# STATISTICAL ANALYSIS PLAN

## A Multicenter, Double-Blind, Randomized, Placebo-Controlled, Phase 3 Study Evaluating the Safety and Efficacy of Intravenous Iloprost in Patients With Systemic Sclerosis Experiencing Symptomatic Digital Ischemic Episodes (AURORA Study)

**Investigational Product:** Iloprost Injection, for intravenous use
**Protocol Number:** ES-301
**Protocol Version**: 4.0, 18 May 2020

**Sponsor:**
Eicos Sciences, Inc.
3 East 3rd Ave, Suite 200
San Mateo, CA 94401
USA
Telephone: 301-968-2390

**SAP Version:** 4.0, 07 June 2021

# SIGNATURE PAGE

**A Multicenter, Double-Blind, Randomized, Placebo-Controlled, Phase 3 Study Evaluating the Safety and Efficacy of Intravenous Iloprost in Patients With Systemic Sclerosis Experiencing Symptomatic Digital Ischemic Episodes (AURORA Study)**

**Protocol Number:** ES-301

We, the undersigned, have reviewed and approve this Statistical Analysis Plan.

| Signature | Date |
|---|---|
| PPD<br>Medpace, Inc. | |
| PPD<br>Medpace, Inc. | |
| PPD<br>Eicos Sciences, Inc. | |
| PPD<br>Eicos Sciences, Inc. | |
| PPD<br>Eicos Sciences, Inc. | |

# VERSION HISTORY

| Version | Date | Description |
|---|---|---|
| 1.0 | 18 November 2019 | Original signed version |
| 2.0 | 06 February 2020 | Updated List of Abbreviations to include biomarkers.<br>Updated 3.1.3 Exploratory efficacy endpoints to include Biomarkers.<br>Updated 3.2.2 Safety Laboratory Evaluations to add chemistry and hematology assessment at Day 5.<br>Updated 5.2 to include handling of multiple ePRO entries.<br>Section 6.4.1 added to describe analysis of analgesic use.<br>Updated 6.5 to add count and percentage patients who has completed the infusions to the summary.<br>Updated 7.2 to clarify the analysis and to add a sensitivity analysis to test robustness of the severity endpoint.<br>Updated 7.4 to include additional subgroup analyses: age, gender, race, ethnicity, baseline blood pressure and baseline kidney function.<br>Section 12 removed as Protocol has been amended and this analysis plan is consistent with current Protocol amendment.<br>Added Appendix B Calculating the total daily dose of opioids. |
| 3.0 | 21 April 2021 | Updated to match the protocol amendment 3.<br>Added cumulative distribution curves for the primary and first two key secondary efficacy endpoints.<br>Added Section 9 to evaluate the impact of Covid-19 pandemic.<br>Added Section 5.2.4 for handling negative duration of symptomatic RP attack.<br>Added summary tables and clarified some details in Section 8.2 clinical laboratory evaluation. Local lab is considered. Added Appendix C for related CTCAE criteria.<br>Added details for clarification in the other sections. |
| 4.0 | 07 June 2021 | Clarified the baseline calculation in section 3.1.<br>Clarified the definition of Per-Protocol Population in section 4.1.3. |

| | | |
|---|---|---|
| | | Updated the rule of handling analgesic log data in section 6.4.1.<br>Added two exploratory analyses for PGIC and PGIS in section 7.3.1.<br>Added some detail about the laboratory summary in section 8.2.<br>Updated appendix B to include Tramadol and Methadone (>80 mg/day) |

# TABLE OF CONTENTS

# LIST OF ABBREVIATIONS

| | |
|---|---|
| AE | Adverse event |
| AESI | Adverse event of special interest |
| ANCOVA | Analysis of covariance |
| ATC | Anatomical Therapeutic Chemical |
| bFGF | basic fibroblast growth factor |
| BMI | Body mass index |
| cAMP | Cyclic adenosine monophosphate |
| CCB | Calcium channel blockers |
| CSR | Clinical Study Report |
| CXCL4 | C-X-C motif ligand 4 |
| EDC | Electronic data capture |
| eGFR | Estimated Glomerular Filtration Rate |
| ePRO | Electronic patient-reported outcomes |
| IL | Interleukin |
| ITT | Intent-to-treat |
| LLN | Lower limit of normal |
| MedDRA | Medical Dictionary for Regulatory Activities |
| MME | Morphine milligram equivalents |
| MMP-2 | matrix metalloproteinase-2 |
| NRS | Numeric Rating Scale |
| PDE5 | Phosphodiesterase-5 |
| PDGF | Platelet-derived growth factor |
| PGIC | Patient Global Impression of Change |
| PGIS | Patient Global Impression of Severity |
| PINP | Amino-terminal propeptide of type 1 collagen |
| PT | Preferred Term |
| RP | Raynaud's Phenomenon |
| SAE | Serious adverse event |

| | |
|---|---|
| SAP | Statistical analysis plan |
| sE-selectin | Soluble E-selectin |
| sICAM-1 | Soluble intercellular adhesion molecule-1 |
| SOC | System Organ Class |
| SSc | Systemic sclerosis |
| TEAE | Treatment emergent adverse event |
| TPA | Tissue plasminogen activator |
| ULN | Upper limit of normal |
| VCAM-1 | Vascular cell adhesion molecule 1 |
| VEGF | Vascular endothelial growth factor |
| VWF | Von Willebrand factor |
| WHO | World Health Organization |

# 1 INTRODUCTION

The purpose of this document is to provide a description of the statistical methods and procedures to be implemented for the analysis of data from Eicos Sciences, Inc. protocol number ES-301. If circumstances arise during the study such that more appropriate analytic procedures become available, the statistical analysis plan (SAP) may be revised. Any revisions to the SAP (both alternative and additional methods) will be made prior to database lock and reasons for such revisions will be described in the final Clinical Study Report (CSR).

# 2 OVERVIEW

## 2.1 Objectives

The primary objective is to evaluate the efficacy of iloprost compared to placebo on the change from baseline in symptomatic digital ischemic episodes, as measured by the weekly frequency of symptomatic Raynaud's phenomenon (RP) attacks, in patients with systemic sclerosis (SSc).

The secondary objectives are the following:

- To evaluate the efficacy of iloprost compared to placebo on the overall severity of RP attack symptoms
- To evaluate the efficacy of iloprost compared to placebo on the weekly total duration of symptomatic RP attacks
- To evaluate the safety and tolerability of iloprost

The exploratory objectives are the following:

- To evaluate the efficacy of iloprost compared to placebo on the worst pain associated with symptomatic RP attacks
- To evaluate the efficacy of iloprost compared to placebo on the worst numbness associated with symptomatic RP attacks
- To evaluate the efficacy of iloprost compared to placebo on the worst tingling associated with symptomatic RP attacks
- To evaluate the efficacy of iloprost compared to placebo on the worst discomfort associated with symptomatic RP attacks
- To evaluate the efficacy of iloprost compared to placebo on the Raynaud's Condition Score
- To evaluate the efficacy of iloprost compared to placebo on symptomatic RP attack duration
- To evaluate the efficacy of iloprost compared to placebo on the patient assessment of overall change in symptomatic Raynaud's attacks (Patient Global Impression of Change [PGIC])

- To evaluate the efficacy of iloprost compared to placebo on the patient assessment of overall severity in symptomatic Raynaud's attacks (Patient Global Impression of Severity [PGIS])
- To evaluate the patient assessment of overall benefit compared to side effects
- To evaluate the effect of iloprost compared to placebo on biomarkers of SSc

## 2.2 Trial Design

This is a Phase 3, multicenter, double-blind, randomized, placebo-controlled study to evaluate the safety and efficacy of iloprost on the frequency of and relief from symptomatic digital ischemic episodes in patients with SSc. Patients are allowed to continue receiving stable standard of care therapies for the management of symptomatic RP (e.g., calcium channel blockers, angiotensin-converting enzyme inhibitors, statins, fluoxetine, and low dose acetylsalicylic acid) or may participate without the use of background standard of care therapies.

The study will target randomizing 180 patients, with the intention to achieve this target by 31 March 2021.

- At any time prior to 31 March 2021, if there are 180 patients randomized, the recruitment for patients will stop. The patients already in screening period would continue to randomization only if such randomization would occur prior to 31 March 2021; such patients, post randomization, would continue through the post treatment efficacy period.
- If there are fewer than 180 patients randomized by 31 March 2021, then randomization will stop on that date.
- Subject enrollment may be halted by the Sponsor at any time for any reason.

The study consists of an up to 30-day screening period during which patients will complete a daily electronic patient-reported outcomes (ePRO) diary to record information regarding all symptomatic RP attacks (e.g., frequency, severity of symptoms and duration) and analgesic medication use (prescription and over-the-counter). The up to 30-day screening period consists of a 5-day eligibility period and an up to 25-day baseline ePRO diary completion period:

- During the 5-day eligibility period, eligible patients will have a minimum of 10 symptomatic RP attacks, documented in the ePRO diary, occurring over at least 3 separate days of the 5-day eligibility period.
- The baseline ePRO diary completion period is a minimum of 10 days and a maximum of 25 days prior to the day of randomization. While 10 days is the preferred baseline ePRO completion period, it may be necessary to extend the baseline ePRO diary completion period beyond 10 days in order to accommodate the preferred Monday to Friday dosing schedule. All eligible patients must complete a minimum of 80% of the daily ePRO diary entry during the baseline period. The ePRO symptomatic RP attack data from this time period will be used for the calculation of the baseline frequency of symptomatic RP attacks.

Eligible patients will be randomized in a 1:1 ratio to iloprost injection for IV use or placebo. Randomization will be stratified based on the use of phosphodiesterase (PDE5) inhibitors at screening. Study drug administration will begin on Day 1 (Visit 2), and patients will receive study drug for 5 consecutive days (e.g., Monday through Friday) as a continuous IV infusion over 6 hours each day via a peripheral line utilizing the NovaCath™ Integrated IV Catheter System or a peripherally inserted central catheter (PICC) using an infusion pump.

Patients must have a systolic blood pressure ≥85 mmHg (sitting position) 15 minutes (±15 minutes) prior to study drug administration each day of administration. On Day 1 (Visit 2), study drug will be initiated at a starting dose of 0.5 ng/kg/min, and dose increases will occur every 30 minutes (±5 minutes) in increments of 0.5 ng/kg/min up to 2.0 ng/kg/min or the individual tolerated dose. If dose-limiting adverse events (e.g., headache, flushing, jaw pain, myalgia, nausea, or vomiting) occur that cannot be tolerated by the patient, then the dose will be reduced in a stepwise manner by 0.5 ng/kg/min every 30 minutes (±5 minutes), until a tolerated dose is determined. If symptomatic hypotension or a dose-limiting adverse event occurs during administration of study drug at the starting dose (i.e., 0.5 ng/kg/min), the study drug infusion will be discontinued and reinitiation of the study drug infusion can be attempted after the event has resolved or been treated. Blood pressure and heart rate will be obtained 15 minutes (±15 minutes) prior to study drug administration and monitored 15 minutes (±5 minutes) after all up-titrations. If the patient experiences symptomatic hypotension or any other adverse event that cannot be tolerated, as determined by the Investigator, during administration of study drug, the dose will be reduced or the study drug infusion will be stopped until the symptoms resolve, at which point the study drug can be reinitiated at a previously tolerated dose. The maximum tolerated dose will be maintained for the remaining 6-hour daily period. At the end of the 6-hour study drug infusion period, the dose will be stopped. Vital signs will be obtained 15 minutes (±5 minutes) after completion of the study drug infusion.

On Days 2 to 5, the study drug infusion will be started using the highest study drug infusion rate tolerated on the previous day without up- or down-titration, unless the patient does not tolerate the study drug infusion or adverse events occur that cannot be tolerated by the patient and necessitate a reduction in the dose (in 0.5 ng/kg/min increments) and subsequent up-titration is allowed to the Day 1 highest tolerated dose. A lower starting dose may be initiated on Days 3 to 5 if the patient does not tolerate the previous days' highest tolerated dose as a starting dose. Vital signs will be obtained 15 minutes (±15 minutes) prior to study drug administration and at 15 minutes (±5 minutes) after all up-titrations during the study drug infusion. Additionally, vital signs will be obtained at 15 minutes (±5 minutes) after completion of the 6-hour study drug infusion.

Patients will be contacted via telephone on Day 8 (+2 days) (Visit 7) to remind patients to continue to complete the daily ePRO diary; patients will complete the ePRO diary through Day 21. On Day 22 (+2 days) (Visit 8), patients will be contacted via telephone to assess adverse

events and reminded to return to the clinic for post-treatment evaluations. A follow-up visit will occur 30 days after the last administration of study drug on Day 35 (+7 days) (Visit 9).

Patients who discontinue study drug early will remain in the study (unless the patient withdraws consent) and complete the daily ePRO diary through Day 21 as well as post treatment study assessments.

The total duration of the study for a patient will be up to approximately 9 weeks.

A detailed schedule of procedures is provided below:

| | Screening Period[a] | | Treatment Visits | | Post-Treatment Visits | | Follow-up Visit | ET Visit[d] |
|---|---|---|---|---|---|---|---|---|
| | Eligibility Period[b] | Baseline ePRO Completion Period[c] | | | | | | |
| **Visit** | **1** | **NA** | **2** | **3, 4, 5, 6** | **7[e]** | **8[f]** | **9** | |
| **Time Point** | **Days -30 to -26** | **Days -25 to -1** | **Day 1** | **Days 2 to 5** | **Day 8** | **Day 22** | **Day 35** | |
| **Window** | | | | | **+2 days** | **+2 days** | **+7 days** | |
| Informed consent | X[g] | | | | | | | |
| Demographics | X | | | | | | | |
| Medical/surgical history[h] | X | | | | | | | |
| Inclusion/exclusion criteria | X | | X[i] | | | | | |
| Physical examination | X | | | | | | | |
| Height and weight[j] | X | | X | | | | | |
| Vital signs (HR and BP) | X | | X[k] | X[k] | | | X | X |
| 12-lead ECG | X | | | | | | | |
| Clinical laboratory assessments[l] | X | | | X[m] | | | X | X |
| Optional biomarkers[n] | X | | | X[o] | | | X | |
| Urine pregnancy test[p] | X | | X[q] | | | | | |
| Distribute ePRO diary[r] | X | | | | | | | |
| ePRO diary completion[s] | X[t] | X | X | X | X | X[u] | | X |
| Randomization | | | X[v] | | | | | |
| Insertion of NovaCath™ Integrated IV Catheter System or PICC[w] | | | X | | | | | |
| Premedicate with acetaminophen[x] | | | X | X | | | | |
| Administer study drug infusion[y] | | | X[z] | X | | | | |
| Digital ulcer assessment[h] | X | | | | | | X | X |
| Return ePRO tablet | | | | | | | X | |
| Adverse events[aa] | X | | X | X | X | X | X | X |
| Prior/concomitant medication | X | | X | X | X | X | X | X |

a. The study consists of an up to 30-day screening period during which subjects will complete a daily ePRO diary to record information regarding all symptomatic RP attacks (eg, severity of symptoms and duration) and analgesic medication use (prescription and over-the-counter). The up to 30-day screening period consists of a 5-day eligibility period and an up to 25-day baseline ePRO diary completion period. Subjects are allowed a 1-time rescreening at least 2 weeks from the last screening assessment.
b. During the 5-day eligibility period, eligible subjects will have a minimum of 10 symptomatic RP attacks, documented in the ePRO diary, occurring over at least 3 separate days of the 5-day eligibility period.
c. The baseline ePRO diary completion period is a minimum of 10 days and a maximum of 25 days prior to the day of randomization. While 10 days is the preferred baseline ePRO completion period, it may be necessary to extend the baseline ePRO diary completion period beyond 10 days in order to accommodate the preferred Monday to Friday dosing schedule. All eligible subjects must complete a minimum of 80% of the daily ePRO diary entry during the baseline period. The ePRO symptomatic RP attack data from this time period will be used for the calculation of the baseline frequency of symptomatic RP attacks.
d. In circumstances where a subject discontinues the study prior to Day 35 (+7 days) (Visit 9), an ET visit will be performed. Contact the IRT system to record the subject's withdrawal from the study.
e. Subjects will be contacted via telephone on Day 8 (+2 days) (Visit 7) to remind subjects to continue to complete the daily ePRO diary; subjects will complete the ePRO diary through Day 21.
f. Subjects will be contacted via telephone on Day 22 (+2 days) (Visit 8) to confirm they completed end of study questions (overall improvement, severity, and benefit) and to confirm they no longer need to complete the ePRO diary.
g. Written informed consent for the study will be obtained from all subjects or legally authorized representative before any protocol-specific procedures are carried out. In situations where electronic consenting is permitted, consent can occur on or before Day-30 and the screening period will begin once the subject has been trained and begins to complete the ePRO diary.
h. Digital ulcers will be assessed at screening and captured as part of the medical history at this visit. New DUs will be captured as an adverse event/SAE. In all cases, subjects that present with DUs will be followed until Day 35 (+7 days) (Visit 9) or the ET visit. Investigator assessment of the DU will consist of location, status, and healing.
i. Eligible subjects must complete a minimum of 80% of the daily ePRO diary entry during the baseline period.
j. Height will be measured at screening only.
k. Vital signs will be obtained 15 minutes (±15 minutes) prior to study drug administration and at 15 minutes (±5 minutes) after all up-titrations during the study drug infusion. Additionally, vital signs will be obtained at 15 minutes (±5 minutes) after completion of the 6-hour study drug infusion. Subjects must have a systolic BP ≥85 mmHg (sitting position) 15 minutes (±15 minutes) prior to study drug administration each day of administration. The arm used for the study drug IV infusion should not be used for blood pressure monitoring.
l. Includes chemistry and hematology.
m. On Day 5 (Visit 6) only, perform clinical laboratory assessments (chemistry and hematology) 15 minutes (±15 minutes) postinfusion.
n. Blood samples for optional biomarker analysis will be collected during the 5-day eligibility period, Day 5 (Visit 6) (15 minutes [±15 minutes] prior to the end of the study drug infusion), and Day 35 (+7 days) (Visit 9). The arm used for the study drug IV infusion on Day 5 (Visit 6) should not be used to draw blood samples. Subjects who consent to the optional biomarker assessments will not have these assessments completed if decentralized study visits are performed for these subjects.
o. Day 5 (Visit 6) only.
p. For women of childbearing potential only.
q. Prior to study drug administration.
r. Subjects will be provided with an ePRO diary at Visit 1 and trained on its use. Subjects will start the daily ePRO diary the day after signing informed consent. Specific questionnaires include the severity of RP attack symptoms (using NRS), Raynaud's Condition Score, overall patient improvement, and overall patient severity.
s. Subjects will complete the daily ePRO diary during the eligibility period and baseline ePRO diary completion period. Subjects randomized will complete the ePRO diary during the treatment and post-treatment period Days 1 (Visit 2) to 21. Overall patient improvement and patient benefit will be completed on Day 21. Overall patient severity will be completed on Day 1 (Visit 2) (prior to initiating study drug) and Day 21.
t. Completion of the daily ePRO diary begins the same day that the ePRO training is completed.
u. Day 21 is the last day of diary entry. Subjects will be contacted via telephone on Day 22 (+2 days) (Visit 8) to confirm they completed end of study questions (overall improvement, severity, and benefit) and to confirm they no longer need to complete the ePRO diary.
v. Contact the IRT system to randomize the subject and obtain study drug assignment. Randomization may take place on Day 1 (Visit 2) or up to 4 days prior.

w. Prior to the first study drug infusion, the NovaCath™ Integrated IV Catheter System will be inserted into a small peripheral vein in the hand or arm using aseptic technique by a trained healthcare professional. If the subject's peripheral veins do not support the placement of a peripheral line, a trained healthcare professional will insert a PICC. When possible, the same NovaCath Integrated IV Catheter System or PICC will be used for the 5 days of treatment. The NovaCath Integrated IV Catheter System (peripheral catheter system) should be replaced when clinically indicated (eg, loss of line patency, signs of phlebitis, etc) or required by institutional policy. If peripheral venous access cannot be obtained in a subject during an infusion visit, the study drug infusion will be missed on that day. If no peripheral venous access is possible, a PICC should be considered.

x. Unless a subject has taken an acetaminophen-containing prescription or nonprescription (over-the-counter) drug product within 4 hours prior to study drug administration, subjects will be premedicated with a single dose of 2 acetaminophen extended-release tablets (650 mg each), 30 minutes (±15 minutes) prior to study drug administration.

y. To be administered for 5 consecutive days (eg, Monday through Friday) as a continuous IV infusion over 6 hours each day via a peripheral line or a PICC using an infusion pump. Study drug will be initiated at a dose of 0.5 ng/kg/min and will be increased every 30 minutes (±5 minutes) in increments of 0.5 ng/kg/min to 2.0 ng/kg/min or the individual tolerated dose. The weight of the subject at screening may be used with the dose rate card (Appendix C or Appendix D) to determine the starting flow rate for each subject. If the weight of the subject at screening is used, it should be confirmed on Day 1 (Visit 2). Subjects with hepatic dysfunction (Child-Pugh Class B and Class C liver disease) will require a starting a reduced starting dose and modified dose titration. Refer to Protocol Section 5.5.3 for further details.

z. Prior to study drug administration, subjects must complete the Patient Global Impression in Severity questionnaire of the ePRO diary.

aa. All adverse events will be collected from the time of informed consent until Day 35 (+7 days) (Visit 9).

BP = blood pressure; DU = digital ulcer; ECG = electrocardiogram; ePRO = electronic patient-reported outcomes; ET = early termination; HR = heart rate; IRT = Interactive Response Technology; IV = intravenous; NA = not applicable; NRS = Numeric Rating Scale; PICC = peripherally inserted central catheter; RP = Raynaud's phenomenon; SAE = serious adverse event.

# 3 STUDY ENDPOINTS

## 3.1 Efficacy Endpoints

For the endpoint calculated using data from the ePRO diary, the baseline value will be the (weekly) average of the inputs during the 10- to 25-day baseline ePRO diary completion period (i.e. Day -25 to Day -1), and the post-treatment efficacy value will be the (weekly) average of the inputs during Days 8 through 21.

For the baseline calculation, any entries between randomization and the first day of study drug administration (Day 1) will be included. Entries after Day 1 (Visit 2) will not be included in baseline calculation.

A symptomatic Raynaud's attack for this study is defined as at least 1 color change of the patient's finger(s) (blue, white, or red) associated with at least 1 symptom (pain, numbness, tingling, and/or discomfort of the finger[s]). The attack is considered over when the color changes back to pre-attack color (normal) and the symptoms return to the patient's pre-attack level.

### *3.1.1 Primary efficacy endpoint*

The primary efficacy endpoint is the absolute change of symptomatic digital ischemic episodes, as measured by the weekly frequency of symptomatic RP attacks, at baseline compared to the end of the efficacy follow-up.

The baseline weekly frequency of symptomatic RP attacks is defined as the average number of weekly symptomatic RP attacks occurring during the 10- to 25-day baseline ePRO diary completion period. The double-blind weekly symptomatic RP attacks is defined as the average number of weekly symptomatic RP attacks during Days 8 to 21, inclusive. The primary efficacy parameter is the change in the weekly frequency of symptomatic RP attacks from baseline (i.e., the mean number of weekly symptomatic RP attacks during Days 8 to 21 compared to during the 10- to 25-day baseline ePRO diary completion period).

For an individual patient, if ePRO diary was not completed (missing RP attack data) on a particular day, the day will be excluded from the calculation of the mean. If the patient answers that there are no symptomatic RP attacks on a day, the number of symptomatic RP attacks is zero on that date. If the patient answers that there are symptomatic RP attacks on a day and the number of symptomatic RP attacks is missing (which is unlikely since the ePRO will not allow patients to proceed without entering the number of attacks), then the number of symptomatic RP attacks on that day will be imputed by the mean of the patient's number of symptomatic RP on the other days with symptomatic RP attacks.

The weekly frequency is calculated by taking the sum of the attacks over the time period, divided by the number of days with data reported (including 0 attacks) in the time period, and multiplied by 7. For the double-blind weekly symptomatic RP attacks, as long as ePRO diary data are

reported for at least seven days during the post treatment efficacy period (Days 8 to 21), the endpoint will be calculated in this way.

During post treatment efficacy period (Days 8 to 21), if there are less than 7 days values available, the average will not be calculated, corresponding endpoints will be considered as missing and imputed. The imputation rule for analysis is specified in section 5.2.2.

### *3.1.2 Secondary efficacy endpoints*

The secondary efficacy endpoints include changes from baseline to the end of the efficacy follow-up in the following:

- Severity of RP attacks as determined by the overall severity of RP attack symptoms (using a Numeric Rating Scale [NRS]). The symptom (pain, numbness, discomfort, or tingling) with the worst average baseline value for each patient will be used for evaluating the patient's overall severity. If more than 1 symptom has the same value, the symptom used for analysis will be based on the following order of rank: pain>numbness>tingling>discomfort.
- Weekly total duration of symptomatic RP attacks.
- Proportion of responders, defined as patients that have at least a 50% reduction in weekly total duration of symptomatic RP attacks and at least a 50% reduction in overall severity from baseline.

**Severity of RP Attack Symptoms**

Raynaud's phenomenon attacks are associated with significant discomfort (pain, numbness, tingling, and/or discomfort). Patients will be asked to select the number that best indicates the severity of their worst pain, numbness, tingling, and discomfort daily using an 11-point NRS (from 0-No pain/ numbness/ tingling/ discomfort to 10-Extreme pain/ numbness/ tingling/ discomfort). The baseline (10- to 25-day baseline ePRO diary completion period) and post treatment efficacy period (Days 8 to 21) severity will be the average severity over the respective time periods. A patient's average score of each symptom (pain, numbness, tingling and discomfort) will be calculated separately for each period. The symptom(s) with the worst (highest NRS) average baseline value will be used to evaluate the patient's overall severity at baseline and post treatment efficacy period. Therefore, different patients' overall severity may be evaluated by the average scores of different symptoms. For example, if a patient has baseline average score pain = 8.2, numbness = 8.2, tingling = 6.3 and discomfort = 5.5, and post treatment efficacy period average score pain = 3, numbness = 4, tingling = 3 and discomfort = 4, then we will use pain to evaluate this patient's overall severity because the pain and numbness have the worst score and the order of rank: pain>numbness. The patient's overall severity is 8.2 at baseline and 3 in post treatment efficacy period.

**Weekly Total Duration of Symptomatic RP Attacks**

The average weekly total duration for baseline (10- to 25-day baseline ePRO diary completion period) and post treatment efficacy period (Day 8 to 21) will be calculated over the respective time periods.

For each period:

Weekly total duration = (Average Duration of Symptomatic RP Attacks per Attack) x (Weekly Frequency of Symptomatic RP Attacks), where the Average Duration of Symptomatic RP Attacks per Attack is described in section 3.1.3.

During post treatment efficacy period (Days 8 to 21), if there are less than 7 days values available, the average will not be calculated, corresponding endpoints will be considered as missing and imputed. The imputation rule for analysis is specified in section 5.2.2.

### *3.1.3 Exploratory efficacy endpoints*

The exploratory efficacy endpoints include changes from baseline to the end of the efficacy

follow-up in the following:

- Proportion of days without symptomatic RP attacks.
- NRS for worst pain associated with symptomatic RP attacks.
- NRS for worst numbness associated with symptomatic RP attacks
- NRS for worst tingling associated with symptomatic RP attacks
- NRS for worst discomfort associated with symptomatic RP attacks
- Proportion of pain responders, defined as patients that have a 50% reduction in worst pain associated.
- Proportion of numbness responders, defined as patients that have a 50% reduction in worst numbness associated.
- Proportion of tingling responders, defined as patients that have a 50% reduction in worst tingling associated.
- Proportion of discomfort responders, defined as patients that have a 50% reduction in worst discomfort associated.
- Proportion of total weekly duration responders, defined as patients that have a 50% reduction in weekly total duration of symptomatic attacks.
- Raynaud's Condition Score
- Duration of symptomatic RP attacks (average duration of an attack)
- Patient assessment of overall change in symptomatic Raynaud's attacks (PGIC)
- Patient assessment of overall severity in symptomatic Raynaud's attacks (PGIS)
- Patient assessment of overall benefit compared to side effects

Additionally, depending on the results of the primary and secondary efficacy analyses, the effect of iloprost on plasma biomarkers may be evaluated.

**Days without symptomatic RP attacks**

Patients will be asked if they had Raynaud's attacks in the last 24 hours within their ePRO diaries daily. The baseline and endpoint will be the proportion of days without symptomatic RP over the respective time periods. If ePRO diary was not completed on a particular day, the day will be excluded from the calculation.

**Worst pain associated with symptomatic RP**

Patients will be asked to rate the severity of the worst pain using 11-point NRS within their ePRO diaries. The baseline and endpoint worst pain will be the average over the respective time periods.

Worst numbness, worst tingling and worst discomfort may also be explored in the same way.

**Raynaud's Condition Score**

The Raynaud's Condition Score asks patients to rate their difficulty with Raynaud's condition on a given day from "No difficulty (0)" to "Extreme difficulty (10)." Patients will be asked to consider the number of attacks they have had on that day and how long each attack lasted. Patients will also be asked to consider how much pain, numbness, or other symptoms the Raynaud's caused in their fingers (including painful sores) and how much the Raynaud's alone affected the use of their hands that day.

The baseline and endpoint Raynaud's Condition Score will be the average score over the respective time periods. If the score is not reported for a particular day, it will be excluded from the mean derivation; no daily scores will be imputed.

**Average Duration of Symptomatic RP Attacks (per Attack)**

Patients are asked to document the duration of each RP attack within their ePRO diaries. For each attack, the patient will record the duration in minutes. The average duration for baseline (10- to 25-day baseline ePRO diary completion period) and post treatment efficacy period (Day 8 to 21) will be calculated over the respective time periods.

For each period:

Average duration of symptomatic RP attacks per attack = (Total duration of symptomatic RP attacks) / (number of symptomatic RP attacks)

**Patient assessment of overall change in symptomatic RP (PGIC)**

At Day 21 (+2 days), patients will be asked to rate their overall change in symptomatic Raynaud's attacks compared to the start of the study. Overall change in symptomatic Raynaud's attacks will be assessed on a 7-point scale from very much better to very much worse.

**Patient assessment of overall patient severity (PGIS)**

At Day 1 (prior to initiating study drug), patients will be asked to rate their overall severity in symptomatic Raynaud's attacks in the last week. At Day 21 (+2 days), patients will be asked to rate their overall severity in symptomatic Raynaud's attacks. Overall severity of symptomatic Raynaud's attacks will be assessed on a 5-point scale from none to very severe.

**Patient assessment of patient benefit**

At Day 21 (+2 days), patients will be asked if the study drug provided a meaningful benefit: Did the study drug provide a meaningful benefit compared to any side effects? Yes/No

**Biomarker**

Blood samples for biomarker analysis will be collected during the 5-day eligibility period, Day 5 (15 minutes prior to the end of the study drug infusion), and Day 35. These biomarkers include C-X-C motif ligand 4 (CXCL4), soluble E-selectin (sE-selectin), vascular endothelial growth factor (VEGF), tissue plasminogen activator (TPA), basic fibroblast growth factor (bFGF), soluble intercellular adhesion molecule-1 (sICAM-1), vascular cell adhesion molecule 1 (VCAM-1), amino-terminal propeptide of type 1 collagen (PINP), matrix metalloproteinase-2 (MMP-2), endostatin, von Willebrand factor (VWF), endothelin-1, (e.g., CXCL4, sE-selectin, VEGF, TPA, bFGF, sICAM-1, VCAM-1, PINP, MMP-2, endostatin, VWF, endothelin-1, angiopoietin-1, angiopoietin-2, thromboplastin, IL-10, IL-6, PDGF, vascular endothelial cadherin, and cAMP).

## 3.2 Safety Variables

### *3.2.1 Adverse Events*

Since RP attacks are recorded as efficacy assessments, these will not be treated as adverse events. However, if an event results in hospitalization or amputation, the event will then be reported as a serious adverse event (SAE). If patients have non-RP-related pain during an RP attack, that event is captured as an adverse event.

Adverse events, which include clinical laboratory test variables, are monitored and documented from the time of informed consent until Day 35.

A Treatment-emergent adverse events (TEAE) is defined as an adverse event with a start date and time on or after the administration of study drug.

Adverse events related to study drug will be those with potential relationship to the study drug.

Hypotension events will be considered as adverse events of special interest (AESI). Additionally, AESIs may be identified by the sponsor. AESIs will be identified before unblinding.

### *3.2.2 Safety Laboratory Evaluations*

Chemistry and hematology will be assessed at Screening, Day 5, Day 35, and Early Termination Visit. A urine pregnancy test will be performed for women of childbearing potential at Screening and on Day 1 prior to study drug administration.

### *3.2.3 Vital Signs*

Vital signs include heart rate and blood pressure and are assessed at each study visit.

On dosing days (Days 1 to 5), vital signs will be obtained 15 minutes (±15 minutes) prior to study drug administration and at 15 minutes (±5 minutes) after all up-titrations during the study drug infusion. Additionally, vital signs will be obtained 15 minutes (±5 minutes) after completion of the 6-hour study drug infusion.

Height will be measured at screening only. Weight will be assessed at screening and the first in-clinic visit (Day 1).

### *3.2.4 Other Safety Variables*

A 12-lead ECG will be performed at screening.

A complete physical examination will be performed at Screening. The physical examination will consist of an examination of the following systems: general appearance, skin, eyes, ears, nose, throat, neck (including thyroid), lymph nodes, chest, heart (including any auscultated cardiac murmurs), abdomen, extremities, and neurologic.

## 4 ANALYSIS POPULATIONS

### *4.1.1 Intent-to-Treat Population*

The Intent-to-Treat (ITT) Population is defined as all randomized patients who initiate (receive) study drug infusion. The ITT Population is the primary efficacy analysis population.

### *4.1.2 Safety Population*

The Safety Population is defined as all randomized patients who initiate (receive) study drug infusion. The Safety Population will be used for all safety analyses.

### *4.1.3 Per-Protocol Population*

The Per-Protocol (PP) Population is defined as all patients in the ITT Population who complete at least 3 days of infusion and who complete at least 7 days of ePRO (primary efficacy endpoint) during both the baseline (Days -25 to -1) and post treatment efficacy (Days 8 to 21) periods, without any major protocol deviations that may impact the primary endpoint assessment. The primary analysis will be repeated on PP Population.

Major protocol deviations that may impact the primary efficacy assessment may include but are not limited to:

- Failure to meet key inclusion/exclusion criteria
- Mis-randomization. Wrong treatment kit used for patient and the corresponding planned treatment is not used. For example, the patient is randomized to the placebo group but the iloprost is used instead or vice versa.
- Use of prohibited medication(s) that may impact the efficacy assessment

Patient data will be reviewed by the clinical team to identify exclusions from the PP Population. The list of patients to be excluded from the PP Population will be finalized prior to database unblinding. For deviations such as mis-randomization that can only be determined after unblinding, the patients will be excluded from the PP Population after unblinding.

# 5 GENERAL CONSIDERATIONS FOR DATA ANALYSIS

## 5.1 Baseline Determination

For efficacy measurements, baselines are defined in section 3.1.

For safety measurements, unless otherwise stated, baseline will be the Day 1 value. If the Day 1 value is missing, the last measurement prior to the first administration of study drug will be the baseline value. Baseline for clinical laboratory assessments will be the value at the Screening Visit.

## 5.2 Handling of Dropouts, Missing Data and Multiple Data Entries

### *5.2.1 Missing or Incomplete Dates*

Dates will be printed in ISO 8601 date format (YYYY-MM-DD). If only year and month are available, date will be displayed as YYYY-MM. If only year, then just YYYY. Dates that are missing because they are not applicable for the patients will be output as “NA”, unless otherwise specified.

For diagnosis/onset dates for Systemic Sclerosis diagnosis and Raynaud's Phenomenon diagnosis, the following imputation rules for partial dates will be used for summary (if the imputed date is later than the date of screening, the date of screening will be used instead):

- If month and year available but day missing, day will be imputed to the 15$^{th}$ (e.g., 2009-06 will be imputed as 2009-06-15)
- If year available but day and month missing, month and day will be imputed to July 1 (e.g. 2010 will be imputed as 2010-07-01).

Adverse events with missing start dates will be considered as treatment-emergent unless the partial date excludes that possibility, e.g. the adverse event month is prior to the treatment infusion month.

Otherwise, the first day of the month will be used to impute missing start days and January will be used to impute missing start months. If the causality for an adverse event is missing, the adverse event will be assumed to be related.

### *5.2.2 Imputation for Efficacy Endpoints*

Day 8 to Day 21 (Post Treatment Efficacy Period)

For primary and secondary efficacy analysis, if there are values available for less than 7 days between Day 8 and Day 21, the missing daily values between Day 8 and Day 21 (inclusive) required to total 7 days values will be imputed by (1) the average value of the patients in the same treatment group on the same study day if it is intermittent missing daily values, or (2) the average value of the patients in the placebo group on the same study day if the patient has discontinued (early termination) from the study. For example, if a patient has values on Day 8 to Day 10, Day 12, Day 13, and discontinues on Day 14, then the missing value on Day 11 will be imputed with method (1) and the missing value on Day 14 will be imputed with method (2). Missing values on Day 15 to Day 21 will not be imputed since there are sufficient (at least 7) daily values to calculate the average. If a patient has values on Day 8 to Day 10 and Day 15 to Day 18, then no imputation is needed since there is already 7 daily values to calculate the average.

Overall severity will be derived after the imputation for each symptom.

### *5.2.3 Data Handling for Multiple ePRO Entries*

Multiple Entries of the same ePRO diary form on a given day

When there are multiple data entries of an ePRO form on a given day, the analysis would be conducted using the first form containing data. The “last local saved time” will be used to compare the time when the entries are saved.

### *5.2.4 Negative duration of symptomatic RP attack*

Some patients may accidentally select a negative number for duration of symptomatic RP attack, but it is evident that the duration should not be less than zero. The absolute value will be used if negative duration of symptomatic RP is entered.

## 6 ANALYSIS OF DISPOSITION AND PATIENT CHARACTERISTICS

### 6.1 Disposition and Analysis Populations

Patient disposition will be provided for all randomized patients. The number and percentage of patients in each of the following disposition categories will be presented by treatment and in total:

- Patients who are randomized,

- Patients who initiated infusion of study drug,
- Patients who completed treatment per protocol
- Patients who completed Day 22 telephone visit (efficacy period) and
- Patients who completed the study.

For patients who did not complete treatment per protocol, a summary will be provided for the reason of discontinuing treatment.

For patients who did not complete the study, a summary will be provided for the reason of discontinuation.

The reason for screen failure will be summarized and listed.

The number and percent of patients in each of the analysis populations will also be provided.

The reason for exclusion from the PP population will be listed.

## 6.2 Protocol Deviations

Protocol deviations as defined in the Medpace Protocol Deviation Plan will be summarized for all patients in the Randomized Population. The number and percentage of patients with a major/CSR-reportable protocol deviation will be summarized for each deviation term and overall, by treatment group and total. The deviations will also be listed by site and patient.

## 6.3 Demographics and Baseline Characteristics

Demographic and baseline characteristics will be provided for the Safety Population.

Descriptive statistics (n, mean, standard deviation, median, minimum, and maximum) for age, baseline weight, baseline body mass index (BMI), height at Screening, time since diagnosis for systemic sclerosis, time since onset for Raynaud's phenomenon and average number of times outdoor daily at baseline will be provided for each treatment group and in total. Counts and percentages of patients for sex, childbearing potential, ethnicity, race, smoking/alcohol use, phosphodiesterase-5 (PDE5) inhibitor use at screening, calcium channel blockers (CCB) use at screening, use both PDE5 and CCB at screening, type of systemic sclerosis and presence of digital ulcers will also be presented. The following baseline efficacy related information will also be provided:

- Weekly frequency of symptomatic RP attacks
- Average duration of symptomatic RP attacks
- Average weekly duration of symptomatic RP attacks
- Average Raynaud's condition score
- Average overall severity of RP attack symptoms
- Patient assessment of overall severity in symptomatic Raynaud's attacks (PGIS)
- Average worst pain\numbness\tingling\discomfort associated with symptomatic RP

Medical history will be coded using the latest version of Medical Dictionary for Regulatory Activities (MedDRA) and summarized by system organ class (SOC) and preferred term (PT) by treatment group and total. Medical history will also be listed. Medical history related to SSc and RP phenomenon will also be listed.

## 6.4 Concomitant Medications and Procedures

Medications will be coded using the latest version of the World Health Organization (WHO) Drug Dictionary version. Prior medications are any medications stopped prior to the first dose of study medication. Concomitant medications are any medications taken on or after the first dose of study medication. The number and percentage of patients taking each concomitant medication will be summarized by preferred term within each anatomical therapeutic chemical (ATC) classification for the Safety Population. This will only include concomitant medications collected in the electronic data capture (EDC) system. The analgesic log recorded in the ePRO diary is used in the following subsection.

### *6.4.1 Analgesic Use*

Patient analgesic medication use, both prescription and over-the-counter, is being captured daily using an analgesic patient log.

Descriptive statistics for analgesic use for each patient will be tabulated for the baseline (10 to 25 Days pre-treatment) and post treatment efficacy period (Day 8 to Day 21) by treatment group for the ITT Population. During each period, this will include:

i. Analgesic use frequency– the proportion of days with an analgesic medication used (prescription and over-the-counter medications will be summarized separately). This will be provided for each treatment group and in total. Analgesic use frequency will be calculated by scoring each day as a 'medication day' (if the patient takes analgesic mediation on that day) or a 'nonmedication day' (if the patient does not take analgesic mediation on that day) and then computing the proportion of medication days (range: 0 to 1.0) as number of 'medication day'/ (number of 'medication day' + number of 'nonmedication day'). If ePRO diary was not completed on a particular day, the day will be excluded from the calculation. Change and percent change from baseline will be provided. If the patient answers "Yes" for taking prescription or over-the-counter medications, we will use it as-is regardless of the input of medication name and dosage.
    a. An analysis of covariance (ANCOVA) model will be used to analyze the change from baseline, including randomized treatment group and randomized stratification (i.e., use of PDE5 inhibitors at screening) as factors and baseline analgesic use frequency as a covariate. The comparison between treatment groups will be estimated together with the 95% confidence interval and p-value. The

analysis will be performed on the ITT Population. Prescription and over-the-counter medications will be analyzed separately.

ii. Analgesic type – number and percentage of patients taking each analgesic medication will be provided by type (opioid and non-opioid analgesic), by ATC and preferred term. This will be summarized for baseline and post-treatment efficacy periods (Days 8 to 21) for each treatment group and in total. No minimal dose or number of days is required. If the patient enters the medication name, we will use it as-is unless the record does not have a corresponding ATC and preferred term.

iii. Average daily dose of opioid analgesic medication – to compare opioid analgesic daily doses more easily over time and across patients, the opioid analgesic medications will be converted to morphine milligram equivalents (MME), based on published Equianalgesic Opioid Dose Conversion tables (Appendix B). After the conversion, the mean MME (daily dose) will be summarized for baseline and post-treatment efficacy periods (Days 8 to 21) by treatment group for a subset of ITT Population patients who have taken opioid analgesic medication during baseline or post-treatment efficacy period. If the ePRO diary was not completed on a particular day, that day will be excluded from the calculation. If a patient does not have any opioid analgesic medication in a certain period, the daily dose for this patient will be zero. Change and percent change from baseline will also be provided for opioid analgesics. The analgesic log is entered by the patient and may contain incorrect information and/or duplicate records. The log will be reviewed by sponsor and medical monitor, and any incorrect/duplicate records will be flagged before database lock. Those flagged as duplicate records will not be used in the average daily dose calculation. Those flagged as incorrect will be reviewed and may be adjudicated (where possible) based on concomitant medications entered in EDC. If adjudication is not possible, the record will not be used in the average daily dose calculation. Records that will be excluded include those in which:

- the medication name entry is in conflict with EDC when the opioid component does not match the concomitant medication record (e.g. “hydrocodone” in analgesic log but oxycodone is confirmed in EDC then hydrocodone will be excluded whereas an entry of hydrocodone in the diary when hydrocodone-acetaminophen is confirmed in EDC will not be excluded)
- there is an "impossible" dose entry based on available medication forms and EDC entry (e.g. oxycodone 0.5 mg)
- there are zeros, blanks, "N/A", "mistake" (or similar text indicating error) or "none" in drug name, dose or frequency fields

a. An analysis of covariance (ANCOVA) model will be used to analyze the change from baseline, including randomized treatment group and randomized stratification (i.e., use of PDE5 inhibitors at screening) as factors and baseline daily dose of opioid analgesic medication as a covariate. The comparison between treatment groups will be estimated together with the 95% confidence interval and p-value. The analysis will be performed on the ITT Population.

### 6.5 Study and Drug Exposure

Study exposure will be calculated as the date of last visit – date of first dose + 1. Study exposure will be summarized with descriptive statistics by treatment group and in total.

The count and percentage of patients receiving study drug infusions and patients who have completed the infusions will be summarized by day (Day 1 through Day 5) for each treatment group and in total.

In addition, the count and percentage of patients will be summarized for each treatment group and in total:

- by the highest study drug infusion dose by day;
- by the last study drug infusion dose by day;
- by the number of days infusion is taken (initiated).

## 7 ANALYSIS OF EFFICACY

All available efficacy data (weekly frequency of symptomatic RP attacks, severity of symptomatic RP attacks, Raynaud's Condition Score, weekly total duration of symptomatic RP attacks, the duration of symptomatic RP attacks (average duration of an attack), overall change in symptomatic RP, overall patient severity in symptomatic RP attack) will be listed for the ITT Population. Efficacy measurements will be summarized with descriptive statistics (n, mean, standard deviation, median, minimum, and maximum) by treatment group (iloprost injection for IV use and placebo). Baseline, change from baseline, and percent change from baseline to endpoint will be presented for each efficacy measurement for the ITT Population.

A fixed sequential testing procedure will be implemented. In a hierarchical step-down manner, the primary efficacy endpoint will be tested at the 2-sided 0.05 level first, followed by testing the secondary efficacy endpoints at the 2-sided 0.05 level in the following hierarchical manner: (1) overall severity of RP attack symptoms, (2) weekly total duration of symptomatic RP attacks, (3) the proportion of responders, defined as patients that have a 50% reduction in weekly total duration of symptomatic RP attacks and 50% reduction in overall severity from baseline. Inferential conclusions about these efficacy endpoints will require statistical significance of the previous endpoints and the primary efficacy endpoint.

For statistical analysis, in the event a subject is stratified incorrectly, “randomized stratum” will be used rather than “actual stratum”.

## 7.1 Primary Efficacy Analysis

The primary efficacy endpoint is the change in symptomatic digital ischemic episodes, as measured by the weekly frequency of symptomatic RP attacks from baseline to the end of the efficacy follow-up.

The number and percentage of patients who have 0, 1, …, and 25-day entries of frequency of symptomatic RP attacks will be provided by treatment group on ITT Population for baseline period and efficacy period (Days 8 to 21). In addition, the number and percentage of patients that have at least 7 days of ePRO dairy days completed will be provided for baseline and post treatment efficacy period will also be provided.

The cumulative distribution of primary efficacy endpoint will be plot by randomized treatment group.

The primary analysis on the primary efficacy endpoint will be performed based on an analysis of covariance (ANCOVA) model, including randomized treatment group and randomized stratification (i.e., use of PDE5 inhibitors at screening) as factors and baseline weekly RP attacks as a covariate. The treatment comparisons will be estimated together with the 95% confidence interval and p-value. The primary analysis will be performed on the ITT Population.

For patients in the ITT Population with a missing primary efficacy endpoint, imputation rule in section 5.2.2 will be used.

The primary analysis will be repeated on the PP Population.

Season effect

The season effect will be explored according to the meteorological definition.

- Fall: September 1 to November 30;
- Winter: December 1 to February 28; and
- Spring: March 1 to May 31

The primary analysis will be repeated on ITT population with date of randomization by season and its interaction with treatment group as an additional fixed factor. The treatment comparisons for each season will be estimated together with the 95% confidence interval.

## 7.2 Secondary Efficacy Analyses

The secondary efficacy analyses will be performed on the ITT Population. Secondary efficacy endpoints in the hierarchical step-down testing procedure include changes from baseline to the end of the efficacy follow-up in the following:

(1) overall severity of RP attack symptoms, (2) weekly total duration of symptomatic RP attacks, (3) the proportion of responders, defined as patients that have a 50% reduction in weekly total duration of symptomatic RP attacks and 50% reduction in overall severity from baseline.

For patients in the ITT Population with a missing secondary efficacy endpoint, imputation rules (as outlined in section 5.2.2) will be used. Overall severity will be derived after the imputation for each symptom. Responders will be derived after the imputation of the missing weekly total duration of symptomatic RP attacks and overall severity. The calculation of each endpoint can be found in section 3.1.2.

### *7.2.1 Overall Severity of RP Attack Symptoms*

An analysis of covariance ANCOVA model will be used to analyze the change from baseline to the end of the efficacy follow-up in overall severity of RP attack symptoms and will include randomized treatment group and randomized stratification as factors and baseline value as a covariate. The treatment comparisons will be estimated together with the 95% confidence interval and p-value.

The cumulative distribution of overall severity of RP symptoms will be plot by randomized treatment group.

A sensitivity analysis will be performed to test the robustness of the endpoint and to assess the impact of using a rank order of symptoms to determine which symptom is used when more than 1 symptom has the same baseline NRS. When there are multiple symptoms with the same worst average (highest NRS) baseline value for a patient, one of these symptoms will be randomly selected to evaluate the patient's overall severity instead of using the order of rank.

Numbers 1, 2, 3 and 4 will be randomly assigned to the four symptoms for each patient (different patients may have different numbers assigned to the same symptom). When there are multiple symptoms with the same worst average baseline value for a patient, the one with the largest number (among these symptoms) will be used to evaluate the overall severity of this patient.

Random seed = 536512 will be used for the random assignment.

### *7.2.2 Weekly Total Duration of Symptomatic RP Attacks*

An analysis of covariance ANCOVA model will be used to analyze the change from baseline to the end of the efficacy follow-up in weekly total duration of symptomatic RP attacks and will include randomized treatment group and randomized stratification as factors and baseline value as a covariate. The treatment comparisons will be estimated together with the 95% confidence interval and p-value.

The cumulative distribution of weekly total duration of symptomatic RP attacks will be plot by randomized treatment group.

### *7.2.3 Proportion of Responders*

A logistic regression model will be used to estimate the proportion of responders within each treatment group. The model will include the randomized treatment group and randomized stratification as factors and the baseline (weekly total duration of symptomatic RP attacks and overall severity) as covariates. The odds ratio between iloprost and placebo will be provided together with the p-value.

## 7.3 Exploratory Efficacy Analyses

The exploratory efficacy endpoints include changes from baseline to the end of the efficacy follow-up in the following:

- Proportion of days without symptomatic RP attacks.
- NRS for worst pain associated with symptomatic RP attacks.
- NRS for worst numbness associated with symptomatic RP attacks
- NRS for worst tingling associated with symptomatic RP attacks
- NRS for worst discomfort associated with symptomatic RP attacks
- Proportion of pain responders, defined as patients that have a 50% reduction in worst pain associated.
- Proportion of numbness responders, defined as patients that have a 50% reduction in worst numbness associated.
- Proportion of tingling responders, defined as patients that have a 50% reduction in worst tingling associated.
- Proportion of discomfort responders, defined as patients that have a 50% reduction in worst discomfort associated.
- Proportion of total weekly duration responders, defined as patients that have a 50% reduction in weekly total duration of symptomatic attacks.
- Raynaud's Condition Score
- Duration of symptomatic RP attacks (average duration of an attack)
- Patient assessment of overall change in symptomatic Raynaud's attacks (PGIC) [PGIC itself is evaluating the change from baseline and thus does not need baseline value]
- Changes from baseline to the end of the efficacy follow-up in patient assessment of overall severity in symptomatic Raynaud's attacks (PGIS)
- Patient assessment of overall benefit compared to side effects

Additionally, depending on the results of the primary and secondary efficacy analyses, the effect of iloprost on plasma biomarkers may be evaluated. The analysis of the biomarkers will be documented separately.

Descriptive statistics (n, mean, standard deviation, median, minimum, maximum) will be provided by treatment group for the exploratory endpoints on ITT Population.

ANCOVA models will be used to analyze the continuous efficacy endpoints and will include randomized treatment group and randomized stratification as factors and baseline value as a covariate. The treatment comparisons will be estimated together with the 95% confidence interval. P-values will be provided.

PGIS and PGIC are considered as continuous variables for analyses. PGIS scores are converted as None = 0, Mild = 1, Moderate = 2, Severe = 3, Very severe = 4. PGIC scores are converted as Very much worse = -3, Moderately worse = -2, A little worse = -1, No change = 0, A little better = 1, Moderately better = 2, Very much better = 3. The ANCOVA model for PGIC will include only randomized treatment group and randomized stratification as factors since there is no PGIC value at baseline.

The same logistic model for secondary efficacy analysis will be used for the exploratory categorical endpoints on ITT Population. The odds ratio between iloprost and placebo will be provided together with the p-value.

Patients with missing value will be excluded. No imputation is needed.

### *7.3.1 Additional analyses*

Log transformed NRS for each symptom

The change from baseline of log transformed NRS for each symptom (pain / numbness/ tingling/ discomfort) associated with symptomatic RP attacks will be analyzed using ANCOVA separately, with the randomized treatment group and randomized stratification as factors and the log transformed baseline value as a covariate. Estimate of geometric mean (percent) change from baseline for each group and difference between treatment groups (percent difference comparing iloprost with placebo group in ratio [efficacy period to baseline]) will be provided with the 95% confidence interval. P-value will be provided for the difference.

For each symptom, if there is any average NRS at baseline or endpoint period for any patient equal to zero, then the logarithm of (NRS+1) will be used for this symptom.

Responders for each symptom

Responders defined as patients achieving at least 50% reduction in each symptom (original scale of worst pain/ numbness/ tingling/ discomfort), duration per RP attack and weekly duration of RP attacks will be analyzed by a logistic regression separately with responder as the dependent variable, randomized treatment group and stratification variable as factors and baseline value as a covariate. Odds ratio between treatment groups will be provided with confidence interval and p-value.

The followings will be analyzed similarly. For PGIS, the baseline score will be included as a factor. The modeling for PGIC will not have the baseline value as a factor.

PGIC: proportion of patients “Moderately better” or “Very Much Better”

PGIS: Proportion of patients who improved at least 2 points (e.g., from "Very severe = 4" to "Moderate = 2")

Digital Ulcers

Digital ulcers will be assessed at screening and captured as part of the medical history at this visit. New digital ulcers will be captured as an adverse event/SAE. In all cases, subjects that present with DUs will be followed until Day 35 (+7 days) (Visit 9) or the Early Termination visit. Number of digital ulcers present at Screening, number of digital ulcers present, number of digital ulcers, new digital ulcers and healed digital ulcers at Follow-up Visit or Early Termination Visit will be summarized by treatment group and total.
Digital ulcers assessments will be listed.

Outdoor times

The average number of times a patient goes outdoors daily will be summarized for baseline and post treatment efficacy period (Days 8 to 21) by treatment group for the ITT Population.

### 7.4 Subgroup Analysis

The primary efficacy endpoint will be summarized for the ITT population by systemic sclerosis subtype (limited cutaneous or sine, diffuse cutaneous), use of PDE5 inhibitors at screening (Yes/No), digital ulcer presents at screening, CCBs at screening, baseline weekly frequency (>75%, ≤75% percentile weekly attack frequency). ANCOVA model will be used to analyze the primary efficacy endpoint, which will include randomized treatment group, randomized stratification, subgroup, interaction of subgroup and randomized treatment group as factors, and baseline value as a covariate. Corresponding forest plot for treatment effect by subgroup will be provided.

Other subgroups such as by age (<65, ≥65), gender (male, female), race (white, nonwhite), ethnicity (Hispanic or Latino, Not Hispanic or Latino), baseline systolic blood pressure (<90 mmHg, ≥90 mmHg) and baseline kidney function [Estimated Glomerular Filtration Rate (eGFR)<90 mL/min/1.73 $m^2$, eGFR≥90 mL/min/1.73 $m^2$] will also be explored.

The corresponding subgroup analysis will be performed only if each category has at least 5 patients in each treatment group.

## 8 ANALYSIS OF SAFETY

The safety parameters include adverse events, physical examination findings, vital sign measurements (heart rate and blood pressure), 12-lead electrocardiogram findings, and standard clinical laboratory measurements (chemistry and hematology). All safety analyses will be performed on the Safety Population. Patients will be analyzed by treatment received.

## 8.1 Adverse Events

An overview of adverse events will be provided that includes the number and percentage of patients from the Safety Population by treatment group and in total for the following categories:

- Any TEAEs,
- Maximum severity/grade of TEAEs,
- Study drug-related TEAEs,
- Maximum severity/grade of study drug-related TEAEs,
- SAEs,
- TE-SAEs,
- Study drug-related TE-SAEs,
- AESI,
- TEAEs leading to study drug discontinuation,
- Study drug-related TEAEs leading to study drug discontinuation, and
- TEAEs leading to death.

Adverse events will be coded using the latest version of MedDRA. The number and percentage of patients with AEs will be summarized by SOC and PT for each treatment group and in total for the following:

- TEAEs,
- Study drug-related TEAEs,
- SAEs,
- TE-SAEs,
- Study drug-related TE-SAEs, and
- TEAEs leading to study drug discontinuation.

The number and percentage of patients with TEAEs will be summarized by reported maximum severity/grade for each MedDRA preferred term within system organ class for each treatment group and in total.

The number and percentage of patients with AESIs will be summarized by preferred term within system organ class for each treatment group and in total.

The number and percentage of patients with TEAEs that are present during efficacy period (Days 8 to 21) will be summarized by reported maximum severity/grade for each MedDRA preferred term within system organ class for each treatment group and in total, including AEs start during the period or start before Day 8 but end after Day 8. The number of AEs will be presented.

All SAEs, TEAEs leading to study drug discontinuation, and TEAEs leading to death will be listed.

## 8.2 Clinical Laboratory Evaluations

Descriptive statistics for each chemistry and hematology parameter will be presented by treatment group and total for baseline, each visit, and the change from baseline, excluding the local laboratory values. The eGFR will be calculated using the following Modification of Diet in Renal Disease (MDRD) equation instead of using the value from laboratory. The LLN of 60 mL/min/1.73 $m^2$ will be used for eGFR.

GFR (mL/min/1.73 $m^2$) = $175 \times (S_{cr})^{-1.154} \times (Age)^{-0.203} \times$ (0.742 if female) × (1.212 if African American)

The number and percentage of patients within each category will be summarized at each visit by treatment group for alanine aminotransferase, aspartate aminotransferase, alkaline phosphatase with categories (<LLN, Normal, >ULN to <=2xULN, >2xULN to <=3xULN, >3xULN), total bilirubin (<=1.1xULN, >1.1xULN) and eGFR (Normal, <LLN).

The number and percentage of patients within each category will be summarized at each visit by treatment group for glucose, bicarbonate, albumin, calcium, blood urea nitrogen, creatinine, chloride, potassium, sodium, total protein, hematocrit, hemoglobin, platelets, mean corpuscular volume, red blood cell count, white blood cell count and differential with categories (Normal, >ULN or <LLN).

The number and percentage of patients within each category will be summarized at each post baseline visit by treatment group for eGFR (decrease from baseline <50%, >=50%). Subjects with laboratory parameter values for hemoglobin, platelets, white blood cells, lymphocytes and neutrophils that meet the CTCAE version 5.0 criteria in Appendix C will be provided in the listing.

If there are more than 5 patients with abnormal post baseline values (less than LLN or larger than ULN) for a parameter (excluding eGFR), a shift table may be provided. The abnormal values will be flagged in the listing.

## 8.3 Vital Signs

Vital signs include height, weight, BMI, heart rate and blood pressure. Vital signs parameters will be summarized using descriptive statistics for each visit for each treatment group. The change from baseline will also be summarized. In addition, blood pressure and heart rate on Day 1 to Day 5 will be summarized by time point (pre-infusion, first measurement during infusion, last measurement after infusion stop for Day 1 to Day 5 and last measurement during infusion for Day 1), including change from baseline and change from pre-infusion on the same day. The pre-infusion value is the last measurement before the infusion. The first measurement during infusion should be within 1-hour of infusion initiation to be included in the summary.

## 8.4 ECG Parameters

The overall interpretation of the ECG at Screening will be summarized with counts and percentages for each treatment group and in total. All ECG parameters will be listed.

### 8.5 Physical Examination

Physical examination data will be listed.

Prior and concomitant medications and procedures will be listed.

## 9 COVID-19 IMPACT

The e-CRF was amended in order to collect data related to COVID-19 impact on trial conduct (e.g., missed study visits or study discontinuations due to COVID-19). This information will be summarized in the clinical study report.

### 9.1 Disposition

To assess the impact of COVID-19 pandemic on the trial, the following summaries will be provided on ITT Population:

- Proportion of patients impacted and withdrawn due to COVID-19 pandemic.
    - the reasons of discontinuation related to COVID-19 will be presented separately.
    - the number of patients affected (patients ongoing or newly enrolled in the study during the COVID-19 pandemic, i.e. on or after 27Jan2020); impacted (i.e. having visits impacted by the COVID-19 pandemic) or infected will also be presented.
- Number of patients with visits impacted by COVID-19 with the corresponding reason.
    - the number of patients impacted due to COVID-19 with the reasons by visit will be presented to assess the impact of COVID-19.
    - the number of patients with visit performed in person at a location other than the research site (home, ambulatory infusion suite, other), performed via virtual method (phone, video, other) or out of window due to COVID-19 will be presented by visit.
- The COVID-19 impacted subjects/visits will be flagged in the listing of visits, study discontinuations and protocol deviations.

In addition, the reason for screen failure related to COVID-19 will be summarized.

### 9.2 Demographic, Baseline Characteristics and Efficacy Endpoints

In this study, the majority of participants were enrolled after the public health emergency related to COVID-19 was declared by the Secretary of Health and Human Services (HHS) on January 31, 2020, effective January 27, 2020, therefore Demographic and Baseline Characteristics will only be provided for the ITT population. Furthermore, as the data for the primary and key secondary endpoints are collected in the ePRO daily diary, COVID-19 related disruption to efficacy data collection is expected to be minimal. As such, any missing data due to COVID-19 will be handled using existing missing data handling rule in the corresponding sections.

### 9.3 Sensitivity analysis

If there are more than 10 patients with 2 or more missing infusion days or with fewer than 7 days of completed ePRO diaries in the post treatment efficacy period (Days 8 to 21) because of COVID-19 pandemic, the primary and key secondary efficacy endpoints analyses will be repeated on ITT Population excluding these patients.

## 10 DATA MONITORING COMMITTEE AND INTERIM ANALYSES

A DMC consisting of 3 members including an SSc expert, a prostacyclin PAH expert, and an independent biostatistician.

The DMC will be responsible for safeguarding the interests of clinical study patients and for enhancing the integrity of the study. To address this mission, the DMC will have ongoing access to efficacy and safety data, and information regarding the quality of study conduct. The DMC will review safety information on a monthly basis. The DMC will also have a planned formal interim analysis meeting, where available efficacy data will be reviewed to enable the interpretation of safety in the context of this emerging efficacy data, and to assess the quality of its capture. The meeting will occur when approximately 80 to 90 patients have been randomized. In addition, the DMC will hold ad hoc teleconference meetings to discuss safety or study conduct information as needed. Based on its insights from emerging evidence, the DMC will provide recommendations to the study Sponsor, including a recommendation regarding study continuation. Given the very short timeframe for the study and the importance of robust evidence, there will not be formal statistical boundaries for early termination for efficacy. In assessing the acceptability of the safety profile, the DMC will consider the totality of information regarding benefits and risks.

To contribute to enhancing the integrity of the study, the DMC may also formulate recommendations relating to the rates of recruitment and eligibility of patients, improving adherence to protocol-specified regimens (e.g., ePRO adherence, study drug infusions, etc.), retention of patients, and the timeliness of data capture and adjudication of study endpoints. The DMC will be advisory to the study Sponsor (and/or the clinical study leadership group, hereafter referred to as the Steering Committee). The Sponsor and/or the Steering Committee will be responsible for promptly reviewing the DMC recommendations, discussing them with the DMC, and making decisions about their implementation.

A separate charter further describes the DMC role.

## 11 STEERING COMMITTEE

The purpose of the Steering Committee is to provide objective and independent scientific and medical input on the design, execution, analysis, and reporting of the study. The Steering Committee will consist of members with medical specialties or scientific expertise (e.g., statisticians) pertinent to the population under study. There will be 7 voting external members of the Steering Committee who are independent of Eicos Sciences, Inc. The Steering Committee statisticians and Sponsor representatives will be nonvoting members. A separate charter describes the Steering Committee role.

## 12 SAMPLE SIZE AND POWER CONSIDERATIONS

The study will target randomizing 180 patients, with the intention to achieve this target by 31 March 2021.

- At any time prior to 31 March 2021, if there are 180 patients randomized, the recruitment for patients will stop. The patients already in screening period would continue to randomization only if such randomization would occur prior to 31 March 2021; such patients, post randomization, would continue through the post treatment efficacy period.
- If there are fewer than 180 patients randomized by 31 March 2021, then randomization will stop on that date.

The study will be stratified based on the current use of oral PDE5 inhibitors. 86 patients per arm, at the end of the efficacy evaluation time point, will provide ≥85% power to detect a 5.5 improvement between iloprost and the placebo group for the mean change from baseline to the end of the efficacy follow-up in the weekly frequency of symptomatic RP attacks, assuming a common standard deviation of 12 and use of a traditional 2-sided 0.05 level test. Assuming a rate of missingness of 5% before Day 8, 90 patients in each treatment group are planned.

## 13 PROGRAMMING SPECIFICATIONS

All available data will be presented in patient data listings, which will be sorted by site number, unique patient identifier and where appropriate, visit number and visit/assessment date.

The programming specifications, including the mock-up validity listings, analysis tables, figures, and data listings, as well as the derived database specifications, will be prepared in stand-alone documents.

# 14 APPENDIX A: LABORATORY TESTS

**Standard Safety Chemistry Panel**

| | |
|---|---|
| Alanine aminotransferase | Albumin |
| Alkaline phosphatase | Aspartate aminotransferase |
| Bicarbonate | Blood urea nitrogen |
| Calcium | Chloride |
| Creatinine | Estimated glomerular filtration rate |
| Glucose | Potassium |
| Sodium | Total bilirubin |
| Total protein | |

**Hematology**

| | |
|---|---|
| Hematocrit | Hemoglobin |
| Mean corpuscular volume | Platelets |
| Red blood cell count | |
| White blood cell count and differential [1] | |

1. Manual microscopic review is performed only if white blood cell count and/or differential values are out of reference range.

**Pregnancy test**

Urine pregnancy test (for women of childbearing potential only)

# 15 APPENDIX B: CALCULATING THE TOTAL DAILY DOSE OF OPIOIDS

To calculate the total daily dose of opioids, refer to https://www.cdc.gov/drugoverdose/pdf/calculating_total_daily_dose-a.pdf and https://www.cms.gov/Medicare/Prescription-Drug-coverage/PrescriptionDrugCovContra/Downloads/Opioid-Morphine-EQ-Conversion-Factors-Aug-2017.pdf:

1. Determine the total daily amount of each opioid medication the patient takes. Calculate the morphine-equivalent dose: Use the equianalgesic dose ratio provided in the table below corresponding to the current drug/route of administration.
2. Convert each to morphine milligram equivalents (MME)—multiply the dose for each opioid by the conversion factor. (see table below)
3. Add them together.

| OPIOID (doses in mg/day except where noted) | CONVERSION FACTOR |
|---|---|
| Codeine | 0.15 |
| Fentanyl transdermal (in mcg/hr) | 2.4 |
| Hydrocodone | 1 |
| Hydromorphone | 4 |
| Methadone | |
| 1-20 mg/day | 4 |
| 21-40 mg/day | 8 |
| 41-60 mg/day | 10 |
| > 60 mg/day | 12 |
| Morphine | 1 |
| Oxycodone | 1.5 |
| Oxymorphone | 3 |
| Tramadol (mg) 0.1 | |

Example: patient taking one tablet of Oxycodone Hydrochloride and Acetaminophen 5 mg/325 mg, two (2) times per day

1. Total daily dose of Oxycodone is 5 mg x 2 = 10 mg
2. MME = 10 mg x 1.5 = 15 mg

# 16 APPENDIX C: RELATED CTCAE CRITERIA

| Parameter | Grade 1 | Grade 2 | Grade 3 | Grade 4 |
|---|---|---|---|---|
| Hemoglobin (Hgb) | <LLN-10.0 g/dL;<br><LLN-6.2 mmol/L;<br><LLN – 100 g/L | <10.0-8.0 g/dL;<br><6.2-4.9 mmol/L;<br><100-80 g/L | <8.0 g/dL;<br><4.9 mmol/L;<br><80 g/L | |
| Platelets | <LLN-75000/$mm^3$;<br><LLN-75.0*$10^9$/L | <75000-50000/$mm^3$;<br><75.0-50.0 * $10^9$/L | <50000-25000/$mm^3$;<br><50.0-25.0*$10^9$/L | <25000/$mm^3$;<br><25.0*$10^9$/L |
| White blood cells | <LLN-3000/$mm^3$;<br><LLN-3.0*$10^9$/L | <3000-2000/$mm^3$;<br><3.0-2.0*$10^9$/L | <2000-1000/$mm^3$;<br><2.0-1.0*$10^9$/L | <1000/mm3;<br><1.0*$10^9$/L |
| Lymphocytes | <LLN-800/$mm^3$;<br><LLN-0.8*$10^9$/L | <800-500/$mm^3$;<br><0.8-0.5*$10^9$/L | <500-200/$mm^3$;<br><0.5-0.2*$10^9$/L; | <200/$mm^3$;<br><0.2*$10^9$/L; |
| Lymphocytes | | >4000/$mm^3$-<br>20000/$mm^3$ | >20000/$mm^3$ | |
| Neutrophils | <LLN-1500/$mm^3$;<br><LLN-1.5*$10^9$/L | <1500-1000/$mm^3$;<br><1.5-1.0*$10^9$/L | <1000-500/$mm^3$;<br><1.0-0.5*$10^9$/L | <500/$mm^3$;<br><0.5*$10^9$/L |
| LLN / ULN = lower / upper limit of the normal range. | | | | |

# 17 REFERENCE

1. Wigley, Fredrick M., et al. "Intravenous iloprost infusion in patients with Raynaud phenomenon secondary to systemic sclerosis: a multicenter, placebo-controlled, double-blind study." Annals of Internal Medicine 120.3 (1994): 199-206.