value will be used.

Abnormal ECG categories will be assigned based on raw findings as shown in Table 6.

Table 6: Abnormal Electrocardiogram Categories

Abnormal ECG Category	Raw Findings
Rhythm	* Atrial premature complexes * Ventricular premature complexes * Sinus arrhythmia * Sinus tachycardia
Conduction	* Nonspecific intraventricular conduction delay * Right bundle branch block * Incomplete right bundle branch block * Left posterior fascicular block * Left bundle branch block * Incomplete left bundle branch block
Morphology/Chamber enlargement	* Left ventricular hypertrophy * Left ventricular hypertrophy with repolarization abnormality
Axis deviation	* Right axis deviation * Left axis deviation

Myocardial infarction	* Pathologic Q waves * Acute ST elevation MI * Acute non-ST elevation MI
ST Segment/T Waves/U Waves	* Early repolarization * Nonspecific ST and T wave abnormality * Nonspecific ST elevation * Nonspecific U wave abnormalities
Miscellaneous	* Long QT interval * Short PR interval

Overall interpretation and abnormal ECG categories will be summarized at each scheduled visit with counts and percentages of subjects or flare-ups as appropriate.

New-onset ECG abnormalities are defined as abnormalities that did not exist at baseline. New-onset ECG abnormalities will be summarized with counts and percentages of subjects or flare-ups as appropriate. A by-subject listing of new-onset ECG abnormalities will be generated.

Incidence of QTc >450 msec, QTc >500 msec, or change from baseline in QTc >60 msec will be summarized with counts and percentages of subjects or flare-ups as appropriate.

Potentially clinically significant (PCS) values are predefined criteria/thresholds presented in Table 7. PCS values at baseline and new-onset PCS values post-baseline will be summarized with counts and percentages of subjects or flare-ups as appropriate. New-onset PCS values are defined as values that were not PCS at baseline and were PCS post-baseline. A listing will be provided for all subjects meeting PCS ECG criteria.

Table 7: Potentially Clinically Significant Electrocardiogram Values

Parameter	PCS Low	PCS High
PR Interval	None	1) >200 ms only OR 2) increase from baseline ≥20 ms only OR 3) >200 ms and increase from baseline ≥20 ms
QRS Interval	None	1) >100 ms only OR 2) increase from baseline ≥10 ms only OR 3) >100 ms and increase from baseline ≥10 ms
QT Interval	None	1) >500 ms only OR 2) increase from baseline ≥60 ms only OR 3) >500 ms and increase from baseline ≥60 ms
QTcF and QTcB Interval	None	1) >500 ms only OR 2) increase from baseline ≥60 ms only OR 3) >500 ms and increase from baseline ≥60 ms

msec = millisecond, QTcF = rate-corrected QT interval using Fridericia's formula (QT interval divided by the cube root of the RR interval), QTcB = rate-corrected QT interval using Bazett's formula (QT interval divided by the square root of the RR interval)

7.7 Physical Examination

Physical examination results and changes from baseline will be summarized at each scheduled visit by body system with counts and percentages of subjects.

7.8 Columbia-Suicide Severity Rating Scale

Suicidal ideation and behavior will be assessed and summarized using the Columbia-Suicide Severity Rating Scale (C-SSRS).

- Type 1-Type 5 suicidal ideation
  - Type 1suicidal ideation is passive suicidal ideation (wish to be dead)
  - Type 2 suicidal ideation is non-specific active suicidal thoughts
  - Type 3 suicidal ideation is active suicidal ideation without plan or intent
  - Type 4 suicidal ideation is active suicidal ideation with some intent no plan
  - Type 5 suicidal ideation is active suicidal ideation with specific plan and intent
- Behavior
  - Non- suicidal self- injurious behavior
  - Suicidal behaviors include actual attempts, interrupted attempts, aborted attempts, preparatory acts or behavior, and suicidal behaviors

Data listings will be provided for

- All C-SSRS results
- Subjects reporting any type 4 or type 5 suicidal ideations, suicidal behavior or non- suicidal self-injurious behavior

7.9 Bone Safety Data

Bone safety analysis based on WBCT and X-Ray safety imaging data include all phase 2 data (PVO-1A-202 and PVO-1A-201). Additional details can be found in the Bone Safety Management Plan.

Baseline is defined as non-missing measures at first assessed timepoint, regardless of which study this was performed in.

Bone safety data will be described for following time periods:

The Original Baseline Period

The original baseline period will be defined by the original baseline through February 2020.

The Interruption (Off-treatment) Period

The baseline for the off-treatment period is the earlier of the first assessment while off-treatment or the assessment within 1 month prior to the last dose of palovarotene. Bone growth assessments during the interruption period included subjects in the off-treatment population, which was defined as subjects <18 years of age at first study entry who interrupted palovarotene after the clinical hold date and did not re-initiate treatment and who had an off-treatment baseline and post-off-treatment baseline visit.

The end of the interruption period occurred either at the end-of-treatment/end of study visit if the subject discontinued from the study, at the end-of-study visit if they completed the study, at the new baseline if the subject re-initiated palovarotene treatment.

The New Baseline Period

A new baseline was established by the re-initiation of palovarotene dosing following the interruption period. New baseline bone growth assessments includes all subjects who had a new baseline and a post-new baseline visit and received at least one dose of palovarotene after the new baseline visit date. The new baseline visit date was defined as the date of the last visit prior to the re-initiation of palovarotene treatment following the interruption period.

Note that bone safety analysis will also be performed based on premature physeal closure status.

7.9.1 Knee and Hand/Wrist Radiographs

Subjects under the age of 18 years at the time of enrollment into Study PVO-1A-202 or PVO-1A-201 with open epiphyseal growth plates will undergo knee and hand/wrist radiographs (AP view) for assessment of epiphyseal growth plate.

Knee and hand/wrist radiographs will be performed remotely (eg, at a local medical facility) for subjects who are unable or unwilling to attend site visits. Once a subject has achieved 100% skeletal maturity (as confirmed by radiography), knee and hand/wrist radiographs will no longer be required.

Summary statistics of bone age and chronological age (values and changes from baseline) for subjects < 18 years will be calculated at each scheduled visit for Enrolled Population. Changes in bone age will be plotted over time.

Two independent radiology reviewers will qualitatively assess the hand/wrist and knee growth plate for abnormalities listed below. Discrepancies in the reads will be adjudicated.

The following abnormalities of the Epiphyseal Growth Plates assessed from X-ray will be assessed:

- EGP Abnormality Location: Wrist or Knee
- Calcinosis
- Cupping
- Dense Metaphyseal Lines
- Frayed Metaphyseal Edge
- Under Mineralisation or Osteopenia of the Adjacent Growing Bone
- Sclerosis of the Adjacent Growing Bone
- Widening of the Epiphyses
- Other

Summary tables of growth plate abnormalities identified in radiographs will include the following:

- Number and percent of subject with epiphyseal growth plates open (by hand/wrist and/or knee).
- Number and proportion of subjects with any abnormality at each visit, for each specific abnormality (e.g. growth recovery lines, Sclerosis, Under mineralization, etc.).
- Number and proportion of subjects with any change from baseline and from prior timepoint for hand/wrist and knee by each abnormality for each visit.
- Number and proportion of subjects with any change from baseline and from prior timepoint for hand/wrist and knee by each abnormality across all post-screening visits. For summaries across all timepoints, the worst case for a given subject will be used in the analysis.
- Number and proportion of subjects with any change from baseline and from prior timepoint for hand/wrist and knee across all abnormalities and post-screening visits. For summaries across all abnormalities and timepoints, the worst case for a given subject will be used in the analysis.

Premature growth plate closure

Listings of all subjects with premature growth plate closure will be provided.

Analyses of Knee Alignment

Distal femoral angle and premature closure of the growth plate will also be summarized for subjects < 18 years old in Enrolled Population.

Summary statistics of lateral distal femoral angles will be calculated for each visit and new-onset potentially clinically significant knee lateral distal femoral angles (i.e. angles outside of the physiological range defined as 78.9 to 84.9 degrees). Listings will also be generated for subjects meeting PCS criteria overall and by PPC status.

7.9.2 WBCT Bilateral Hand/Wrist and Knee Growth Plate Morphology Assessment

Imaging was performed for pediatric subjects, i.e. subjects <18 years old who underwent WBCT scan at study enrollment and during regularly scheduled follow-up visits thereafter. Although the WBCT scans obtained in the trial are used primarily to quantify change in HO volume, they will also be used to assess bone safety, including bilateral hand/wrist and knee growth plate abnormalities on all subjects age < 18 years.

The following abnormalities of the Epiphyseal Growth Plates assessed from WBCT will be assessed:

- EGP Abnormality Location: Wrist or Knee
- Calcinosis
- Cupping
- Dense Metaphyseal Lines
- Frayed Metaphyseal Edge
- Under Mineralisation or Osteopenia of the Adjacent Growing Bone
- Sclerosis of the Adjacent Growing Bone

- Widening of the Epiphyses
- Other

Counts of abnormalities of epiphyseal growth plates (yes/no/not evaluable) detected will be summarized for each timepoint. (selecting the "last case" in the hierarchy for each abnormality, in order of the terms listed below):

- Open Epiphyseal Growth Plates
  - Yes, Open
  - Yes, Partially Closed
  - No (Closed)
  - Not Evaluable
  - Missing
- Open Epiphyseal Growth Plates Location
  - Wrist or Knee
  - Wrist and Knee
  - Wrist Only
  - Knee Only
  - Not Evaluable
  - Closed

Counts of abnormalities detected will be summarized as follows:

- Any Abnormality of Epiphyseal Growth Plates
  - Yes
  - No
  - Not Evaluable

Frequency of each abnormality for each timepoint and changes from baseline and from prior timepoint will also be summarized.

7.9.3 WBCTs Measurements of Femur and Tibia Lengths

Although the WBCT scans obtained in the trial are used primarily to quantify change in HO volume, they will also be used to assess the length of the femur and tibia bilaterally on all subjects age <18 years.

Two independent radiology reviewers will measure the femur and tibia lengths. The summary statistics of femur and tibia lengths (values and changes from baseline) for subjects < 18 years will be calculated for each visit for Enrolled Population. Femur and tibia length will be summarized for four different subgroups:

- Male, Age < 10, 10 - <14, 14 - < 18
- Female, Age < 8, 8 - <14, 14 - <18

Figures will include plots of the femur and tibia length by chronological age at assessment by gender.

Leg length discrepancies will be calculated by the difference in right and left femur and tibia lengths as measured by low-dose WBCT. Potential clinically meaningful leg length differences (ie, >1.5 cm difference) will be provided in a listing.

7.9.4 Evaluation of Bilateral Hips by WBCT Safety Read

Although the WBCT scans obtained in the Trial are used primarily to quantify change in HO volume, they will also be used to assess potential adverse effects on the femoral head related to avascular necrosis of the hip (AVN) in all subjects.

Two independent radiology reviewers will qualitatively assess the hip joints on WBCT scan to determine whether there is evidence of AVN. Discrepancies in the reads will be adjudicated. Each hip will be assessed for presence or absence of AVN and by specific abnormalities (i.e. cysts, subchondral lucency, subchondral collapse, osteosclerosis, flattening/collapse femoral head, …), if present, the degree of change from the baseline scan and previous scan.

WBCT Hip abnormalities identified in WBCT safety reads will be summarized for Enrolled Population include the following:

- Number and proportion of subjects with No, Yes (possible, probable, or definite) AVN or not evaluable at baseline, and by specific abnormality
- Number and proportion of subjects with any new post-baseline or prior time point AVN, and by specific abnormality
- Number and proportion of subjects with any new post-baseline or prior time point worsening in AVN, and by specific abnormality
- Number and proportion of subjects with improvement in baseline or prior time point AVN, and by specific abnormality

Listings of any subject with possible, probable, definite AVN present will be provided.

7.9.4.1 Assessments for Spinal Health

WBCT scan assessments described above, will also be used to analyze spinal bone mineral density, bone strength, bone mineral content and fractures, utilizing biomechanical computed tomography (BCT) analysis. Additionally descriptive statistics will be provided for each timepoint.

8 Pharmacokinetics Data

Pharmacokinetics of palovarotene dosing will be assessed twice during flare-up based treatment: once during the 20-mg regimen at any time between Study Days 4 to 28, and once during the 10-mg regimen at any time between Study Days 32 to 84, for the first treated flare-up only. If this is not possible for the first treated flare-up, PK blood samples can be obtained during any subsequent flare-up dosing cycle. Pharmacokinetic blood samples will be collected at pre-dose and 3, 6, 10, and 24 hours post-dose. Subjects who underwent PK assessment under PVO-1A-202 Protocol Amendment 3 will not have PK assessed again under PVO-1A-202 Amendment 4.

Blood samples for PK assessment of non-flare-up dosing will be collected at the first 3-month safety assessment at pre-dose and 3, 6, 10, and 24 hours post-dose; if samples cannot be or were not obtained at the 3-month safety assessment, or if a subject is on flare-up based treatment, then the PK blood sample for non-flare-up based treatment can be obtained during any subsequent 3-month safety visit.

The allowed time deviation windows for pre-dose blood sample is 1 hour before dosing. The concentrations outside the sampling window will be listed but not used for PK parameter calculation and summary statistics. Plasma palovarotene concentrations will be summarized with descriptive statistics (number of non-missing values, arithmetic mean, standard deviation, standard error, coefficient of variation, geometric mean, median, minimum, and maximum).

All plasma concentrations below the limit of quantification (LOQ) will be set to zero for the purpose of calculating descriptive statistics. If at any time point, 1/3 or more of subjects have <LOQ values, descriptive statistics will not be calculated at that time point. PK concentrations will be imputed using the following:

- If 0-hour concentration missing, then 24-hour concentration used to replace the missing concentration.
- If 24-hour concentration missing, then 0-hour concentration used to replace the missing concentration.
- Invalid 24-hour concentrations replaced with the 0-hour concentrations. Twenty-four-hour concentrations are determined to be invalid if concentration is greater than the 10-hour value (i.e. assumed to have been obtained after next day dose instead of before next day dose).

For some subject who has both 0-hour and 24-hour concentration missing, the missing values will not be imputed. Data will be listed but excluded from PK parameter calculation.

If any other blood sample collections are missing, PK parameters will not be assessed. Actual sampling times will be used for the PK analysis, if available. Nominal sampling times will be used for descriptive statistics.

Individual concentration-time profiles for palovarotene will be presented (linear and semi-logarithmic scales). Arithmetic means (+/- standard deviation) concentration-time profiles will also be presented (linear and semi-logarithmic scales).

An exploration of the effects of body weight on PK will be performed as the data allows.

The PK analysis of palovarotene will be conducted using model-independent methods as implemented in WinNonlinTM. The PK parameters analyzed will include $AUC_{0\text{-}24(ss)}$, $C_{max(ss)}$, $C_{min(ss)}$, $T_{max(ss)}$, $T_{min(ss)}$, CL_{ss}/F, λ_z, and $t_{1/2}$ at steady-state. The PK parameters are defined as follows in Table 8:

Table 8: PK Parameters

$AUC_{0\text{-}24(ss)}$	Area under the concentration versus time curve over the 24-hour dosing interval; calculated using linear trapezoid rule.
$C_{max(ss)}$	Maximum measured plasma concentration at steady-state.
$C_{min(ss)}$	Minimum measured plasma concentration at steady-state.
$T_{max(ss)}$	Time of maximum measured plasma concentration at steady-state, obtained by inspection.

$T_{min(ss)}$	Time of minimum measured plasma concentration at steady-state, obtained by inspection.
CL_{ss}/F	Clearance defined as: Dose/$AUC_{0-24(ss)}$.
λ_z	Terminal rate constant
$t_{1/2}$	Apparent terminal elimination half-life; calculated as ln(2)/ λ_z. The number of data points included in the regression will be determined by visual inspection, but a minimum of three data points in the terminal phase, excluding C_{max}, will be required to estimate λ_z.

The λ_z will not be presented for subjects who do not exhibit a terminal elimination phase in their concentration-time profiles. In order to estimate the first-order terminal elimination constant, λ_z, linear regression of concentration in logarithm scale vs. time will be performed using at least 3 data points. Uniform weighting will be selected to perform the regression analysis to estimate λ_z.

The constant λ_z will not be assigned if one of the following happens:

- T_{max} is equals to one of the 3 last data points,
- The estimated elimination rate indicates a positive slope, or
- The terminal elimination phase is not linear (as appears in a semi-logarithmic scale) based on visual inspection.

PK parameters will be summarized with descriptive statistics (number of non-missing values, arithmetic mean, standard deviation, standard error, coefficient of variation, geometric mean, median, minimum, and maximum).

9 Changes from Protocol introduced in SAP Version 0.3 (13 July 2020)

For the purposes of clarity and consistency of nomenclature, the study populations described in the protocol were revised to align with the study populations from Part B.

The protocol describes the primary efficacy comparison of outcomes from data collected from the Natural History Study (NHS) and the Phase 2 study. These comparisons will be performed in a future analysis and the analysis described in this SAP will instead focus on comparisons solely from Study PVO-1A-202 Part C.

10 Changes from Protocol introduced in SAP Version 1.1 (16 December 2022)

The following substantive changes from protocol Section 8 ('Statistical and Analytical Plans') are documented:

Location/Section Number	Change	Rationale
8.1 General Methods	Per the protocol, data from Part B and Part C were to be combined. Unless otherwise specified, Part B and Part C data will now be summarized separately, with the focus of the final CSR being data collected after	Data were collected in separate clinical databases. The Interim CSR summarized Part A, Part B and Part C independently, due to the substantial differences in the

	28 February 2020, i.e. cutoff for the interim CSR.	dosing regimens in each part of the study.
8.3 Study Populations	This SAP amendment adds Restart Full Analysis Set (RFAS) includes all enrolled subjects in the FAS who restarted palovarotene therapy after January 2020 and have at least two HO volume measurements in Part C trial on or after restart, with the first such WBCT imaged at least 30 days after the end of the previous flare-up. The Per-Protocol Set (PPS) is not intended to be summarized.	The RFAS population is most similar to those described in other palovarotene studies. Due to the small sample size, particularly when the additional restriction that baseline WBCT must be taken at least 30 days after the end of the previous flare-up (as was the case for Study 301 and NHS), formal per-protocol analyses are not planned to be performed.
8.7.1 Primary Efficacy	Per the protocol, the primary efficacy analysis was to be a wLME comparing annualized HO volume collected in Study 202 Parts B and C with that in Study 001 (NHS). In the final CSR, propensity-score based comparisons of Study 202 Part B and Part C vs NHS, as well as Study 202 Part C alone vs NHS, will be performed. Additional analyses more similar to those specified in the protocol, e.g. GEE, will also be performed.	Efficacy summaries will be based on analyses prepared at the request of regulators.
8.8.1 Adverse Events	Per the protocol, retinoid associated adverse events were to be summarized. Consistent with the interim CSR, this is no longer planned.	The pre-specified retinoid AEs outlined in the protocol did not encompass all AEs which might be considered retinoid associated AEs. SMQ analysis was more comprehensive.

11 Changes from SAP Version 0.3 (13 July 2020) introduced in SAP Version 1.1 (16 Dec. 2022)

The following key substantive changes from SAP Version 0.3 are documented:

Location/Section Number	Change	Rationale
1.1 Context for Version 1.1	Amendment is preparation for final CSR.	Version 0.3 described the data summaries to be used for the interim CSR.
3.1.4 Baseline for Subjects who Restarted Palovarotene after dosing interruption in Part C	Definition added.	Needed for RFAS summaries.
3.4 Adjustment for Covariates	Efficacy analyses comparing annualized new HO in Study 202 vs NHS requested by FDA are described.	Efficacy was summarized descriptively in the interim CSR.
4.4 Flare-up Per Protocol Population	Formal per-protocol analyses are not planned to be performed.	See Section 10
4.6 Restart Full Analysis Set	RFAS added.	See Section 10
4.7 Part D Analysis Set	Part D Analysis Set added.	Needed to summarize post-interruption data for skeletally immature subjects who stopped taking palovarotene before completion and had imaging to allow for their study.
7 Safety Data	Safety data summaries in the final CSR will be based on the safety data presentations in recent documents, including the Safety Update included in the NDA and the Day 90 Safety Update; the primary difference will be that those summaries included data from all palovarotene studies, while the Study 202 presentations will only include data collected in Study 202.	Given the complicated history of palovarotene, basing the CSR summary on previous discussions will ensure consistency.

12 References

1. Hays RD, Spritzer KL, Amtmann D, et al. Upper-extremity and mobility subdomains from the Patient-Reported Outcomes Measurement Information System (PROMIS) adult physical functioning item bank. *Arch Phys Med Rehabil.* 2013;94(11):2291-2296. doi:10.1016/j.apmr.2013.05.014.

13 APPENDIX

Appendix 1. PROMIS T-Score Conversions

The following conversion tables allow a user to convert Global Physical Health, Global Mental Health, and Total scores into T-scores. T-score distributions are standardized such that a 50 represents the average (mean) for the US general population, and the standard deviation around that mean is 10 points. A high score always represents more of the concept being measured. Thus, a subject who has a T-score of 60 is one standard deviation better (healthier) than the general population.

Table 9: PROMIS T-Score Conversions for Adult Global Physical and Mental Health

Adult Global Physical Health			Adult Global Mental Health		
Raw Score	**T-Score**	**Standard Error**	**Raw Score**	**T-Score**	**Standard Error**
4	16.2	4.8	4	21.2	4.6
5	19.9	4.7	5	25.1	4.1
6	23.5	4.5	6	28.4	3.9
7	26.7	4.3	7	31.3	3.7
8	29.6	4.2	8	33.8	3.7
9	32.4	4.2	9	36.3	3.7
10	34.9	4.1	10	38.8	3.6
11	37.4	4.1	11	41.1	3.6
12	39.8	4.1	12	43.5	3.6
13	42.3	4.2	13	45.8	3.6
14	44.9	4.3	14	48.3	3.7
15	47.7	4.4	15	50.8	3.7
16	50.8	4.6	16	53.3	3.7
17	54.1	4.7	17	56	3.8
18	57.7	4.9	18	59	3.9
19	61.9	5.2	19	62.5	4.2
20	67.7	5.9	20	67.6	5.3

Table 10: PROMIS T-Score Conversions for Pediatric Global Health

Pediatric Self-Completed Total			Pediatric Proxy-Completed Total		
Raw Score	**T-Score**	**Standard Error**	**Raw Score**	**T-Score**	**Standard Error**
7	16	3.4	7	14.7	2.9
8	17.1	3.6	8	15.3	3.1
9	18.3	3.7	9	16	3.2
10	19.7	3.8	10	16.9	3.4
11	21.2	3.8	11	18.1	3.6
12	22.8	3.7	12	19.4	3.7
13	24.4	3.6	13	21	3.8
14	26.1	3.6	14	22.7	3.8
15	27.6	3.5	15	24.4	3.7
16	29.2	3.5	16	26.1	3.7
17	30.8	3.5	17	27.7	3.7
18	32.4	3.6	18	29.4	3.8
19	34	3.6	19	31.2	3.8
20	35.6	3.6	20	32.9	3.8
21	37.2	3.6	21	34.6	3.8
22	38.8	3.6	22	36.2	3.8
23	40.4	3.6	23	37.9	3.9
24	42.1	3.7	24	39.7	4
25	43.9	3.7	25	41.7	4
26	45.7	3.6	26	43.6	3.9
27	47.5	3.6	27	45.4	3.8
28	49.2	3.6	28	47.3	3.9
29	51.1	3.7	29	49.3	4.1
30	53.3	3.9	30	51.8	4.4
31	55.7	4.2	31	54.5	4.7
32	58.3	4.5	32	57.3	5
33	61.1	4.9	33	60.2	5.4
34	64.2	5.4	34	63.2	6
35	67.5	6.1	35	66.1	6.5

Appendix 2. Flare-up Location Mapping

Table 11: Flare-up Location Mapping

List	Flare Location	Flare Other Specify	Adjusted Location: Combine Left and Right
1	ABDOMEN		Lower spine/abdomen
2	ANTERIOR NECK		Head/Neck
3	CERIVCAL SPINE		Cerivcal spine
4	CHEST		Upper Spine/Chest
5	HEAD		Head/Neck
6	JAW		Jaw
7	LEFT ANKLE OR FOOT		Distal lower extremities
8	LEFT DISTAL LOWER EXTREMETIES		Distal lower extremities
9	LEFT DISTAL LOWER EXTREMITIES		Distal lower extremities
10	LEFT DISTAL UPPER EXTREMETIES		Distal upper extremities
11	LEFT ELBOW		Elbow
12	LEFT HIP		Hip
13	LEFT KNEE		Knee
14	LEFT SHOULDER		Shoulder
15	LEFT WRIST OR HAND		Distal upper extremities
16	LOWER BACK		Lower spine/abdomen
17	LOWER SPINE/ABDOMEN		Lower spine/abdomen
18	LUMBAR SPINE		Lower spine/abdomen
19	NECK		Head/Neck
20	OTHER	ABOVE LEFT KNEE	Knee
21	OTHER	BOTH RIGHT KNEE AND RIGHT HIP	Hip
22	OTHER	LEFT ANTERIOR THIGH	Hip
23	OTHER	LEFT KNEE/ LEFT CALF	Knee
24	OTHER	LEFT NECK	Head/Neck
25	OTHER	LEFT NECK AND SUBMANDIBULAR AREA	Head/Neck
26	OTHER	LEFT POSTERIOR CALF, FOCUS OF TENDERNESS 8 CM BELO	Knee
27	OTHER	LEFT TRICEPS	Shoulder
28	OTHER	NECK (RIGHT AND BELOW THE CLIN)	Head/Neck
29	OTHER	NECK RIGHT AND RIGHT ARM	Shoulder
30	OTHER	RIGHT ARM	Shoulder
31	OTHER	RIGHT BICEP	Shoulder
32	OTHER	RIGHT PECTORAL AND AXILLARY AREA	Upper Spine/Chest
33	OTHER	RIGHT THIGH AND CERVICAL SPINE	Hip
34	OTHER	RIGHT TRICEP	Shoulder

List	Flare Location	Flare Other Specify	Adjusted Location: Combine Left and Right
35	OTHER	UPPER BACK AND RIGHT SCAPULA	Upper back
36	OTHER	LEFT SHOULDER AND LEFT SIDE OF NECK	Shoulder
37	OTHER	RIGHT THIGH	Hip
38	OTHER	RIGHT TRICEPS	Shoulder
39	OTHER	LEFT DISTAL MEDIAL THIGH	Knee
40	OTHER	LEFT SCAPULA	Upper back
41	OTHER	LEFT HAMSTRING	Hip
42	OTHER	LEFT HAMSTRING	Hip
43	OTHER	LEFT FOREARM	Elbow
44	OTHER	LEFT ANTERIOR THIGH	Hip
45	OTHER	LEFT FOREARM	Elbow
46	OTHER	LEFT HAMSTRING	Hip
47	OTHER	RIGHT THIGH, NOT INCLUDING HIP OR KNEE	Hip
48	OTHER	R THIGH	Hip
49	OTHER	RIGHT UPPER THIGH	Hip
50	OTHER	SOURCE ONLY STATED "BACK"	missing
51	OTHER	LEFT CHEEK	Head/Neck
52	OTHER	LEFT ARM	Elbow
53	OTHER	3 SITES: NECK, SKULL, SHOULDER	Head/Neck
54	OTHER	RIGHT GROIN	Hip
55	OTHER	RIGHT CHEEK	Head/Neck
56	OTHER	HEAD, NECK, ARM	Head/Neck
57	OTHER	RIGHT GREAT TOE	Distal lower extremities
58	OTHER	RIGHT LOWER BACK	Lower spine/abdomen
59	OTHER	BICEPS, RIGHT SIDE	Shoulder
60	OTHER	THUMB (RIGHT)	Distal upper extremities
61	OTHER	RIGHT CALF	Knee
62	OTHER	SUBMAXILAR AREA	Jaw
63	OTHER	THE FLARE UP IS LOCATED IN THE RIGHT THUMB.	Distal upper extremities
64	OTHER	RIGHT SHOULDER	Shoulder
65	OTHER	THIGH AND LEFT LEG	Hip
66	OTHER	HEAD AND RIGHT ARM	Head/Neck
67	OTHER	RIGHT SHOULDER, RIGHT ELBOW, LEFT ELBOW AND LEFT ARM	Shoulder
68	OTHER	UPPER BACK (LEFT SIDE) AND CERVICAL SPINE (LEFT SIDE)	Upper back
69	OTHER	BELOW THE CHIN	Head/Neck
70	OTHER	NECK - RIGHT SIDE	Head/Neck

List	Flare Location	Flare Other Specify	Adjusted Location: Combine Left and Right
71	OTHER	LEFT THIGH	Hip
72	OTHER	RIGHT THIGH	Hip
73	OTHER	LEFT SCAPULAR AREA	Upper back
74	OTHER	RIGHT INNER THIGH	Hip
75	OTHER	FOREHEAD / BRIDGE OF THE NOSE	Head/Neck
76	OTHER	SACRUM/COCCYX	Lower spine/abdomen
77	OTHER	3RD AND 4TH METATARSALS	Distal lower extremities
78	OTHER	LEFT GREAT TOE	Distal lower extremities
79	OTHER	RIGHT GROIN.	Hip
80	OTHER	LEFT POSTERIOR CALF, FOCUS OF TENDERNESS 8 CM BELOW POPLITEAL FOSSA	Knee
81	RIGHT ANKLE OR FOOT		Distal lower extremities
82	RIGHT DISTAL LOWER EXTREMETIES		Distal lower extremities
83	RIGHT DISTAL LOWER EXTREMITIES		Distal lower extremities
84	RIGHT DISTAL UPPER EXTREMITIES		Distal upper extremities
85	RIGHT ELBOW		Elbow
86	RIGHT HIP		Hip
87	RIGHT KNEE		Knee
88	RIGHT SHOULDER		Shoulder
89	RIGHT WRIST OR HAND		Distal upper extremities
90	SUBMANDIBULAR AREA		Head/Neck
91	THORACIC SPINE		Upper Spine/Chest
92	UPPER BACK		Upper back
93	UPPER SPINE/CHEST		Upper Spine/Chest
94		02-002 E (202B) LOCATION OF FLARE-UP 2 IS LEFT SUBSCAPULAR/ MID-BACK REGION. THIS IS NOT AN OPTION TO CHOOSE AND IS AN EXCEPTION MADE BY PPD .	Upper Spine/Chest
95	BILATERAL LOWER LEGS		Distal lower extremities
96	CERVICAL SPINE		Cerivcal spine
97	L AXILLA, EXTENDING INTO SCAPULAR REGION		Shoulder
98	LEFT GROIN		Hip
99	LEFT HIP		Hip
100	LEFT LATERAL NECK		Head/Neck
101	LEFT NECK		Head/Neck
102	LEFT PECTORAL MUSCLE		Upper Spine/Chest
103	LEFT POSTERIOR THIGH		Hip
104	POSTERIOR NECK		Head/Neck

List	Flare Location	Flare Other Specify	Adjusted Location: Combine Left and Right
105	RIGHT FOREARM		Elbow
106	RIGHT ILIAC CREST		Hip
107	RIGHT THIGH		Hip
108	RIGHT UPPER ARM		Shoulder
109	RIGHT UPPER THIGH		Hip
110	Other Location, Specify	AROUND RIGHT SCAPULA	Upper back
111	Other Location, Specify	BETWEEN SHOULDER BLADES	Upper back
112	Other Location, Specify	CERVICAL SPINE AND THORACIC SPINE	Upper Spine/Chest
113	Other Location, Specify	FRONTAL	missing
114	Other Location, Specify	HEAD - FOREHEAD	Head/Neck
115	Other Location, Specify	INNER RIGHT ARM EXTENDING FROM ELBOW TO SHOULDER. TRANSIENT SWELLINGS OVER BACK AND BACK OF HEAD CONSTANTLY HAPPENING; FALL 2012 SWELLING OVER OCCIPUT.	Elbow
116	Other Location, Specify	LEFT LEG (11/2016) AND BILATERAL BUTTOCKS (11/2016)	Hip
117	Other Location, Specify	LEFT SHOULDER, EXTENDING TO LEFT ELBOW	Shoulder
118	Other Location, Specify	LEFT SHOULDER, LEFT ELBOW	Shoulder
119	Other Location, Specify	LEFT THIGHT	Hip
120	Other Location, Specify	LOWER BACK AND CERVICAL SPINE	Lower spine/abdomen
121	Other Location, Specify	NECK	Head/Neck
122	Other Location, Specify	NECK (RIGHT SIDE)	Head/Neck
123	Other Location, Specify	RIGHT FIRST DIGIT	missing
124	Other Location, Specify	RIGHT SHOULDER AND LEFT SHOULDER. FLARES ARE SYMMETRICAL AND TEND TO OCCUR ON BOTH SIDES. FIRST FLARES WERE TOPS OF SHOULDERS, LOOKED LIKE AN OFFENSIVE LINEMAN.	Shoulder
125	Other Location, Specify	RIGHT SHOULDER, LEFT SHOULDER, AND CHEST	Shoulder
126	Other Location, Specify	THIGH FOLLOWING VACCINATION INTO RIGHT THIGH	Hip
127	Other Location, Specify	THORACIC SPINE AND LUMBAR SPINE	Lower spine/abdomen
128	Other Location, Specify	TOE (AFTER SURGERY FUSED TOGETHER)	Distal lower extremities
129	Other Location, Specify	UNDER CHIN	Head/Neck
130	Other Location, Specify	UPPER BACK, NECK, ABDOMEN	Upper back
131	Other Location, Specify	UPPER BACK/ NECK, ANTERIOR	Upper back
132	Other Location, Specify	UPPER LEG (RIGHT SIDE)	Hip

List	Flare Location	Flare Other Specify	Adjusted Location: Combine Left and Right
133	OTHER	ANTERIOR NECK AND ANTERIOR RIGHT THORAX	Head/Neck
134	OTHER	ANTEROLATERAL RIGHT THIGH	hip
135	OTHER	BACK	missing
136	OTHER	BACK OF HEAD	Head/Neck
137	OTHER	BACK OF NECK	Head/Neck
138	OTHER	BILATERAL ANKLES	Distal lower extremities
139	OTHER	BILATERAL ANKLES, CAN'T REMEMBER WHICH WAS FIRST	Distal lower extremities
140	OTHER	BILATERAL GROIN	hip
141	OTHER	BOTH LOW LEGS	Distal lower extremities
142	OTHER	CENTRAL MID-BACK ALONG THE SPINE	lower spine/abdomen
143	OTHER	CLAVICULA	Upper Spine/Chest
144	OTHER	FACE	Head/Neck
145	OTHER	FEMUR, RIGHT SIDE	hip
146	OTHER	FORE HEAD	Head/Neck
147	OTHER	FOREHEAD	Head/Neck
148	OTHER	FOREHEAD AND PERIORBITAL	Head/Neck
149	OTHER	JAW AND NECK	jaw
150	OTHER	L CALF	knee
151	OTHER	L POSTERIOR RIBCAGE	Upper back
152	OTHER	L SCAPULA	Upper back
153	OTHER	LEFT AXILLA	shoulder
154	OTHER	LEFT LEG	missing
155	OTHER	LEFT POSTERIOR THIGH	hip
156	OTHER	LEFT SHOULDER AND THORACIC SPINE	shoulder
157	OTHER	LEFT SUPERIOR MEMBER	missing
158	OTHER	LEFT UPPER BACK	Upper back
159	OTHER	MIDDLE BACK	lower spine/abdomen
160	OTHER	MOUTH	jaw
161	OTHER	NECK (SIDE)	Head/Neck
162	OTHER	NECK AND SHOULDERS	Head/Neck
163	OTHER	NECK/THROAT	Head/Neck
164	OTHER	ONE HIP, CANNOT REMEMBER SIDE	hip
165	OTHER	R LOWER ABDOMEN/R UPPER THIGH	lower spine/abdomen
166	OTHER	R SUBSCAPULAR REGION	Upper back
167	OTHER	R UPPER THIGH/LEG	hip
168	OTHER	RIGHT BACK SIDE, SHOULDER	shoulder
169	OTHER	RIGHT BACK, SHOULDER	shoulder
170	OTHER	RIGHT HEEL	Distal lower extremities

List	Flare Location	Flare Other Specify	Adjusted Location: Combine Left and Right
171	OTHER	RIGHT LATEROCERVICAL	Head/Neck
172	OTHER	RIGHT NECK	Head/Neck
173	OTHER	RIGHT SIDE NECK	Head/Neck
174	OTHER	RIGHT THIGH/KNEE	hip
175	OTHER	SACRAL SPINE AREA	lower spine/abdomen
176	OTHER	SHOULDERS, ELBOWS, CHEST AND BACK	shoulder
177	OTHER	SKULL	Head/Neck
178	OTHER	TOE (AFTER SURGERY, FUSED TOGETHER)	Distal lower extremities
179	OTHER	UPPER AND LOWER BACK	Upper back
180	OTHER	UPPER BACK/HEAD/NECK (OCCIPITAL REGION)	Upper back

PVO-1A-202C_ SAP_V1.1 CLEAN

Final Audit Report 2022-12-17

Created:	2022-12-16
By:	PPD
Status:	Signed
Transaction ID:	PPD

"PVO-1A-202C_ SAP_V1.1 CLEAN" History

Document created by PPD
2022-12-16 - 10:30:49 PM GMT

Document emailed to PPD for signature
2022-12-16 - 10:32:20 PM GMT

Email viewed by PPD
2022-12-16 - 10:35:10 PM GMT

Signer PPD entered name at signing as PPD
2022-12-16 - 10:35:32 PM GMT

PPD has agreed to the terms of use and to do business electronically with IPSEN PHARMA SAS.
2022-12-16 - 10:35:34 PM GMT

Document e-signed by PPD
Signature Date: 2022-12-16 - 10:35:34 PM GMT - Time Source: server

Document emailed to PPD for signature
2022-12-16 - 10:35:36 PM GMT

Email viewed by PPD
2022-12-16 - 10:49:00 PM GMT

Signer PPD entered name at signing as PPD
2022-12-17 - 9:14:17 PM GMT

PPD has agreed to the terms of use and to do business electronically with IPSEN PHARMA SAS.
2022-12-17 - 9:14:19 PM GMT

Document e-signed by PPD
Signature Date: 2022-12-17 - 9:14:19 PM GMT - Time Source: server

Agreement completed.

2022-12-17 - 9:14:19 PM GMT