



**Protocol No. DMVT-505-3002**  
**IQVIA Biotech Study No. FZA91720**

**A PHASE 3 EFFICACY AND SAFETY STUDY OF TAPINAROF FOR THE  
TREATMENT OF PLAQUE PSORIASIS IN ADULTS**

**Statistical Analysis Plan Amendment 2, version 3.0**

Document last saved: 4-May-2020 2:22 PM

## Revision History

Date	Reason for update	Section Updated
05AUG2019	Original Version 1.0	NA
04OCT2019	Amendment 1 Version 2.0	<ul style="list-style-type: none"><li>• Section 9 Handling of Missing Data:<ul style="list-style-type: none"><li>○ Added LOCF and TF imputation method to handle missing efficacy data.</li></ul></li><li>• Section 12.1 Subject Disposition:<ul style="list-style-type: none"><li>○ Added summary for screen failures along with screen failure reasons.</li></ul></li><li>• Section 12.6 Medical History and Concurrent Procedures:<ul style="list-style-type: none"><li>○ Added a summary for psoriasis history, cardiovascular risk factors and liver disease family history;</li><li>○ Updated duration of disease categories from “&lt;1, 1-5 years and 5 years” to “&lt;5, 5-10 years, and 10 years”</li></ul></li><li>• Section 12.7.3 Exploratory Efficacy Endpoints:<ul style="list-style-type: none"><li>○ Added endpoint mean % change in % BSA affected from baseline to weeks 2, 4, and 8</li></ul></li><li>• Section 12.7.4 Functional Outcomes and Health-related Quality of Life Endpoints:<ul style="list-style-type: none"><li>○ Added description for PSD total score calculation.</li><li>○ Clarified the RAND SF-36 version was used in the study.</li></ul></li><li>• Section 12.8 Efficacy Analyses:<ul style="list-style-type: none"><li>○ Added LOCF on ITT population as sensitivity analysis for primary and secondary endpoints.</li><li>○ Added additional subgroup analysis for primary and all secondary endpoints: age group, sex, race, duration of disease, total %BSA affected, %BSA affected by body region.</li></ul></li><li>• Section 12.9.1 Adverse Events:<ul style="list-style-type: none"><li>○ Added treatment-related serious TEAE and serious TEAE leading to study drug discontinuation to overall AE summary.</li><li>○ Added additional summaries for treatment-</li></ul></li></ul>

Date	Reason for update	Section Updated
04MAY2020	Amendment 2 Version 3.0	<p>related TEAEs; TEAEs leading to study discontinuation; serious TEAEs</p> <ul style="list-style-type: none"> <li>○ Added subgroup analysis for TEAEs and serious TEAEs: age group, sex and race</li> <li>○ Clarified duration summary for AESIs.</li> </ul> <ul style="list-style-type: none"> <li>● Added additional summary and analysis due to the impact of the novel Coronavirus Disease 2019 (COVID-19):           <ul style="list-style-type: none"> <li>○ Section 12.11 Additional Analyses – COVID-19</li> </ul> </li> <li>● Section 8 Analysis Populations:           <ul style="list-style-type: none"> <li>○ Added a sentence that additional criteria that are not documented as protocol deviations may be added to determine PP population prior to unblinding the study database.</li> <li>○ Updated exclusion criteria to specify use of oral corticosteroids</li> </ul> </li> <li>● Section 9 Handling of Missing Data:           <ul style="list-style-type: none"> <li>○ Clarified baseline PGA score and baseline value of the corresponding endpoint are included in multiple imputation model</li> <li>○ Updated the example SAS code for multiple imputation</li> </ul> </li> <li>● Section 11.4 Analysis Visit Window:           <ul style="list-style-type: none"> <li>○ Updated analysis visit window for Week 12 and Week 16</li> </ul> </li> <li>● Section 12.4 Study Exposure and Compliance:           <ul style="list-style-type: none"> <li>○ Renamed the parameter “number of days actually dosed” to “number of doses administered”</li> </ul> </li> <li>● Section 12.7.4 Functional Outcomes and Health-related Quality of Life Endpoints:           <ul style="list-style-type: none"> <li>○ DLQI: Added texts for handling of missing data in domain scores</li> <li>○ SF-36: Removed physical component score and mental component score; Added 8 domain scores.</li> </ul> </li> <li>● Section 12.8 Efficacy Analyses:           <ul style="list-style-type: none"> <li>○ Added percent change from baseline for peak</li> </ul> </li> </ul>

Date	Reason for update	Section Updated
		<p>NRS, DLQI and PSD.</p> <ul style="list-style-type: none"><li>○ Added a new responder analysis for Peak NRS</li><li>○ Added new analysis for itch scores (in-clinic and at-home diary): change from baseline, treatment difference and a plot</li><li>○ Added country as new subgroup analysis</li><li>○ Added a Kaplan-Meier figure for time to PGA success</li></ul> <ul style="list-style-type: none"><li>● Section 12.9.1 Adverse Events:<ul style="list-style-type: none"><li>○ Clarified all AEs with onset on the Week 12 visit will be included in the summary</li><li>○ Added more extensive summary for each AESI.</li><li>○ Added Kaplan-Meier figure for each AESI of time to first event</li></ul></li><li>● Section 12.9.2 Local Tolerability Scale (LTS):<ul style="list-style-type: none"><li>○ Added a summary for LTS at the sensitive areas</li></ul></li><li>● Section 12.9.3 Clinical Laboratory Testing:<ul style="list-style-type: none"><li>○ Clarified normal reference range and reference range indicators are provided by central lab.</li><li>○ Clarified only urinalysis lab tests with numeric results reported are included in shift tables.</li><li>○ Added mapping table between central lab reported reference range indicators and indicators in shift tables</li></ul></li><li>● Section 12.10 Pharmacokinetic Analyses:<ul style="list-style-type: none"><li>○ Added additional summary for PK concentration divided by baseline total %BSA</li><li>○ Added subgroup summaries for PK concentrations</li><li>○ Removed geometric mean from the PK concentration summaries.</li></ul></li><li>● Section 13 Changes From the Protocol and Planned Analyses:<ul style="list-style-type: none"><li>○ Updated to indicate SF-36 analysis changed from the protocol</li><li>○ Clarified additional analyses added due to COVID-19</li></ul></li></ul>

## Table of Contents

Revision History .....	2
Table of Contents .....	5
List of Abbreviations and Definitions of Terms .....	7
1. INTRODUCTION.....	9
2. STUDY OBJECTIVES.....	9
3. STUDY DESIGN.....	9
4. HARDWARE AND SOFTWARE.....	10
5. DATABASE LOCK.....	10
6. RANDOMIZATION.....	10
7. SAMPLE SIZE DETERMINATION.....	11
8. ANALYSIS POPULATIONS .....	11
9. HANDLING OF MISSING DATA .....	12
10. INTERIM ANALYSIS .....	14
11. DATA CONVENTIONS FOR ANALYSIS.....	14
11.1 General Statistical Principles.....	14
11.2 Study Day.....	14
11.3 Baseline and Change from Baseline .....	14
11.4 Analysis Visit Window.....	14
11.5 Study Centers .....	15
11.6 Multiple Comparisons .....	15
12. STATISTICAL EVALUATION.....	15
12.1 Subject Disposition.....	15
12.2 Protocol Deviations .....	16
12.3 Demographic and Baseline Characteristics.....	16
12.4 Study Medication Exposure and Compliance .....	16
12.5 Prior and Concomitant Medications.....	17
12.6 Medical History and Concurrent Procedures .....	17
12.7 Efficacy Endpoints .....	18
12.7.1 Primary Efficacy Endpoint .....	18
12.7.2 Secondary Efficacy Endpoints.....	18
12.7.3 Exploratory Efficacy Endpoints .....	18

---

<b>12.7.4</b>	<b>Functional Outcomes and Health-related Quality of Life Endpoints</b> .....	<b>19</b>
<b>12.8</b>	<b>Efficacy Analyses</b> .....	<b>20</b>
<b>12.8.1</b>	<b>Primary Efficacy Analyses</b> .....	<b>21</b>
<b>12.8.2</b>	<b>Secondary and Exploratory Efficacy Analyses</b> .....	<b>22</b>
<b>12.8.3</b>	<b>Other Efficacy Analyses</b> .....	<b>23</b>
<b>12.8.4</b>	<b>Subgroup Analyses</b> .....	<b>23</b>
<b>12.9</b>	<b>Safety Analysis</b> .....	<b>23</b>
<b>12.9.1</b>	<b>Adverse Events</b> .....	<b>23</b>
<b>12.9.2</b>	<b>Local Tolerability Scale (LTS)</b> .....	<b>25</b>
<b>12.9.3</b>	<b>Clinical Laboratory Testing</b> .....	<b>25</b>
<b>12.9.4</b>	<b>Vital Signs</b> .....	<b>26</b>
<b>12.9.5</b>	<b>Physical Examination</b> .....	<b>26</b>
<b>12.10</b>	<b>Pharmacokinetic Analyses</b> .....	<b>26</b>
<b>12.11</b>	<b>Additional Analyses - COVID-19</b> .....	<b>27</b>
<b>13.</b>	<b>CHANGES FROM THE PROTOCOL AND PLANNED ANALYSES</b> .....	<b>27</b>
<b>14.</b>	<b>FORMATTING</b> .....	<b>27</b>
<b>15.</b>	<b>ARCHIVING AND RETENTION OF DOCUMENTS</b> .....	<b>28</b>

## List of Abbreviations and Definitions of Terms

<b>Abbreviation</b>	<b>Definition</b>
ADaM	Analysis dataset model
AE	Adverse event
AESI	Adverse event of special interest
ANCOVA	Analysis of covariance
ATC	Anatomical Therapeutic Chemical Classification System
BLQ	Below limit of quantitation
BMI	Body mass index
BSA	Body surface area
CI	Confidence interval
CMH	Cochran-Mantel-Haenszel
COVID-19	Coronavirus Disease 2019
DLQI	Dermatology Life Quality Index
eCRF	Electronic Case Report Form
EOS	End of study
HADS	Hospital Anxiety and Depression Scale
ITT	Intent-to-Treat
LOCF	Last observation carried forward
LTS	Local Tolerability Scale
MedDRA	Medical Dictionary for Regulatory Activities
MI	Multiple imputation
NRS	Numeric rating scale
OC	Observed cases
ODS	Output Delivery System
PASI	Psoriasis Area and Severity Index
PGA	Physician's Global Assessment
PK	Pharmacokinetic
PP	Per protocol
PSD	Psoriasis Symptom Diary
PT	Preferred Term
QoL	Quality of Life
RTF	Rich-text-formatted
SAE	Serious adverse event
SAP	Statistical analysis plan
SD	Standard deviation
SDTM	Study data tabulation model
SE	Standard error

---

<b>Abbreviation</b>	<b>Definition</b>
SF-36	36-item Short Form Survey
SOC	System Organ Class
TEAE	Treatment-emergent adverse event
TF	Treatment failure
WHO-DD	World Health Organization – Drug Dictionary



## **1. INTRODUCTION**

This Statistical Analysis Plan (SAP), based on the study protocol amendment 1 dated May 8<sup>th</sup>, 2019, provides additional details concerning the statistical analyses outlined in the protocol and reflects any changes to the protocol from amendments. This plan will not repeat all the definitions given in the protocol but will provide further details of the summaries and analyses planned therein.

The scope of this plan includes analyses pertaining to all data collected in the Double-Blind, Vehicle-Controlled Phase of the study. Analyses pertaining to data collected in the Open-Label Long-Term study will be described in a separate plan.

## **2. STUDY OBJECTIVES**

The primary objective is:

- To evaluate the efficacy of tapinarof cream, 1% compared with vehicle control in adults with plaque psoriasis.

The secondary objectives are:

- To further characterize the efficacy of tapinarof cream, 1% compared with vehicle control over time
- To describe the effect of tapinarof cream, 1% on psoriasis symptom severity and the associated impact on daily activities and attitudes in adults with plaque psoriasis
- To evaluate the safety and tolerability of tapinarof cream, 1% in adults with plaque psoriasis
- To characterize the pharmacokinetics (PK) of tapinarof in adults with plaque psoriasis

## **3. STUDY DESIGN**

This is a double-blind, randomized, vehicle-controlled, Phase 3, multicenter study to evaluate the efficacy and safety of topical tapinarof cream, 1% compared with vehicle-control cream in adults with plaque psoriasis.

Following a 34-day screening period, eligible subjects will be randomized at a 2:1 ratio to receive once daily treatment with tapinarof cream, 1% or vehicle cream. Subjects will return to the clinic at Weeks 2, 4, 8, and 12 for efficacy and safety assessments. Additionally, subjects will be contacted by phone at Weeks 6 and 10 to assess adverse events (AEs) and concomitant medications, to review study drug administration instructions, and to confirm subject's continued participation in this study. Study drug will be dispensed to subjects during the clinic visits and will be administered at home between clinic visits as instructed by site personnel. Subjects will

be instructed to apply study drug once daily to all affected areas, including newly appearing lesions and lesions/areas that improve during the study. Subjects will apply sufficient study drug to cover completely each lesion with a thin layer of study drug and will record the time of study drug application in a daily diary provided by the study site. (Note that subjects are allowed, but not required, to treat fingernails, toenails, palms, soles, and scalp lesions with study drug; however, efficacy analyses will not include assessment of improvement of psoriasis in these areas.) Subjects will be instructed to maintain the dosing time chosen at the beginning of the study for their full study participation. At the phone contacts at Weeks 6, 10, and 14, subjects should be reminded to complete their daily diary and bring it with them to the next clinic visit.

At the end of 12 weeks of treatment in this study, subjects will have the option to enroll in a separate open-label long-term safety and efficacy study for an additional 40 weeks of treatment. Subjects who choose not to participate in the open-label long-term study, who fail to qualify for participation in the open-label long-term study, or who qualify to participate in the open-label long-term study but ultimately elect not to enroll in that study, will complete a Follow-up Visit approximately 4 weeks (Week 16 visit) after the end of treatment in this study. Subjects who withdraw from the study before Week 12 will complete an Early Termination Visit.

#### **4. HARDWARE AND SOFTWARE**

Statistical analysis will be performed following IQVIA Biotech standard operating procedures and on the IQVIA Biotech computer network. All statistical analysis will be performed using SAS Version 9.4 or higher with program code prepared specifically for the project by qualified IQVIA Biotech statisticians and SAS programmers.

#### **5. DATABASE LOCK**

After completion of all data review procedures, validation of the project database, and approval of the data review document by the study sponsor, the clinical database will be locked and the study will be unblinded. Any change to the clinical database after this time will require written authorization, with explanation, by the Sponsor and the project Biostatistician.

#### **6. RANDOMIZATION**

For the double-blind, vehicle-controlled phase of the study, subjects will be randomized at a ratio of 2:1 to receive tapinarof cream, 1% or vehicle cream as follows:

<b>Randomization in Double-Blind, Vehicle-Controlled Phase</b>	
<b>Regimen</b>	<b>Number of Subjects</b>
Tapinarof cream, 1% once daily for 12 weeks	Approximately 333 Subjects
Vehicle cream once daily for 12 weeks	Approximately 167 Subjects

Randomization will be stratified by baseline PGA score so that subjects with mild and severe psoriasis (PGA scores of 2 and 4, respectively) will be limited to approximately 10% each of the total randomized population, and approximately 80% will have a PGA score of 3, signifying moderate disease.

The randomization lists will be generated using a validated system, which involves a pseudo-random number generator so that the resulting sequence of treatments will be both reproducible and non-predictable. Access to the codes will be controlled and documented.

## 7. SAMPLE SIZE DETERMINATION

It is assumed that the proportion of subjects who achieve PGA score of clear (0) or almost clear (1) with a minimum 2-grade improvement from Baseline at Week 12 will be 40% for subjects receiving tapinarof compared with 15% for subjects receiving vehicle control. With 500 subjects randomized in a 2:1 ratio (approximately 333 subjects receiving tapinarof cream, 1% and approximately 167 subjects receiving vehicle cream) there will be > 99% power for statistical significance (2-sided  $p < 0.05$ ). If tapinarof response rate is 35%, and vehicle response rate is 20%, the power will be > 94%. Power comes from a Fisher Exact sample size calculation which is conservative. It is assumed that up to 25% of the subjects will be lost to follow-up by 12 weeks. These subjects will be included in the primary analysis in the ITT population using the multiple imputation method as detailed below.

## 8. ANALYSIS POPULATIONS

The following populations are defined:

- Safety population: All randomized subjects who receive at least 1 application of study drug will be included in the Safety population. Subjects will be analyzed as treated.
- Intent-to-treat (ITT) population: All randomized subjects will be included in the Intent-to-treat population. Subjects will be analyzed as randomized.
- Per Protocol (PP) population: All subjects in the ITT population who did not have any major protocol deviations or other events that may impact the interpretation of the primary efficacy endpoint will be included in the PP population. Exclusion criteria for the PP Population may include:
  - Failure to meet the enrollment inclusion/exclusion criteria
  - On-study use of oral corticosteroids for >3 days and/or at >10mg dose
  - Failure to have a Week 12 assessment that occurs  $\leq 3$  days post last dose and also between study day 78 to 91 (inclusive) for the primary efficacy endpoint. Alternative days post dose and visit window may be used upon review of the actual blinded data before database lock.

- Failure to be compliant with treatment regimen (as defined in Section 12.4, i.e., compliance < 80%)

Additional criteria that may or may not be documented as protocol deviations may be added to determine PP population. The composition of the PP population will be determined and documented in blind reviews of the database conducted prior to unblinding the study database.

The ITT and PP Populations will be used for the analyses of efficacy endpoints. The ITT population is considered as the primary population for efficacy analyses and the PP will be considered as supportive. If a subject receives an incorrect study treatment other than the intended study treatment, the analysis of that subject's data will be based on the actual treatment received for the safety and PP analyses and the randomized assignment for the ITT analyses. If a subject receives more than 1 study treatment, the treatment group assignment for the safety and PP analyses will be determined on a case by case basis.

- Pharmacokinetic (PK) population: All subjects who undergo plasma PK sampling and have evaluable concentration-time data for analysis will be included in the PK population. A sample that is below the limit of quantification (BLQ) of the assay is considered evaluable.

## 9. HANDLING OF MISSING DATA

### Imputation of Start and End Dates of Adverse Event (AE)

To calculate duration of Adverse Events, the following rules will be used where applicable to impute partial or completely missing start dates or end dates:

- If only the day is missing for a start date, the 1<sup>st</sup> of the month will be imputed. If the new estimated date falls before the date of first dose, while the known month and year match the month and year of the first dose, the date of first dose will be used as the new estimated date. The AE will be considered as a treatment-emergent AE (TEAE).
- If only the day is missing for an end date, the last day of the month will be imputed. If the new estimated date falls after the date of last study visit, the date of last study visit will be used as the new estimated date. Last study visit is defined as the Week 12 visit for those who elect to participate in the open-label long-term study, or the Week 16 Follow-up visit for those who fail to qualify for participation in the open-label study, or who qualify to participate but ultimately elect not to enroll.
- If both the day and the month are missing for a start date or end date, no imputation will be used, and the duration will not be calculated. However, if the year of start is the same or greater than the year of the first dose date, the AE will be considered as treatment emergent.

- If the start date or end date is completely missing, duration will not be calculated. However, an event with completely missing start date will be considered as treatment emergent.

#### Imputation of Missing Efficacy Data

The primary method of handling of missing efficacy data will utilize Multiple Imputations (MI). Last observation carried forward (LOCF) method will be used as supportive analysis of the primary and secondary endpoints. For sensitivity analysis of the primary and secondary endpoints, missing data will be imputed as Treatment Failures (TF). Only missing data up to Week 12 will be imputed. Missing data at the Week 16 Follow-up visit will be analyzed as Observed Cases (OC) only.

**MI:** For the MI model, 100 imputations will be generated using PROC MI of SAS. Fully Conditional Specification (FCS) model using the regression method will be used with the response (e.g., PGA score) at prior post-baseline visits, baseline PGA score, baseline value of the corresponding endpoint, and treatment group as covariates, The ROUND and MINIMUM options will be utilized to ensure imputed values are non-negative integers. The seed to be used in all MI model is 20180330.

One set of MI datasets will be generated for the PGA scores, PASI scores and %BSA results separately. Response status (Yes, No) will be derived from the imputed scores in each MI dataset. Analysis of each primary, secondary and exploratory outcome will then be performed for each of the relevant MI datasets. The results of the 100 analyses will be transformed into a normal statistic and combined into a single analysis using PROC MIANALYZE.

Example SAS code for the PGA scores is as follows:

```
PROC MI DATA = PGA1 OUT = PGA2 SEED = 20180330 NIMPUTE = 100 NOPRINT;  
  CLASS TRTPN BASEPGA;  
  FCS LOGISTIC(BASEPGA = TRTPN / DETAILS);  
  FCS REG(PGA3 = TRTPN BASEPGA / DETAILS);  
  FCS REG(PGA4 = TRTPN BASEPGA PGA3 / DETAILS);  
  FCS REG(PGA5 = TRTPN BASEPGA PGA3 PGA4 / DETAILS);  
  FCS REG(PGA6 = TRTPN BASEPGA PGA3 PGA4 PGA5 / DETAILS);  
  VAR TRTPN BASEPGA PGA3-PGA6;  
RUN;
```

The imputation models may be modified based on the actual data if there is an issue in model convergence.

**LOCF:** The last observed value will be carried forward for any subsequent missing values. Baseline values will not be carried forward.

**TF:** missing data will be imputed as treatment failures for responder endpoint.

Missing functional outcomes, quality of life endpoints or safety endpoints will not be imputed.

## **10. INTERIM ANALYSIS**

No interim analysis is planned.

## **11. DATA CONVENTIONS FOR ANALYSIS**

### **11.1 General Statistical Principles**

All statistical tests will be performed at the 0.05 (two-sided) level of significance.

All observed and derived variables (e.g., change from baseline, percentage change from baseline, and response status) that are analyzed or summarized will be listed by subject. Descriptive statistics will provide an overview of the efficacy and safety results. For categorical parameters, the number and percentage of subjects in each category will be presented. The denominator for percentage will be appropriate for the purpose of the analysis. For continuous parameters, descriptive statistics will include number of subjects, mean, standard deviation (SD), median, minimum, and maximum.

### **11.2 Study Day**

Day 1 is the date of first study drug administration. As subjects apply study drug under supervision on the day of randomization, this will generally be the day of the Baseline Visit. Study day is calculated relative to the date of Baseline (Day 1).

### **11.3 Baseline and Change from Baseline**

Baseline value is defined as the last non-missing value prior to the first dose of study drug. Change from baseline is defined as the post-baseline value minus the baseline value unless otherwise specified. Percent change from baseline is calculated as follows: Percent change = (Change from baseline / Baseline) \* 100.

### **11.4 Analysis Visit Window**

All efficacy and safety endpoints will be analyzed according to the nominal visits (i.e. actual visit) except for assessments collected on early termination and unscheduled visits. Early termination and unscheduled visits will be re-numbered to an analysis visit based on their windowed visits defined by actual study day. If more than one visit occurs within a single visit window, then the analysis will take the one closest to the target day. If the 2 visits are equidistant from the target day, the visit with later date and time will be used.

The following analysis visit windows will apply to early termination and unscheduled visits:

Visit	Week	Target Day	Analysis Window
2	Baseline	1	On or prior to Day 1
3	2	15	Post first dose – Day 21
4	4	29	Day 22 – Day 42
5	8	57	Day 43 – Day 70
6	12	85	Day 71 – Maximum (last day of treatment + 3 days, Day 87)
7	16	113	> Maximum (last day of treatment + 3 days, Day 87)

## 11.5 Study Centers

There will be approximately 50-60 study sites in the United States (US) and Canada.

The primary and secondary efficacy endpoints will be summarized by study site using descriptive statistics. Exploratory analyses of the primary and secondary endpoints may be performed to elucidate any identified sources of heterogeneity in the results.

## 11.6 Multiple Comparisons

Multiple comparisons of the secondary endpoints will be controlled using the Fixed-Sequence method. Hypotheses testing for the secondary efficacy endpoints will only be conducted if the primary efficacy endpoint has demonstrated statistical significance at 0.05 two-sided in favor of tapinarof. Testing of the secondary endpoints will be performed sequentially following the below pre-specified order (see details in Section 12.7.2):

- Proportion of subjects with  $\geq 75\%$  improvement in PASI from Baseline at Week 12
- Proportion of subjects with a PGA score of 0 or 1 at Week 12
- Mean change in %BSA affected from Baseline to Week 12
- Proportion of subjects with  $\geq 90\%$  improvement in PASI score from Baseline to Week 12

Hypothesis testing for secondary endpoints will stop if statistical significance at 0.05 two-sided in favor of tapinarof is not observed.

## 12. STATISTICAL EVALUATION

### 12.1 Subject Disposition

The number and percentage of subjects screened, screen failed along with reason for screen failures, randomized, included in each analysis population, completing 12-Week Double-Blind treatment, completing the study, withdrawing from treatment (together with the reasons for withdrawal), withdrawing from the study (together with the reasons for withdrawal), and subjects

excluded from PP population (together with the reasons for exclusion) will be summarized using frequencies and percentages by treatment group. For subjects who have completed the 12-Week treatment, the number and percentage of those who elect to enroll in the open-label study will be summarized.

Enrollment and disposition will be listed by subject. Screen failures and subjects not randomized will also be listed by subject.

The number of days in the study (date of study completion / discontinuation minus date of Day 1 plus 1) will be summarized descriptively by treatment group.

## **12.2 Protocol Deviations**

Protocol deviations will be listed by subject. Protocol deviation categories (major vs minor) will be summarized by treatment group.

## **12.3 Demographic and Baseline Characteristics**

Demographics and baseline characteristics will be summarized overall and by treatment for the ITT, PP and Safety populations. The following demographic and baseline variables will be included:

- Age
- Gender
- Race
- Ethnicity
- Weight
- Height
- BMI
- Baseline PGA
- Baseline PASI score
- Baseline %BSA Affected

## **12.4 Study Medication Exposure and Compliance**

The following exposure and compliance parameters will be summarized descriptively by treatment group:

- Total number of days exposed, defined as date of last dose of study drug minus date of first dose of study drug plus 1.



- Number of doses administered, calculated from the subject dose diary. If a subject does not return diary or returns the diary with missing dosing record, the number of doses taken during that period is assumed to be 0.
- Grams (g) of study drug administered, total and average per day. Drug administered is calculated as the summation of the difference between dispensed weight and returned weight for all returned tubes. Unreturned tubes will be assumed unused and will be included as 0 gram in amount drug used calculation.
- Percent compliance will be calculated as the number of doses administered over the number of days exposed.
- Subject compliance, defined as  $\geq 80\%$  compliance over the entire duration of study. If the percentage of study medication compliance is unknown, the subject is assumed to be non-compliant with study medication.

## **12.5 Prior and Concomitant Medications**

Prior (within the 34 days before Visit 1 and with stop dates prior to first dose of study drug) and concomitant (ongoing or with stop dates on or after first dose of study drug) medications will be listed by subject for each treatment group. If the medication is ongoing or the stop year is missing, the medication will be considered as received for the entire duration of the study. Medications will be coded using WHO-DD terminology.

To distinguish prior vs concomitant medications, the following rules for stop dates will apply:

- If only year was recorded, and it is before Baseline, it is a prior medication; if year is same or after Baseline, it is assumed to be a concomitant medication.
- If day is missing, but month and year are before Baseline, it is a prior medication; if month and year are the same as Baseline, it is assumed to be a concomitant medication; if month and year are after Baseline, it is a concomitant medication.
- If start date is after Baseline, it is a concomitant medication regardless.

Prior and concomitant medications will be summarized separately by treatment, WHO-DD Anatomical-Therapeutic-Chemical (ATC) classification and preferred term (PT).

## **12.6 Medical History and Concurrent Procedures**

Medical history (including previous and ongoing medical conditions) will be coded using MedDRA and listed by subject. Medical history will be summarized by system organ class (SOC) and PT for each treatment.

Psoriasis History, Cardiovascular Risk Factors and Liver Disease Family History will be summarized and listed by subject separately.

The following will be summarized by treatment group for psoriasis history:

- Duration (year) of disease, defined as year of Screening visit minus year of initial psoriasis diagnosis. Duration of disease will be reported as <5, 5-10, and >10 years for summary purpose.
- Area affected (scalp, pam, fingernails, toenails, soles)

## **12.7 Efficacy Endpoints**

### **12.7.1 Primary Efficacy Endpoint**

The primary efficacy endpoint is the proportion of subjects who achieve a Physician Global Assessment (PGA) treatment response, defined as a PGA score of clear (0) or almost clear (1) with a minimum 2-grade improvement from Baseline at Week 12.

The PGA is a static 5-point assessment of overall disease severity, as determined by the Investigator, ranging from 0 (clear) to 4 (severe).

### **12.7.2 Secondary Efficacy Endpoints**

The secondary endpoints are:

- Proportion of subjects with  $\geq 75\%$  improvement in Psoriasis Area and Severity Index from Baseline (PASI-75) at Week 12
- Proportion of subjects with a PGA score of 0 or 1 at Week 12
- Mean change in percent of total body surface area (%BSA) affected from Baseline to Week 12
- Proportion of subjects with  $\geq 90\%$  improvement in PASI score from Baseline (PASI-90) to Week 12.

The PASI score assesses the severity of psoriasis taking into account the overall severity of erythema (redness), induration (plaque thickness), and scale, and the extent of %BSA affected with psoriasis. The 3 clinical signs are each graded on a 5-point scale (0 to 4) and the %BSA affected is scored on a 7-point scale (0 to 6) for each of the 4 specified body regions (head, upper extremities, trunk, and lower extremities). The individual scores are multiplied by a weighted factor for each body region; the sum of these scores gives the overall PASI score.

### **12.7.3 Exploratory Efficacy Endpoints**

The exploratory endpoints to further characterize the efficacy of tapinarof cream, 1% compared with vehicle control over time are:

- Time (days) to achieving a PGA treatment response, defined as days from date of first dose until achieving a PGA score of 0 or 1 with a minimum 2-grade improvement from Baseline. Subjects who have no PGA response will be censored at the date of their last PGA assessment.
- Proportion of subjects with a PGA score of 0 or 1 at Weeks 2, 4, and 8
- Proportion of subjects who achieve a PGA treatment response at Weeks 2, 4, and 8
- Mean absolute and percent change in PASI score from Baseline to Weeks 2, 4, 8, and 12
- Proportion of subjects with  $\geq 50\%$  improvement in PASI score (PASI-50) from Baseline to Weeks 2, 4, 8, and 12
- Proportion of subjects with  $\geq 75\%$  improvement in PASI score (PASI-75) from Baseline to weeks 2, 4, and 8
- Proportion of subjects with  $\geq 90\%$  improvement in PASI score (PASI-90) from Baseline to weeks 2, 4, and 8
- Mean change in %BSA affected from Baseline to Weeks 2, 4, and 8
- Mean % change in %BSA affected from Baseline to Weeks 2, 4, 8, and 12.

#### 12.7.4 Functional Outcomes and Health-related Quality of Life Endpoints

To describe the effect of tapinarof cream, 1% on functional outcomes and the associated impact on daily activities and attitudes in adults with plaque psoriasis:

- Change in Peak Pruritus-Numeric Rating Scale (Peak Pruritus-NRS) from Baseline at Weeks 2, 4, 6, 8, 10, 12 and 16
- Change in psoriasis impact on daily activities, as measured by the Dermatology Life Quality Index (DLQI) total and individual dimension scores from Baseline at Weeks 4, 12, and 16

**The DLQI** is a 10-item PRO measure that assesses the extent to which the skin condition has affected the subject's quality of life over the past week, including 6 domains (daily activities, personal relationships, symptoms and feelings, leisure, work/school, and treatment). The total score (0-30) will be calculated as the sum of 10 questions with each ranging from 0 to 3. The scoring of each question is as follows:

Very much	Scored 3
A lot	Scored 2
A little	Scored 1
Not at all	Scored 0
Not relevant	Scored 0

Question unanswered	Scored 0
Question 7: “prevented work or studying”	Scored 3

If one question is left unanswered this is scored 0 and the scores are summed and expressed as usual out of a maximum of 30. If two or more questions are left unanswered the total DLQI score is missing. If question 7a is answered ‘yes’ or ‘not relevant’, then the answer of 7b will be ignored.

When summarizing the domain scores, if a domain contains two questions and the answer to one of the questions is missing, that domain will be scored as the answer to the non-missing question. If all the answers to questions in a domain are missing, the domain will not be scored.

- Change in psoriasis symptoms, as measured by the Psoriasis Symptom Diary (PSD), including total and 16 individual scores from Baseline at Weeks 2, 4, 8, 12, and 16.

PSD is a 16-item assessment measures self-reports of psoriasis symptoms, and impact on functional health. The total score will be calculated as the sum of 16 questions with each ranging from 0 to 10. If any of the 16 questions is unanswered, the total score will be considered as missing.

- Change in the eight domain scores, as measured by the RAND Short Form-36 (SF-36) questionnaire from Baseline at Weeks 4, 12, and 16.

The 8 domain scores below will be computed using the scoring rules for the RAND 36-Item Health Survey (Version 1.0):

- Physical functioning
- Role limitations due to physical health
- Role limitations due to emotional problems
- Energy/fatigue
- Emotional well-being
- Social functioning
- Pain
- General health

## 12.8 Efficacy Analyses

All efficacy outcomes will be analyzed in the ITT population. Only the primary and secondary efficacy endpoints will be analyzed in the PP population as supportive analyses, as summarized in the Table below.

Efficacy Endpoints	Primary	Sensitivity Analysis
--------------------	---------	----------------------

	Analysis	
<u>Primary Endpoint</u> <ul style="list-style-type: none"> <li>• PGA response at Week 12</li> </ul>	ITT, MI	ITT, OC ITT, TF ITT, LOCF PP, OC
<u>Secondary Endpoints</u> Categorical Endpoint: <ul style="list-style-type: none"> <li>• PASI-75 at Week 12</li> <li>• PGA 0 or 1 at Week 12</li> <li>• PASI-90 at Week 12</li> </ul> Continuous Endpoint: <ul style="list-style-type: none"> <li>• Change in %BSA at Week 12</li> </ul>	ITT, MI	ITT, OC ITT, TF (categorical endpoint) ITT, LOCF PP, OC
<u>Exploratory Endpoints</u> <ul style="list-style-type: none"> <li>• Time to PGA response</li> </ul>	ITT, OC	N/A
<u>Exploratory Endpoints</u> <ul style="list-style-type: none"> <li>• PGA 0 or 1 at Weeks 2, 4, and 8</li> <li>• PGA response at Weeks 2, 4, and 8</li> <li>• Change in PASI from Baseline at Weeks 2, 4, 8, and 12</li> <li>• % change in PASI from Baseline at Weeks 2, 4, 8, and 12</li> <li>• PASI-50 at Weeks 2, 4, 8, and 12</li> <li>• PASI-75 at Weeks 2, 4, and 8</li> <li>• PASI-90 at Weeks 2, 4, and 8</li> <li>• Change in %BSA from Baseline at Weeks 2, 4, and 8</li> <li>• % change in %BSA from Baseline at Weeks 2, 4, 8, and 12</li> </ul>	ITT, MI	ITT, OC
<u>Other Endpoints</u> <ul style="list-style-type: none"> <li>• Change from Baseline in Peak NRS at Weeks 2, 4, 6, 8, 10, 12 and 16</li> <li>• % change from Baseline in Peak NRS at Weeks 2, 4, 6, 8, 10, 12 and 16</li> <li>• Peak NRS 4-point Responder at Weeks 2, 4, 6, 8, 10, 12 and 16</li> <li>• Change from Baseline in DLQI at Weeks 4, 12, and 16</li> <li>• % change from Baseline in DLQI Total Score at Weeks 4, 12 and 16</li> <li>• Change from Baseline in PSD at Weeks 2, 4, 8, 12, and 16</li> <li>• % change from Baseline in PSD Total Score at Weeks 2, 4, 8, 12 and 16</li> <li>• Change from Baseline in SF36 at Weeks 4, 12, and 16</li> </ul>	ITT, OC	N/A

### 12.8.1 Primary Efficacy Analyses

The primary efficacy endpoint will be analyzed using a Cochran–Mantel–Haenszel (CMH) test stratified by PGA score at Baseline (PGA scores of 2, 3 or 4) based on the ITT population using the multiple imputation approach.

- For each multiple imputation dataset, the CMH option in PROC FREQ will calculate the adjusted estimate of the common relative risk of success (tapinarof/placebo) over all strata with its 95% confidence interval (CI).
- The logit-based version of the estimate and its confidence bounds will be log-transformed and the SE of the estimate of the log(relative risk) will be back-calculated from the log(confidence bounds) by assuming an asymptotic normal distribution.
- The estimates of the log(relative risk) and their standard errors for all the multiple imputation datasets will be input to PROC MIANALYZE to test the null hypothesis that tapinarof is not different from placebo. The p-value of the combined results will be reported for interpretation.
- PROC MIANALYZE will output an overall estimate of the log(relative risk) and its 95% CI. These will be exponentiated to produce an overall estimate of the relative risk and its 95% CI.

Sensitivity analyses in the ITT population will be conducted in which missing values are not imputed, i.e., Observed Cases (OC), or imputed as treatment failures (TF), and efficacy is tested with the CMH statistic as described above. Supporting analyses of the primary endpoint will also be performed

- in the PP population based on OC only,
- in the ITT population with missing values imputed by LOCF,

### **12.8.2 Secondary and Exploratory Efficacy Analyses**

The same methods as discussed for the primary analyses will be used to analyze all dichotomized and continuous secondary endpoints. The secondary efficacy endpoints will be tested sequentially in the order listed. Testing will stop if non-significance (2-sided  $p \geq 0.05$ ) is observed as described in Section 11.6.

Other exploratory efficacy endpoints will be analyzed using CMH test for proportions, and analysis of covariance (ANCOVA) model for continuous variables. CMH tests will be stratified by baseline PGA score, and ANCOVA models will include baseline PGA score as a covariate (2, 3 or 4), and baseline value of the endpoint as a continuous covariate. For proportions, the same MI approach describe above will be taken to deal with missing data. For continuous variables, the standard MI procedures for ANCOVA will be followed. Least squared means of treatment difference (tapinarof – placebo) along with the 95% CI will be reported.

Time to achieving a PGA success will be analyzed with the Kaplan Meier method in the ITT population based on Observed Cases. Treatment groups will be compared using a log-rank test. Estimates of median time to response from Kaplan Meier distribution will be computed. A Kaplan-Meier figure will also be generated.

### **12.8.3 Other Efficacy Analyses**

Change from baseline in DLQI, PSD, SF-36, and Peak Pruritus-NRS will be analyzed using an ANCOVA model with treatment as a main effect, baseline PGA score as a categorical covariate, and baseline value of the endpoint as a covariate. Percent change from baseline in DLQI, PSD and Peak Pruritus-NRS will be summarized and analyzed using the same method.

In addition, the proportion of subjects with at least a 4-point reduction from baseline in the Peak Pruritus-NRS among subjects with a baseline NRS score  $\geq 4$  will be summarized and analyzed following the same method as the dichotomized secondary endpoints. A plot will be used to illustrate this information.

Subject itch diary and in-clinic itch score assessments will be listed and summarized using descriptive statistics by day. If a subject completed diary itch score on the date of in-clinic visit, the itch score assessed in-clinic will be used in the summary. Change from baseline, treatment difference (Tapinarof cream 1% – Vehicle Cream) in mean change from baseline values and its associated 95% CI will be summarized by day. In addition, the treatment difference and its associated 95% CI will be plotted by day.

### **12.8.4 Subgroup Analyses**

The primary and all secondary efficacy endpoints will be summarized descriptively in each baseline PGA subgroup, study center, age (<65,  $\geq 65$ ), sex, race (White, Other), duration of disease (<5 Years, 5-10 Years, >10 Years), %BSA affected (<10%,  $\geq 10\%$ ) and country (USA and Canada). The %BSA Affected continuous outcome will be further explored with summaries of change and % change in %BSA Affected at Week 12 by body region. The implications of any significant heterogeneity on the assessment of overall efficacy will be explored.

## **12.9 Safety Analysis**

All safety analyses will be conducted in the Safety Population. Data will be listed by subject and treatment and summarized by treatment. No formal statistical comparisons will be made for safety data.

Single 12-lead ECGs, Hospital Anxiety and Depression Scale (HADS) collected at Screening visit will be presented in a by-subject listing.

### **12.9.1 Adverse Events**

AE terms will be coded using the MedDRA dictionary. A treatment-emergent AE (TEAE) is defined as an AE that starts after the first dose of study medication. For subjects not rolling over to Open-label study, all reported AEs will be included in the Double-Blind phase summaries. For subjects rolling over to the Open-label study, all AEs with an onset date on or before the Week 12 visit date will be included in the Double-Blind phase summaries.

If relationship to treatment is missing, the event will be conservatively treated as related to study drug. Missing severity will be imputed to the highest severity level (most severe).

All AEs will be listed by treatment and by subject, detailing the verbatim term given by the investigator, the preferred term (PT), system organ class (SOC), onset date and time, end date and time, duration (days), CTCAE grade, outcome, relationship to study drug/treatment, action taken with study drug/treatment, other action taken to treat the event, seriousness and criteria for seriousness. Serious AEs (SAEs), TEAEs leading to study discontinuation, TEAEs related to study drug will also be listed separately. The following adverse event of special interest (AESIs) will be identified and listed separately: contact dermatitis, folliculitis, and headache.

AEs will be summarized by treatment and overall as incidence rates of:

- Any AE
- Any TEAE
- Any AESIs
- Any treatment-related TEAE, including definitely, probably, and possibly related
- Any TEAE leading to study drug discontinuation
- Any SAE (non-Fatal)
- Any serious TEAE
- Death
- Treatment-related Serious TEAE
- Serious TEAE leading to study drug discontinuation

Individual PTs will be summarized by treatment according to the following:

- All TEAEs by SOC in alphabetical order and PT in descending order of combined frequency (also for treatment-related TEAEs, TEAEs leading to study discontinuation and serious TEAEs)
- All TEAEs by SOC, PT, and maximum CTCAE grade
- All TEAEs by SOC, PT, and maximum causality (not related, related) to the study drug

In addition, TEAEs and serious TEAEs will be summarized by SOC (alphabetical order) and PT (descending order) for the following subgroups:



- Age (<65 and >= 65 years old)
- Sex
- Race (White, Other)

At each level of summarization, a subject will be counted once if he/she reported one or more events. The severity of TEAEs and relationship to study drug will be summarized in a similar manner. For summaries of relationship to study drug, a subject will be classified according to the closest relationship. For summaries of TEAE CTCAE grade, a subject will be classified according to the worst grade.

For AESIs, summarization will be more extensive, reflecting the more detailed information collected. Information summarized will include number of events per subject, earliest onset day, duration (in days), causality, grade and seriousness of AESIs, outcome, actions taken with study drug, assorted physical characteristics of the AESIs, and demographic/baseline characteristics and PGA status of the subjects experiencing them. Each type of AESI will be summarized separately. If a subject has more than one treatment-emergent occurrence of an AESI, the subject's maximum duration, highest levels of causality and seriousness, maximum grade and generally the most extreme level of each characteristic will be summarized. If an AESI was ongoing at EOS, it will not be included in the duration summary.

A Kaplan-Meier figure will be generated for each AESI of time to first event. Subjects not experiencing the AESI will be censored at the date of study completion or discontinuation.

### **12.9.2 Local Tolerability Scale (LTS)**

LTS scores will be summarized by treatment and visit for subject overall assessment (scores from 0=None to 4=Strong/Severe) and Investigator overall assessment (scores from 0=No irritation to 4=Very Severe) separately. LTS scores at the sensitive areas will be summarized by visit, area and treatment group for Investigator overall assessment, as well as listed by subject and anatomical site.

### **12.9.3 Clinical Laboratory Testing**

All clinical laboratory (hematology, clinical chemistry, urinalysis) values will be listed by subject. The reference normal ranges and reference range indicators (e.g. high, low, normal etc.) will be supplied by the central laboratory and displayed in the data listings. For the purpose of shift table summaries, the central lab reported reference range indicators will be classified as low, normal and high accordingly based on the table below. Change from baseline in abnormality status will be summarized using shift tables. For urinalysis, only lab tests with numeric results reported are included in shift tables. Separate listings will identify subjects with markedly abnormal values. For quantitative measures, observed values and changes in clinical laboratory values will be summarized descriptively by visit and treatment group.

Reference Range Indicator in Shift Tables	Reference Range Indicator Reported by Central Lab
Low	Low, Low Level 1 Alert, Low Level 2 Alert
Normal	Normal
High	High, High Level 1 Alert, High Level 2 Alert

### 12.9.4 Vital Signs

Observed values and changes in vital sign parameters (systolic and diastolic blood pressure [SBP, DBP], pulse rate, and body temperature) will be summarized descriptively by visit and treatment group. Vital sign values will be classified as normal, low, high, based on reference ranges as per Table below. Subjects with markedly abnormal changes will be listed and tabulated separately.

	Absolute Values			Change (Absolute) from Baseline	
	Low	Normal	High	Abnormal Change	Markedly Abnormal Change
SBP	<90 mmHg	90-140 mmHg	>140 mmHg	≥ 20 mmHg	≥ 40 mmHg
DBP	<50 mmHg	50-90 mmHg	>90 mmHg	≥ 10 mmHg	≥ 20 mmHg
Pulse	<50 bpm	50-100 bpm	>100 bpm	≥ 10 bpm	≥ 30 bpm

### 12.9.5 Physical Examination

Brief physical examination results including weight, height, and BMI will be listed.

### 12.10 Pharmacokinetic Analyses

Plasma concentration data at Weeks 4 and 12 will be listed and summarized for the tapinarof cream, 1% group. Listings will be sorted by treatment, subject, day, and time.

The plasma concentrations divided by baseline total %BSA affected in cm<sup>2</sup> will be summarized descriptively as well. For this calculation, 1% BSA is equal to 185 cm<sup>2</sup>.

Descriptive summaries will include n, mean, SD, median, minimum, and maximum. Concentration values below the limit of quantification will be treated as 0 in the summary.

In addition, plasma concentration at Weeks 4 and 12 will be summarized for the following subgroups:

- Age (<65 and ≥ 65 years old)
- Sex
- Race (White, Other)

Exploratory analyses may include relationships of plasma concentrations with patient demographics or AEs.

### **12.11 Additional Analyses - COVID-19**

In order to describe the impact of COVID-19 on current study, the following disposition events will be summarized in the tables separately:

- Subjects discontinued from the treatment/study as a result of a positive COVID-19 diagnosis.
- Subjects discontinued from the treatment/study due to other reasons related to COVID-19. This is excluding COVID-19 diagnosis but may include reasons such as site closure, travel restrictions, fear of infection, etc.
- Subjects with study visits altered (including modified in-clinic visit, virtual and phone visits) and missed due to COVID-19

COVID-19 related protocol deviations will be summarized separately. The impact of COVID-19 (including protocol deviation, visit alteration, treatment/study discontinuation and diagnosis of COVID-19) will also be flagged at subject-level in a data listing. Subject profile will be used to compile all COVID-19 related information for affected subjects.

An additional sensitivity analysis of the primary endpoint will be added in the ITT population based on OC only and excluding any virtual PGA assessments at Week 12.

Also, all COVID-19 related symptoms and confirmed cases occur during the study will be reported as AEs and included in the summaries.

## **13. CHANGES FROM THE PROTOCOL AND PLANNED ANALYSES**

The protocol specified change over time in physical component score and mental component score as measured by the SF-36 questionnaire as an exploratory endpoint. Scoring rules for the RAND 36-Item Health Survey (Version 1.0) are provided for the eight domain scores but not for the physical component score or the mental component score. Hence, summaries of the eight domain scores are planned.

No other changes from the protocol in planned analyses are noted. Any additional statistical analyses to handle modifications to study conduct and/or missing data due to the COVID-19 pandemic have been documented in Section 12.11 above in this SAP and will be finalized prior to database lock and unblinding of the study.

## **14. FORMATTING**

Each page of the analysis will show the sponsor's name, and the protocol number. Report tables will be embedded in the MS Word report document from SAS program output without change.

The footer of each table will show the name of the SAS program module which generated it and the date and time the table was generated.

The SAS programs will generate rich-text-formatted (RTF) output with the “RTF” extension using the SAS Output Delivery System (ODS). The summary tables and listings will be formatted using the Courier New 9-point font. The RTF output is included in report documents prepared with Microsoft Word and converted to PDF format without typographical change.

Datasets will be created and taken as input to validated SAS programs to generate the report-ready tables, listings, and figures. Each output display will show the names of the data sets and SAS program used to produce it.

## **15. ARCHIVING AND RETENTION OF DOCUMENTS**

After finalization of the analysis, the following will be archived at IQVIA Biotech and/or with the study sponsor:

- SAP and any amendments
- All SAS code used in the project for statistical analysis, report tables generation, and analysis data set creation
- Tables, listings and figures as included in the clinical study report
- SAS study data tabulation model (SDTM) and analysis dataset model (ADaM) datasets