MC2 Therapeutics

Statistical Analysis Plan
for
Clinical Study Reporting

of the Phase III Clinical Study MC2-01-C7

A Randomized, Multicentre, Investigator-Blind, Parallel-Group Trial to Evaluate the Efficacy and Safety of MC2-01 Cream Compared to Vehicle and Active Comparator in Subjects with Mild-to-Moderate Psoriasis Vulgaris

FINAL Version 1.0 as of 7 October 2019

Author: Frank Freischläger, Senior Principal Statistical Writer, Estimondo GmbH

Contributor: Ladislav Pecen, Senior Principal Statistical Scientist, Estimondo GmbH

For review and approval, see the MC2-01-C7 SAP for CSR Version 1.0 Review and Approval Sheet on file.

Confidential

Document History

Version	Effective Date	Description of Changes
1.0	7 Oct 2019	New

Table of Contents

List of Tables

List of Abbreviations and Acronyms

Abbreviation	Term
ADaM	Analysis Data Model
AE	Adverse event
ANCOVA	Analysis of covariance
ATC	Anatomic therapeutic chemical
BSA	Body surface area
CDISC	Clinical Data Interchange Standards Consortium
CRF	Case report form
CSR	Clinical study report
CTP	Clinical trial protocol
DLQI	Dermatology Life Quality Index
EQ-5D	EuroQOL five dimensions
FAS	Full analysis set
ICH	International Council on Harmonization
IP	Investigational product
LOCF	Last observation carried forward
LSR	Local skin reaction
MCMC	Markov Chain Monte Carlo
MedDRA	Medical dictionary for regulatory activities
MI	Multiple imputation
mPASI	Modified Psoriasis Area and Severity Index
NRI	Non-responder imputation
PGA	Physician's Global Assessment
PPS	Per-protocol set
PT	Preferred term
PTCS	Psoriasis Treatment Convenience Scale
QOL	Quality of life
SAE	Serious adverse event
SAP	Statistical analysis plan
SAS®	Statistical Analysis System®
SAS/STAT®	SAS® Statistics software module
SD	Standard deviation
SDTM	Standard Data Tabulation Model
SGA	Subject global assessment of disease severity
SOC	System Organ Class
TFL	Tables, figures, listings
TPA	Tipping-point analysis
VAS	Visual analogue scale

1 Introduction

The Phase III trial MC2-01-C7 will be analyzed and reported in accordance with the clinical trial protocol (CTP) [1] and the International Conference on Harmonization (ICH) guidelines ICH E3 [2], ICH E6 [3] and ICH E9 [4]. MC2-01-C7 is a trial to compare the MC2 product in psoriasis vulgaris with a vehicle and an active comparator focusing on psoriasis on the body (trunk and/or limbs).

This statistical analysis plan (SAP) should be read in conjunction with the CTP and the electronic case report form (CRF) [5]. This version of the plan has been developed using the CTP version 2.0 dated 16 October 2018 and the SDTM-annotated CRF dated 15 August 2019. Any further changes to the protocol or CRF may necessitate updates to the SAP.

Based on this SAP, deliverables are derived analysis datasets following the CDISC ADaM standard, tables, figures and listings (TFL).

2 Study Design and Objectives

2.1 Study Design

This is a prospective, investigational, investigator-blind, vehicle and active-controlled, parallel-group, 3-arm, randomized, phase-III trial conducted at approximately 34 sites in the European Union. The trial population consists of subjects fulfilling all inclusion and no exclusion criteria as defined in the CTP, Sections 5.2 and 5.3 with any exceptions documented [1]. Random allocation to one of three parallel treatment groups in a 3:1:3 ratio (MC2-01 cream; MC2-01 vehicle; active comparator) is stratified by site.

A total of 476 subjects are planned to be enrolled, randomized and observed until final evaluation and completion. Subjects prematurely terminating the trial will not be replaced.

The maximum trial duration for each subject will be approximately 14 weeks and includes a screening period of up to 30 days (if washout of prohibited medications is required), an 8-week treatment period, and a follow-up period of 2 weeks.

A total of seven visits and a telephone interview of subjects are planned. The telephone interview is planned for Week 2 (Visit 3), while the visits will be held at screening (Visit 0, Day -30 to 0), randomization (Visit 1, Week 0, Day 0), Visits 2, 4, and 5 at Weeks 1, 4 and 6, respectively, planned for Days 7, 28, and 42 plus/minus 2 days, respectively, Visit 6 as end of treatment Visit (Week 8, Day 56+/-2, or at early termination), and Visit 7 as follow-up visit (Week 10, Day 70+/-2). The follow-up visit is planned as an on-site visit in case of abnormal laboratory values or unresolved adverse events (AE). Otherwise, the follow-up visit is a telephone interview. At any time during the trial, if judged necessary by the investigator, unscheduled visits may be performed. For a visit schedule, see CTP Section 7.0 [1].

2.2 Treatments, Study Groups and Treatment Assignment

Subjects will be randomly allocated to either receive *MC2-01 cream*, a combination product of calcipotriol and betamethasone, a *MC2-01 vehicle* or an *active comparator*, a marketed calcipotriol and betamethasone gel/topical suspension. Subjects apply the investigational product (IP) topically once daily preferable in the evening for 8 weeks to affected areas on the trunk (including the neck), the limbs, i.e., arms (including the back of the hands), the legs (including the buttocks and the top of the feet) and the scalp.

Subjects should apply enough IP to treat the entire affected areas. Up to 4 tubes/bottles of 60 gram will be dispensed for a treatment period of 2 weeks including the allowed visit window of 2 days. The weekly dose is not to exceed 100 g, and the treated area should not exceed 30% of the body surface area (BSA).

Subjects classified as *clear* at any of the on-treatment visits may stop the IP treatment at the investigator's discretion. They should remain in the trial and attend all visits up to and including the follow-up visit. The IPs will continue to be dispensed to the subject, and IP treatment may be restarted at the subject's discretion. The subjects should not discontinue IP treatment themselves between visits but are only allowed to stop using the IP treatment on the advice of the investigator at a scheduled visit.

Subjects will be uniquely identified by a five-digit subject number composed of a two-digit site number, followed by three digits. The investigators and the entire study team will be kept blinded to treatment allocation information until the time of unblinding. Due to difference in formulation and packaging, it is not possible to double blind the IPs. A designated non-blinded staff member will be responsible for handling the IP (e.g. dispensing, returning, drug accountability). The investigators and the rest of the blinded study team will be kept blinded to treatment formulation information until the time of unblinding.

2.3 Study Objectives

2.3.1 Primary Objectives

The primary objective is to evaluate the efficacy of MC2-01 cream compared to active comparator in subjects with psoriasis vulgaris.

2.3.2 Secondary Objectives

The secondary objective is to characterise the safety profile of MC2-01 cream in subjects with psoriasis vulgaris.

2.4 Study Endpoints

In the following, endpoints are described by label, and by unit or valid categories. A categorical endpoint is of yes-no-type, if no categories are specified. A continuous or numerical endpoint is a score with or without a unit, if no unit is specified.

2.4.1 Efficacy Including Subject-Reported Outcomes

The primary endpoint is the percentage change from baseline in the modified Psoriasis Area and Severity Index (mPASI) on the body (trunk and/or limbs) at Week 8.

Based on a physician's global assessment (PGA) and the Psoriasis Treatment Convenience Scale (PTCS), secondary endpoints include

- *PGA success* on the body at Week 8, defined as a minimum 2-point decrease from baseline to Week 8 on the PGA of psoriasis severity on the body (trunk and/or limbs) to a PGA of psoriasis severity of 0 (clear) or 1 (almost clear); i.e., a score of 0 (clear) or 1 (almost clear) for subjects with moderate disease at baseline; or a score of 0 (clear) for subjects with mild disease at baseline.
- Subject assessment of treatment convenience at Week 8 using the PTCS total score,
- Number of subjects with at least 75% reduction in mPASI from baseline (mPASI 75) on the body (trunk and/or limbs) at Week 8.
- *PGA success* on the scalp at Week 8, defined as a minimum 2-point decrease from baseline to Week 8 on the PGA of psoriasis severity on the scalp to a PGA of psoriasis severity of 0 (clear) or 1 (almost clear); i.e., a score of 0 (clear) or 1 (almost clear) for subjects with severe or moderate disease at baseline; or a score of 0 (clear) for subjects with mild disease at baseline.

Primary and secondary endpoints will be evaluated confirmatorily by hierarchical testing as described in Sections 5.2 and 5.3 below. Other endpoints will be exploratorily analyzed.

Other endpoints include

- Percentage change from baseline in mPASI score on the body (trunk and/or limbs) at Week 4,
- mPASI 75 on the body (trunk and/or limbs) at Week 4,
- PGA success on the body at Week 4,
- PGA success on the scalp at Week 4,
- Change from baseline in Subject Global Assessment (SGA) at Week 4 and Week 8,
- Subject assessment of treatment convenience at Week 4 using the PTCS total score,
- Change from baseline in the Dermatology Life Quality Index (DLQI) score at Week 4 and Week 8,
- Change from baseline in EQ-VAS at Week 4 and Week 8,
- BSA involvement on the body and on the scalp at Week 4 and Week 8.

mPASI and mPASI 75

At all on-site visits (Screening, Week 0, Week 1, Week 4, Week 6, Week 8, and Early Termination), the investigator will assess the extent and severity of the subjects' psoriasis on the body (trunk and/or limbs) using an mPASI scoring system for each of the 3 areas (arms, trunk, and legs). The extent (E) is scored 0 to 6. The severity of the psoriatic lesions in each of the 3 areas will be recorded for each of the signs of redness (R), thickness (T), and scaliness (S). For each clinical sign, a single score, reflecting the average severity of all psoriatic lesions on given body region, will be determined. Redness, thickness and scaliness scores are ranging from 0 to 4. The meaning of scores is summarized in the following table.

Table 1: Overview of the meaning of mPASI scores assessed for arms, trunk and legs

Score	Extent (E)	Redness (R)	Thickness (T)	Scaliness (S)
0	no involvement	none (no erythema)	none (no plaque elevation)	none (no scaling)
1	<10%	mild (faint erythema, pink to very light red)	mild (slight, barely perceptible elevation)	mild (sparse, fine-scale lesions, only partially covered)
2	10% to 29%	moderate (definite light red erythema)	moderate (definite elevation but not thick)	moderate (coarser scales, most of lesions covered)
3	30% to 49%	severe (dark red erythema)	severe (definite elevation, thick plaque with sharp edge)	severe (entire lesion covered with coarse scales)
4	50% to 69%	very severe (very dark red erythema)	very severe (very thick plaque with sharp edge)	very severe (very thick coarse scales, possibly fissured)
5	70% to 89%	-	-	-
6	90% to 100%	-	-	-

Per area, the scores for extent, redness, thickness and scaliness are combined to $A = E \cdot (R+T+S)$. Ranging from a minimum score of 0 and a maximum score of 64.8, the mPASI is calculated from the scores of the extent and severity of the disease by use of the following equation:

$$mPASI = 0.2 \cdot A_{arms} + 0.3 \cdot A_{trunk} + 0.4 \cdot A_{legs}.$$

At any post-baseline visit, mPASI 75 is defined as an indicator of a reduction in mPASI from baseline of 75% or more.

PGA and PGA Success on the Body and on the Scalp

PGA measures the investigator's impression of the disease at a single point using a defined, 5-point, static PGA scale (0 = clear, 1 = almost clear, 2 = mild, 3 = moderate or 4 = severe) at Screening, Weeks 0, 1, 4, 6, 8, and Early Termination.

The PGA assessment will be made separately for the body (trunk and/or limbs) and the scalp. It will represent the *average* lesion severity on the two body locations.

At any post-baseline visit, *PGA success* is defined as a minimum 2-point decrease from baseline to the respective visit on the PGA of psoriasis severity to 0 (clear) or 1 (almost clear).

PTCS

PTCS will be completed by the subjects before any other assessments are performed at Weeks 1, 4, 8, and Early Termination.

The aim of PTCS is to assess the impact and convenience of psoriasis treatment. The scale consists of 6 disease-specific, self-report questions rated on a 10-point scale ranging from 1 to 10. The questions are

1. How easy was the treatment to apply to the skin? Answered by 1 = very difficult to 10 = very easy.
2. How greasy was the treatment when applying it to the skin? Answered by 1 = very greasy to 10 = not greasy.
3. How moisturised did your skin feel after applying the treatment? Answered by 1 = not moisturized to 10 = very moisturized.
4. How greasy did your skin feel after applying the treatment? Answered by 1 = very greasy to 10 = not greasy.
5. How much did treating your skin disrupt your daily routine? Answered by 1 = very disturbing to 10 = not disturbing.
6. Overall, how satisfied were you with the medical treatment? Answered by 1 = not satisfied to 10 = very satisfied.

Ranging from 5 to 50, a PTCS total score is the sum of the scores on questions 1 to 5. If no more than two questions are answered, the PTCS total score is missing. If one or two questions remain unanswered, the missing scores are replaced by the average of the answered scores for the summation.

SGA

Subjects will grade the overall severity of their symptoms according to the following 5-point scale at Weeks 0, 1, 4, 6, 8, and Early Termination:

0 = Clear; no psoriasis symptoms at all.
1 = Very mild; very slight psoriasis symptoms that do not interfere with daily life.
2 = Mild; slight psoriasis symptoms that interfere with daily life only occasionally.
3 = Moderate; definite psoriasis symptoms that interfere with daily life frequently.
4 = Severe; intense psoriasis symptoms that interfere or restrict daily life very frequently.

DLQI

The DLQI is a validated questionnaire consisting of 10 questions relating to the degree to which the subject's skin condition affected their daily activities, assessed at Weeks 0, 1, 4, 8, and Early Termination [6]. The questions are

1. Over the last week, how itchy, sore, painful or stinging has your skin been? Answered by 0 = not at all to 3 = very much.
2. Over the last week, how embarrassed or self-conscious have you been because of your skin? Answered by 0 = not at all to 3 = very much.
3. Over the last week, how much has your skin interfered with you going shopping or looking after your home or garden? Answered by 0 = not at all/not relevant to 3 = very much.
4. Over the last week, how much has your skin affected the clothes you wear? Answered by 0 = not at all/not relevant to 3 = very much.
5. Over the last week, how much has your skin affected any social or leisure activities? Answered by 0 = not at all/not relevant to 3 = very much.
6. Over the last week, how much has your skin made it difficult for you to do any sport? Answered by 0 = not at all/not relevant to 3 = very much.

7. Over the last week, has your skin prevented you from working or studying? Answered by no, 0 = not relevant or 3 = yes. If no, over the last week, how much has your skin been a problem at work or studying? Answered, by 2 = a lot, 1 = a little, 0 = not at all.
8. Over the last week, how much has your skin created problems with your partner or any of your close friends or relatives? Answered by 0 = not at all/not relevant to 3 = very much.
9. Over the last week, how much has your skin caused any sexual difficulties? Answered by 0 = not at all/not relevant to 3 = very much.
10. Over the last week, how much of a problem has the treatment for your skin been, for example by making your home messy, or by taking up time? Answered by 0 = not at all/not relevant to 3 = very much.

DLQI domain scores are defined as the sum of the scores on questions 1 and 2 as symptoms and feelings domain, questions 3 and 4 as daily activities domain, questions 5 and 6 as leisure domain, and questions 8 and 9 as personal relationships domain, respectively. Work and school domain is covered by question 7, treatment domain by question 10.

Ranging from 0 to 30, the DLQI score is the sum of the 10 scores. If one question remains unanswered, the missing score is replaced by 0 for the summation. If more than one question remains unanswered, the DLQI score is missing.

EQ-5D

The EQ-5D is intended to measure health-related quality of life (QOL) [7]. It consists of 5 items and a visual analogue scale (EQ-VAS). The EQ-VAS records the subject's perceptions of their own current overall health scoring in a range from 0 to 100. The 5 self-assessment questions describe the subject's current health in 5 dimensions i.e., mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. The items follow a 5-point Likert scale.

BSA Involvement

BSA involvement will be determined on the body, on the scalp and total as the extent (percentage) of subject's psoriatic involvement on the total BSA at Week 0 (Day 0, Baseline), Week 4, Week 8, and at Early Termination.

2.4.2 Safety and Tolerability

Local Skin Reaction

Assessments of local skin reaction (LSR) will be performed by the investigator at all on-site post-screening visits. The treatment area and/or immediate surrounding for the following identified signs will be assessed:

- Perilesional erythema, scaling, edema, atrophy, vesicles and erosion/ulceration;
- Lesional vesicles, and erosion/ulceration.

The intensity of each local skin reaction category is to be graded according to the scale in the following table. The most severe intensity observed for each category of the local skin reaction assessment is recorded.

Table 2: Overview of intensity grading of local skin reaction by area and sign

Category		Intensity			
Area	**Sign**	**0 = absent**	**1 = mild**	**2 = moderate**	**3 = severe**
Lesional	Erosion/ ulceration	none	barely visible erosion	distinct erosion	Ulceration
	Vesicles	none	barely visible vesicles	distinct vesicles	Bullae
Perilesional	Erythema	none	barely visible erythema	distinct erythema	dark red erythema
	Scaling	none	barely visible scaling	distinct scaling	gross scales
	Edema	none	barely visible swelling	easily palpable swelling	gross swelling
	Atrophy	none	barely visible thinning	distinct thinning	Striae
	Vesicles	none	barely visible vesicles	distinct vesicles	Bullae
	Erosion/ ulceration	none	barely visible erosion	distinct erosion	Ulceration

An LSR sum score is defined as the sum of intensity grades across all areas and signs per visit. Per sign, the most intense reaction is defined as the maximum intensity grade across visits.

Adverse Events

All adverse events (AE) are recorded with the following attributes:

- severity (mild, moderate, severe),
- location of AE to treatment area,
- action taken (none, concomitant therapy, concomitant non-drug therapy),
- action taken with IP (dose not changed, dose reduced, dose increased, dose interrupted, drug withdrawn, not applicable, unknown, other)
- outcome (recovered, resolved with sequelae, recovering/ resolving, ongoing, death),
- withdrawal due to AE,
- start and end dates,
- relationship to IP (definitely related, probably related, possibly related, not related),
- relationship to study procedure,
- seriousness, criterion for seriousness (death, life-threatening, hospitalization, disability/incapacity, congenital anomaly/birth defect, other).

Any AE is treatment-emergent, if fulfilling the definition in Section 4.4.5.

Any AE is a drug-related AE, if relationship to IP is not assessed as *not related*.

Any AE is leading to drug withdrawal, if action taken with IP is *drug withdrawn*.

Any AE is located in the treatment area, if location of AE to treatment area is *inside treatment area* or *combined*.

Verbatim terms of the description of AEs will be coded according to MedDRA version 22.0 and classified according to preferred terms (PT) and primary system organ classes (SOC).

Laboratory Parameters

For all subjects, the following tests will be performed at Week 0 (Day 0, Baseline), Week 4 and Week 8, and at Early Termination:

- Serum biochemistry: serum calcium (albumin corrected), serum albumin, serum alkaline phosphatase, serum phosphate, plasma parathyroid hormone, 25-OH Vitamin D (only at Week 0).
- Urinalysis: urinary calcium, urinary phosphate, urinary calcium:creatinine ratio.

Laboratory parameters are reported in or converted to SI units [8]. According to normal ranges, laboratory values are categorized as *abnormal low*, *within normal range*, or *abnormal high*.

During the course of the trial, urine pregnancy tests will be performed at Screening, Week 0, Week 4, Week 8, and at Early Termination.

Vital Signs

Systolic and diastolic blood pressure, pulse rate and body temperature (oral or ear) will be measured at Week 0 (Day 0, Baseline), Week 4, Week 8, and at Early Termination. Values outside notable ranges will be flagged. For a definition of notable ranges, see Section 4.6.4 below.

Physical Examination

Abbreviated physical examination and complete dermatological examination will be assessed at Screening (Visit 0), Week 0 (Day 0, Baseline), Week 4, Week 8 and at Early Termination. In addition, height and weight are measured at Week 0 (baseline).

2.4.3 Extent of Exposure and Study Treatment Compliance

Records of IP used, and dosages administered will be kept during the trial. The *duration of study treatment* is derived as the time in days from the date of first study drug application to the date of last study drug application. The *extent of exposure* in days is determined as the duration of study treatment minus the number of days without dosing. The number of days without dosing is derived from dose documentation between the date of first study drug application and the date of last study drug application. Days without dosing are days with reported non-administration.

Subjects will be asked to return all used and unused tubes/bottles at each visit. All returned tubes/bottles that had been dispensed to a subject and are returned with a broken seal will be weighed to determine the amount of the IP used.

The *amount of drug used* in grams is derived as the sum of the difference in drug weight dispensed and returned per kit. The total amount of drug used will be evaluated and also the amount used during the first week, the three weeks from Week 1 to Week 4, and the four weeks from Week 4 to Week 8.

Duration of study treatment, extent of exposure, and amount of drug used will be presented by summary tables by treatment group.

Treatment Compliance

At all on-treatment visits, the subject will be asked if he/she has used the medication as prescribed. If this is not the case, the degree and nature of noncompliance will be collected in the CRF. In addition, subjects will be asked to complete a dosing diary during the treatment period as a measure of treatment compliance.

On a daily basis in the dosing diary, if non-administration due to approved discontinuation is reported, treatment compliance is concluded, and if missed administration is reported, treatment non-compliance is concluded.

For the entire treatment period to the Week 8 visit and for the periods from Week 0 to th Week 4 visit and from the Week 4 visit to the Week 8 visit, the percentage of *treatment compliance* is defined as 100 times the duration of the respective period minus the number of days with non-compliance divided by the duration of the respective period. Treatment compliance will be categorized to less than 80% versus 80% or more and to less than 70% versus 70% or more.

For the two periods, treatment compliance, as percentage and in categories, and non-administrations will be presented by summary tables by treatment group.

2.4.4 Concomitant Therapies and Procedures

All therapies, including non-drug therapies, and procedures taken within 30 days prior to the start of and/or during the trial will be recorded at Screening and throughout the course of the trial. Information regarding the total daily dose, route of administration, start and discontinuation dates, location and indication are recorded.

All verbatim terms of drug and non-drug therapies will be encoded according to the WHO Drug Dictionary version Q4/2018 and classified by PT and primary Anatomic Therapeutic Chemical (ATC) class level 2. All verbatim terms of procedures will be encoded according to MedDRA version 22.0.

A therapy or procedure is *concomitant*, if not ending before or at Week 0 (Day 0, Visit 1) and not starting after the date of last study drug application. A therapy or procedure is post-treatment, if starting after the date of last study drug application.

Per ATC level 2 and PT, frequencies of subjects with at least one concomitant therapy and at least one post-treatment therapy will be presented by summary tables by treatment group and overall, respectively.

2.5 Baseline Characteristics

2.5.1 Baseline Disease Characteristics

All patients suffer from mild to moderate psoriasis vulgaris on the body.
At screening and at randomization, the trial population is described by

- time in years since diagnosis, defined as the number of years from diagnosis to randomization at Week 0 (Visit 1),
- Fitzpatrick Skin Type class at Screening (Visit 0), classified as I = pale white skin, blue/hazel eyes, blond/red hair, skin always burns and does not tan; II = fair skin, blue eyes, skin burns easily and tans poorly; III = darker white skin that tans after initial burn; IV = light brown skin that burns minimally and tans easily; V = brown skin that rarely burns and tans darkly easily; or VI = dark brown or black skin that never burns and always tans darkly,
- previous psoriasis treatment per type of treatment,
- BSA involvement on the body and on the scalp at Week 0 (Visit 1),
- mPASI score at Week 0 (Visit 1),
- PGA on psoriasis severity on the body and the scalp at Week 0 (Visit 1),
- SGA on disease severity at Week 0 (Visit 1).

Baseline disease characteristics will be presented by summary tables by treatment group and overall. Baseline disease characteristics will also be presented per country and per site.

2.5.2 Demographics

At Screening, gender (male, female), year of birth, race (American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, White, other) and ethnic origin (Hispanic or Latino, not Hispanic or Latino) will be recorded. Age is derived in full years from birth to Screening.

Demographics will be presented by summary tables by treatment group and overall.

2.5.3 Concurrent Diagnoses and Medical History

Entries of medical history, including concurrent diagnoses and surgeries, are recorded with start and end times. Verbatim terms of conditions, diagnoses and surgeries will be coded according to MedDRA version 22.0 classified according to PTs and primary SOCs and presented in summary tables by treatment group and overall. Per primary SOC and PT, frequencies of subjects with at least one entry in the medical history and with at least one concurrent diagnosis will be displayed. An entry in the medical history is a concurrent diagnosis if ongoing at Screening.

2.5.4 Previous Therapies and Procedures

Among recorded therapies and procedures (cp. Section 2.4.4), a therapy or procedure is *previous* if started before Week 0 (Visit 1).

Per ATC level 2 and preferred name, frequencies of subjects with at least one previous therapy will be presented by summary tables by treatment group and overall.

2.5.5 Vital Signs, Height and Weight at Baseline

Systolic and diastolic blood pressure, pulse rate and body temperature at Week 0 (Visit 1), height and weight from physical examination at Week 0 (Visit 1) and the body mass index, derived as weight in kg divided by the square of height in meters will be presented by summary tables by treatment group and overall.

3 General Specifications for Analysis

3.1 Analysis Software

All analyzes will be carried out on the Statistical Analysis System (SAS®) Analytics Pro, version 9.4 including statistical procedures in the SAS/STAT® version 14.1 module [9].

3.2 Descriptive Statistics

All endpoints will be described using summary statistics for each treatment group, by visit where appropriate.

3.2.1 Summarizing Categorical Data

Categorical including binary data will be summarized by means of frequency counts and percentages. The denominator of percentages will be indicated if ambiguous. All categories will be displayed in summary tables including those with zero counts. A category of missing counts will be added.

3.2.2 Summarizing Continuous Data

Continuous data including scores will be summarized by means of arithmetic mean, standard deviation (SD), median, quartiles and ranges. Quartiles are given as *(lower quartile, upper quartile)* and ranges as *(minimum, maximum)*. The number of non-missing values will also be displayed.

3.3 Disposition of Subjects

A subject is screened if at least Visit 0 data are collected, randomized if Week 0 (Visit 1) data are collected and a randomization number given.

Subject disposition will be summarized by the number (counts and percentages) of patients

- screened
- failed screening
- randomized
- treated with an IP
- prematurely discontinued treatment
- completed treatment
- prematurely discontinued trial

Subject disposition will be summarized by treatment group and overall. Subject disposition will also be summarized per country and per site.

3.4 Premature Discontinuation of Treatment and/or Study Participation

On subject disposition (see Section 3.3) counts and percentages will also be given by reason for screen failure and premature discontinuation of treatment and/or trial participation, respectively.

3.5 Protocol Deviations

All observed protocol deviations will be listed. Following definitions in CTP Section 8.3 [1] and Section 4.2.3 below, protocol deviations will be classified as major or minor in a minuted meeting for blind review of the database conducted prior to final database lock and unblinding. All protocol deviations and major protocol deviations will be presented in frequency tables.

4 Study Specifications for Analysis

4.1 Interim Analysis

No interim analysis is planned.

4.2 Analysis Populations

4.2.1 Full Analysis Set

The full analysis set (FAS) includes all subjects randomly assigned to any IP following the intention-to-treat principle. Subjects in the FAS are evaluated by treatment randomized to.

4.2.2 Safety Set

The safety set consists of all randomized subjects, excluding subjects who return all the trial medication sealed. Subjects in the safety set are evaluated by treatment actually administered.

4.2.3 Per-Protocol Set

The per-protocol set (PPS) includes all subjects of the FAS who complete the trial without major protocol deviations deemed to affect the primary endpoint analysis.

If at least one major protocol deviation is observed in a subject of the FAS, the subject will be excluded from the PPS. Decisions on excluding subjects from the PPS will be taken and documented before unblinding the study.

Subjects may be excluded from the PPS if any of the following criteria are met:

- Failure to meet key inclusion/exclusion criteria; i.e. deviation of inclusion/exclusion criteria that according to medical judgement may impact the primary endpoint analysis,
- Usage of restricted medications/treatments; i.e. medications/treatments that according to medical judgement may impact the primary endpoint analysis,
- Nonadherence to the visit schedule at Week 8: 49 to 66 days.
- Noncompliance with the trial treatment regimen; adherence to the treatment schedule defined as application of the randomized IP as specified by the investigator for at least 80% of the days from Week 0 to Week 8 visit, and at least 80% of the days from Week 4 visit to Week 8 visit.

4.2.4 Usage of Analysis Populations

The FAS will be used as randomized for all efficacy analyses. As randomized, the PPS will only be used in sensitivity analyses for the non-inferiority evaluations and for the PTCS score evaluation. PGA success on the scalp will only be analyzed in subjects that at Week 0 (Visit 1) has a PGA of psoriasis severity on the scalp assessed as at least mild and a scalp involvement of at least 0.3% BSA.

Safety endpoints will be analyzed in the safety set as treated, i.e. by trial medication actually taken.

Demographics and baseline disease characteristics will be described for the FAS, the safety set and the PPS. Other baseline characteristics, exposure and concomitant therapies will be described for the FAS.

4.3 Subgroups and Confounding Factors for Analysis

No subgroup analysis will be done.

The potential confounding by trial sites will be considered by adjustment in statistical models. By country, trial sites with less than 14 FAS subjects will be pooled to form *analysis sites*. For trial sites with 14 FAS subjects or more, the trial site will be the analysis site.

The potential confounding by PGA at baseline will be considered by adjustment in all statistical models.

Depending on the endpoint as described below, the analysis may be adjusted by further confounding factors, particularly the baseline values of the respective endpoints in the models as covariates.

4.4 Time Points and Time Periods

4.4.1 Periods, Visits and Study Days

The screening period will last from Screening (Visit 0) to randomization (Visit 1/Week 0, Day 0). The treatment period will last from randomization (Visit 1/Week 0) to end of treatment (Visit 6/Week 8 or Early Termination). The follow-up period will last from end of treatment (Visit 6/Week 8) to end of follow-up (Visit 7/Week 10).

Up to six on-site visits are scheduled per subject, at Screening (Visit 0), and Weeks 0, 1, 4, 6 and 8. Visit 1/Week 0 is Day 0, the reference date. In addition, at Week 2, there is a telephone contact planned. Another on-site visit, or a telephone interview, will be held at end of follow-up. Further unscheduled visits may be done as deemed necessary.

Per visit, actual study days will be derived as the time passed in number of days since the reference date plus 1.

4.4.2 Dates

The date of Visit 1 is the reference date, Day 0. Start and end dates, and the dates of last study drug application and last contact will be related to Day 0, given as *study day*. Negative study days indicate dates before Visit 1.

It is assumed that treatment in the treatment period occurs from the evening of Day 0 to the day before Visit 6 (Week 8) unless otherwise reported.

4.4.3 Identification of Baseline Values

Per variable, baseline values are the values reported at randomization (Visit 1/Week 0, Day 0). If values at randomization are missing, values at Screening (Visit 0) are considered as baseline values.

4.4.4 Definitions of Change from Baseline

Changes from baseline are defined as post-baseline value minus baseline value. Reductions are negative changes. Percentage changes from baseline are defined as 100 times changes from baseline divided by the baseline value.

4.4.5 Treatment Emergency of Adverse Events

An AE is treatment-emergent if it occurs or deteriorates at or after the first application of the IP. Otherwise, an AE is not treatment-emergent.

4.4.6 Treatment Relation of Concomitant Medication

A medication or therapy is related to treatment by timing. It is concomitant or post-treatment as defined in Section 2.4.4 above.

4.5 Units and Unit Conversion

Dates are recorded by day, month and year. For imputation of partial dates, see Section 4.7.1 below. Days are calculated following rules in Sections 4.4.1, 4.4.2 and 4.6.1.

It is assumed that no unit conversion is required. Laboratory data will be reported in or converted to SI units [8].

4.6 Data Derivations

For the derivation of total and sum scores in efficacy endpoints see Section 2.4.1 above. Age, in full years, is derived as the number of calendar years from year of birth to the year of screening.

4.6.1 Durations

Time durations, i.e. lengths of time periods, will be derived as end date minus start date plus 1 day in days.

4.6.2 Analysis Visit Windowing

Post-baseline visits are to be held at Weeks 1, 2, 4, 6, 8 and 10, namely at study Day 7, 14, 28, 42, 56 and 70, within two days from target day. In addition, unscheduled visits and visits of Early Termination may be given. Comprehensive non-overlapping visit windows are defined for Week 4 from Day 14 to Day 42 and for Week 8 as Day 43 or later, respectively, excluding any follow-up visits.

Primary and secondary efficacy endpoints will be summarized and analysed by *analysis visits*. For Week 4 and Week 8, analysis visits are defined as the visits held closest to the target day per analysis visit window as defined above. Any other visits within the same visit window will not be considered for analysis of primary or secondary efficacy endpoints. If the time distances of two visits from the target day are identical, the later visit is selected for analysis.

Other efficacy endpoints and safety endpoints will be summarized and analysed by nominal visits as appropriate.

4.6.3 Classifications

Laboratory parameter values are abnormal low, if below the normal range, normal if within the normal range and abnormal high if above the normal range.

4.6.4 Notable Ranges

Notable ranges are defined for vital signs according to Table 3. Notable values are post-treatment values outside the notable ranges, whether above or below, and are indicated as *notably abnormal*.

Table 3: Definition of Notable Ranges of Vital Signs

Parameter	Notable Values and Changes from Baseline
Systolic blood pressure	90 mmHg or lower and decreased by 20 mmHg or more, 180 mmHg or greater and increased by 20 mmHg or more
Diastolic blood pressure	50 mmHg or lower and decreased by 15 mmHg or more, 110 mmHg or greater and increased by 15 mmHg or more
Pulse rate	50 bpm or lower and decreased by 15 bpm or more, 120 bpm or greater and increased by 15 bpm or more

4.7 Handling of Missing Data and Data Imputation

4.7.1 Partial Dates

If a date relevant for analysis is only partially available, it will be imputed by the central day of the time period of uncertainty, if this rule will be in line with potential dependencies of other complete dates in chronology. The time period of uncertainty is, for example, the calendar month, if the day is missing, or the calendar year, if day and month are missing. If the imputed date must be before or after a reference date, or between two reference dates, and the central day of the time period of uncertainty would not, the date will be imputed by the day of the time period of uncertainty closest to the central day but fulfilling the dependencies.

The approach of partial date imputation will be applied to visit dates, start and end dates of adverse events, start and end dates of concomitant medication, and the dates of last study drug application and of last contact, if partial.

4.7.2 Missing and Invalid Data

Data is missing if unavailable. Data is invalid if not considered for analysis. Invalidity of data is defined below, where applicable. On some analyses as indicated below, missing or invalid data will be replaced by the last observation carried forward (LOCF), by non-responder imputation (NRI) or by multiple imputation (MI).

Single Imputation by LOCF

On LOCF, a missing value is imputed by the last available value carried forward to the visit with the value missing until and including the last visit.

Single Imputation by NRI

On NRI, a missing value is imputed by a value representing non-response. Non-response is defined per endpoint below.

Multiple Imputation

Post-baseline missing or invalid data will be multiply imputed within treatment groups in two steps [10]. At first, any potential intermediate missing data will be imputed 100 times using a Markov Chain Monte Carlo (MCMC) method to obtain a monotone missing data pattern. In a second step, imputation is done using a monotone regression method, fitting an analysis of covariance (ANCOVA) model stepwise visit by visit. The ANCOVA model adjusts for the effects of analysis sites and baseline PGA as factors and the baseline value of the respective endpoint as a covariate. For each treatment group and each copy of the dataset, the estimated parameters and their variances are used to impute the endpoint. Starting with Week 1, the procedure is repeated on Week 4, Week 6 and Week 8, adding the possibly imputed values of the previous visits as covariates. On MI, missing score data is handled like missing continuous data.

If MI is applied to a dichotomized endpoint, the procedure is followed before dichotomization.

For analyses on the PPS or the safety set, no imputations will be made for missing data. For analyses on the FAS, LOCF, NRI and MI will be made for missing and invalid data.

5 Statistical Methods

5.1 Conventions for Inference

The two-sided level of significance is set to $\alpha = 0.05$, the one-sided level of significance to $\alpha = 0.025$.

5.2 Statistical Hypotheses

Let μ_M, μ_V and μ_C be estimators of treatment effect per endpoint for MC2-01 cream, MC2-01 cream vehicle and active comparator, respectively. The following four statistical hypotheses will be tested.

a. H_{0a}: $\mu_M = \mu_V$ versus H_{1a}: $\mu_M \neq \mu_V$,
b. H_{0b}: $\mu_C = \mu_V$ versus H_{1b}: $\mu_C \neq \mu_V$,
c. H_{0c}: $\mu_M - \mu_C \leq -\Delta$ versus H_{1c}: $\mu_M - \mu_C > -\Delta$,
d. H_{0d}: $\mu_M = \mu_C$ versus H_{1d}: $\mu_M \neq \mu_C$.

where Δ is a non-inferiority margin. In two-sided H_{0a}, H_{0b} and H_{0d}, if the hypothesis is rejected, superiority of the group with greater estimate is concluded. On one-sided H_{0c}, if the lower limit of the two-sided 95% confidence interval of the estimate of the treatment difference is greater than $-\Delta$, non-inferiority of MC2-01 cream as compared to active comparator is concluded.

The non-inferiority margin Δ is 12 percentage points on the primary endpoint mPASI and 10 percentage points on the first secondary endpoint PGA success.

5.3 Overall Strategy and Issues of Multiplicity

The statistical hypotheses on the primary and secondary endpoints will be tested confirmatorily in a hierarchical approach. The hierarchical approach ensures that the familywise type 1 error is controlled. According to the definitions in Section 5.2, hypotheses will be hierarchically tested starting with H_{0a} and H_{0b}. If, and only if H_{0a} and H_{0b} are, H_{0c} is tested confirmatorily. Hierarchical testing, H_{0a} to H_{0c}, will be done on the primary endpoint, mPASI, at first. If, and only if, non-inferiority of MC2-01 cream versus active comparator is claimed on the primary endpoint, the hierarchical testing procedure is continued on the first secondary endpoint, PGA success on the body. If, and only if, non-inferiority of MC2-01 cream versus active comparator is claimed on the first secondary endpoint, H_{0d} is tested on the second secondary endpoint, PTCS, on the primary endpoint, the first secondary endpoint, the third secondary endpoint, mPASI 75, where also H_{0a} is tested, and the fourth secondary endpoint, PGA success on the scalp, if, and only if, leading to rejections on the previous steps. In case hypotheses cannot be rejected, testing on this and all preceding hypotheses is exploratory only. The strategy is summarized in the following display.

Table 4: Overview on Hierarchical Hypothesis Testing Strategy with Different Endpoints and Hypotheses

Endpoint	Step	Hypothesis
%change in mPASI at Week 8	1	H_{0a} for superiority of MC2-01 cream over vehicle
	2	H_{0b} for superiority of active comparator cream over vehicle
	3	H_{0c} for non-inferiority of MC2-01 cream and active comparator
PGA success on the body at Week 8	4	H_{0a} for superiority of MC2-01 cream over vehicle
	5	H_{0b} for superiority of active comparator cream over vehicle
	6	H_{0c} for non-inferiority of MC2-01 cream and active comparator
PTCS at Week 8	7	H_{0d} for superiority of MC2-01 cream over active comparator
%change in mPASI at Week 8	8	H_{0d} for superiority of MC2-01 cream over active comparator
PGA success on the body at Week 8	9	H_{0d} for superiority of MC2-01 cream over active comparator
mPASI 75 at Week 8	10	H_{0a} for superiority of MC2-01 cream over vehicle
	11	H_{0d} for superiority of MC2-01 cream over active comparator
PGA success on the scalp at Week 8	12	H_{0a} for superiority of MC2-01 cream over vehicle
	13	H_{0d} for superiority of MC2-01 cream over active comparator

For definitions of hypotheses, see Section 5.2

Efficacy endpoints will be analyzed by different estimands, following a treatment-policy strategy (primary estimand) and a hypothetical strategy (secondary estimand). Only on PTCS, a while-on-treatment strategy will be followed.

According to the *treatment-policy strategy*, all available data will be used, regardless of treatment and treatment adherence. Missing data will be imputed.

According to the *hypothetical strategy*, only data will be used under the IP and while adhering to the IP. If any other psoriasis treatment is used, i.e. treatment that according to medical judgement may improve psoriasis, data during and after the period of other psoriasis treatment intake is invalid and will not be used. Furthermore, data will not be used of periods of potential non-adherence to the IP. Non-adherence will be given if the IP is applied on less than 70% of the days from Day 0 to Week 4, Week 4 to Week 6, and Week 6 to Week 8, respectively. Non-adherence to the IP is derived from days with treatment non-compliance. Missing and invalid data will be imputed.

According to the *while-on-treatment strategy*, all available data will be used with treatment within 7 days before assessment. Data without treatment within 7 days before assessment is invalid and will not be used.

Hierarchical testing applies to the primary analyses per endpoint, not to any sensitivity analyses as described below.

5.4 Approaches for Analysis

All collected, recorded and derived data for analysis will be listed in individual subject data listings. On vital signs, notably abnormal values will be flagged.

5.4.1 Categorical Data

Categorical data will be described by summary tables of counts and percentages. This includes zero counts of all possible categories, if any. Selected categorical variables will be visualized by bar charts by treatment group and visit.

Logistic Regression

Selected binary endpoints will be analysed using a logistic model with effects of treatment, analysis site, and the baseline value of the endpoint as factors. The estimated log odds ratios between the treatments at Week 8 will be derived together with the associated standard error and 95% confidence intervals.

Logistic regression is used for superiority testing only. The non-inferiority analysis is done without modelling by means of unadjusted risk differences between treatment groups. 95% confidence intervals will be computed using normal approximation.

5.4.2 Continuous Data

Continuous data will be described by summary tables of sample characteristics by visit and for changes between visits as defined in Section 3.2.2 above. Selected continuous variables will be visualized by box plots by treatment group and visit.

ANCOVA

Comparisons will be done using an ANCOVA model including treatment, analysis site and baseline PGA as factors, and the baseline value of the respective endpoint as covariate.

The statistical model is

$$\text{endpoint} = \beta_0 + \beta_1 \cdot \text{baseline value} + \text{treatment} + \text{analysis site} + \text{baseline PGA} + \text{error},$$

where β_1 is a parameter of fixed effect and β_0 an intercept. The estimated differences between MC2-01 cream and MC2-01 cream vehicle, and between active comparator and MC2-01 cream vehicle, will be reported together with the associated standard error, the 95% confidence interval, and a p-value of a test on the respective contrast corresponding to the hypothesis as either no difference between treatments (superiority) or difference between treatments within the tolerance limits (non-inferiority) (cp. Section 5.2 above).

5.5 Application of Methods

Statistical hierarchical testing will be applied as described in Sections 5.2, 5.3 and 5.4.2. The following subsections describe the application of methods per endpoint, summarized in Section 5.5.5.

Multiple Imputation Procedure

Applicable to all endpoints with MI, the procedural steps are as follows. On 100 multiple complete imputed datasets as produced according to the instructions in Section 4.7.2, each dataset will be analyzed the same way. After pooling the estimates and standard errors, results will be aggregated following Rubin's rule to yield one estimate and associated standard error, and to calculate the 95% confidence interval for the

treatment difference or odds ratio between treatments, respectively [11]. In data listings, multiply imputed values are described by mean values.

5.5.1 Primary Efficacy Endpoint

The primary endpoint, percentage change in mPASI on the body (trunk and/or limbs) from baseline to Week 8, will be analysed for superiority and non-inferiority by means of ANCOVA as described in Section 5.4.2.

Primary Analysis

The primary analysis is for superiority of MC2-01 cream versus MC2-01 cream vehicle and of active comparator versus MC2-01 cream vehicle as well as for non-inferiority of MC2-01 cream and active comparator using the treatment-policy strategy on the FAS using MI to impute missing data. The estimated pairwise differences will be reported together with the associated standard error, the 95% confidence interval, and p-values of testing hypotheses H_{0a}, H_{0b} and H_{0c}.

Sensitivity Analyses

Three sensitivity analyses are planned on the primary endpoint, for internal validity a tipping-point analysis following the treatment-policy strategy on the FAS using MI, for external validity an analysis following the hypothetical strategy on the FAS using MI, and an analysis following the treatment-policy strategy on the PPS.

Tipping-point analysis for treatment-policy strategy on the FAS

The primary analysis of mPASI using MI assumes that data are missing at random (MAR) within each treatment arm. For each hypothesis, the sensitivity of the MAR assumption is investigated using a tipping-point analysis [10]. A penalty (a number) is added to all imputed Week 8 values in the MC2-01 cream treatment group or (when testing superiority of active comparator vs. vehicle) in the active comparator group, and the analysis is repeated on the FAS for the treatment-policy estimand as described for the primary analysis. The penalty is gradually increased until the point, the tipping point, where the conclusion (superiority or non-inferiority depending on the analysis) no longer is demonstrated. The tipping point will be reported. If the tipping point is considered a clinical plausible difference, the tipping-point analysis does not support the primary analysis.

Hypothetical strategy on the FAS

mPASI is analysed following the hypothetical strategy as described in Section 5.3 on the FAS using MI for missing and unused data as described in Section 4.7.2.

Treatment-policy strategy on the PPS

mPASI is analysed following the treatment-policy strategy as in the primary analysis, but on the PPS without using any imputation approach.

5.5.2 Secondary Efficacy Endpoints

If the primary analysis of the primary endpoint yields rejection of all hypotheses, hierarchical testing is continued on secondary efficacy endpoints, with PGA success at first, followed by PTCS.

PGA Success

The first and fourth secondary efficacy endpoints are the percentage of subjects with PGA success on the body and on the scalp at Week 8, respectively, as defined in Section 2.4.1.

Primary Analysis

Considering PGA as a continuous endpoint, missing data will be imputed using MI as described in Section 4.7.2. On derived PGA success following the treatment-policy strategy, hypotheses will be tested, and comparisons of treatment groups done using a model of logistic regression as described in Sections 5.2 and 5.4.1.

Sensitivity Analyses

Two sensitivity analyses are planned on the first secondary endpoint, an analysis following the treatment-policy strategy on the FAS using NRI, and an analysis following the hypothetical strategy on the FAS using MI. On PGA success, non-response is defined as no PGA success.

In a first sensitivity analysis, missing data will be imputed using NRI as described in Section 4.7.2. On PGA success following a treatment-policy strategy, hypotheses will be tested, and comparisons of treatment groups done using a model of logistic regression as described in Sections 5.2 and 5.4.1.

In a second sensitivity analysis, missing and unused data will be imputed using MI as described in Section 4.7.2. On PGA success following a hypothetical strategy, hypotheses will be tested, and comparisons of treatment groups done using a model of logistic regression as described in Sections 5.2 and 5.4.1.

PTCS

The second secondary efficacy endpoint is the PTCS total score at Week 8. The PTCS total score will be analysed by means of an analysis of variance with treatment, analysis site and baseline PGA as factors.

Analysis of PTCS is following a while-on-treatment strategy using data as described in Section 5.3. Missing and invalid data will be imputed using LOCF as described in Section 4.7.2.

Primary Analysis

The primary analysis of PTCS will be based on the FAS following the while-on-treatment approach using LOCF to impute missing and invalid data.

Sensitivity Analyses

A sensitivity analysis of PTCS will be based on the PPS following the while-on-treatment approach using LOCF to impute missing and invalid data.

mPASI 75

mPASI 75 at Week 8 will be analyzed on the FAS like PGA success at Week 8 including mPASI at baseline as a covariate.

5.5.3 Other Efficacy Endpoints

Efficacy and subject-related outcomes will be exploratorily analyzed on the FAS at Week 4 as described for Week 8 above, without formal statistical hypothesis testing.

PGA success at Week 4 and mPASI 75 at Week 4 will be analyzed on the FAS like PGA success at Week 8 including mPASI at baseline as a covariate (cp. Section 5.4.1).

Observed changes from baseline in SGA at Week 4 and Week 8, and changes from baseline in BSA involvement on the body and the scalp will be presented descriptively on the FAS.

DLQI domain and total scores will be summarized by visit including changes from baseline on the FAS. Changes from baseline in the DLQI total score at Week 4 and Week 8 will be analyzed on the FAS by means of ANCOVA as described in Section 5.4.2 utilizing MI. Imputations will be made on the domain level to derive the DLQI total score.

Observed EQ-5D dimension assessments and VAS score will be summarized by visit on the FAS. Changes from baseline in EQ-VAS score at Week 4 and Week 8 will be analyzed on the FAS by means of ANCOVA as described in Section 5.4.2 utilizing MI.

5.5.4 Safety Endpoints

Safety endpoints will be described by summary tables on the safety set.

Treatment-emergent AEs, treatment-emergent drug-related AEs, treatment-emergent serious AEs (SAE), treatment-emergent drug-related SAEs, treatment-emergent AEs leading to drug withdrawal and treatment-emergent drug-related AEs leading to drug withdrawal will be presented in summary tables of the number and percentage of patients with at least one event, and the number of events, by primary SOC and by PT within SOC. For treatment-emergent AEs and treatment-emergent drug-related AEs, these summaries will also be provided by maximum severity. Maximum severity is defined by the most severe AE per subject, SOC and PT.

Local skin reaction sum scores, vital signs and laboratory parameters will be summarized by visit including changes from baseline. Frequencies of most intense local skin reactions will also be presented.

Using normal ranges, laboratory parameters will be presented in shift tables of frequencies of all categorized post-baseline versus categorized baseline values per visit (cp. Section 4.6.3).

5.5.5 Overview of Analyses

Analyses of safety endpoints are descriptive and exploratory. Planned analyses of primary, secondary and other endpoints for efficacy, including subject-related outcomes, are summarized in the following table.

Table 5: Overview of analyses of endpoints for efficacy (with strategy, analysis set and imputation)

Sequence	Endpoint	Primary Analysis	Sensitivity Analyses
Primary	%change in mPASI at W8	treatment policy, FAS, MI	1: treatment policy, FAS, MI, TPA; 2: hypothetical, FAS, MI; 3: treatment policy, PPS
Secondary	PGA success on the body at W8	treatment policy, FAS, MI	1: treatment policy, FAS, NRI; 2: hypothetical, FAS, MI
	PTCS* at W8	while on treatment, FAS, LOCF	while on treatment, PPS, LOCF
	mPASI 75	treatment policy, FAS, MI	-
	PGA success on the scalp at W8	treatment policy, FAS, MI	1: treatment policy, FAS, NRI; 2: hypothetical, FAS, MI
Other**	%change in mPASI at W4	treatment policy, FAS, MI	-
	PTCS at W4	while on treatment, FAS, LOCF	-
	Change in DLQI	treatment policy, FAS, MI	-
	Change in EQ-VAS	treatment policy, FAS, MI	-

* for superiority of MC2-01 cream versus active comparator, ** additionally EQ-5D, SGA and BSA involvement descriptively only
TPA = tipping-point analysis, W = Week

On primary analysis, for an overview on hierarchical testing of statistical hypotheses on the primary endpoint and the secondary endpoints, see Table 4 in Section 5.3.

5.6 Changes and Clarifications to Analyses Planned in the Study Protocol

Analyses are planned according to protocol. Defined as other endpoints in the protocol, mPASI75 at Week 8 and PGA of psoriasis severity on the scalp at Week 8 are promoted to secondary endpoints at the end of the hierarchical testing strategy on secondary endpoints. For PGA on the scalp, the same definition of PGA success as on the body will be applied, on a subpopulation of subjects with psoriasis on the scalp at baseline. To clarify, BSA assessment is listed under other endpoints. Notable laboratory normal ranges will not be defined.

5.7 Implementation

The SAP will be implemented by means of statistical programming specifications with technical details of implementation and the level of validation of programming. Derived datasets will be produced following raw-to-derived data specifications, tables, figures and listings (TFL) according to approved TFL shell versions.

6 References

[1] MC Therapeutics: Clinical Trial Protocol, Version 2.0, dated 16 October 2018.

[2] International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (1995). Topic E3. Structure and Content of Clinical Study Reports.

[3] International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (2016). Integrated Addendum to ICH E6(R1): Guideline for Good Clinical Practice. Topic E6(R2). Guideline for Good Clinical Practice.

[4] International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (1998). Topic E9. Statistical Principles for Clinical Trials.

[5] MC Therapeutics: Case Report Form, Version 2.0, dated 21 January 2019.

[6] Finlay AY, Khan GK. Dermatology Life Quality Index (DLQI) - simple practical measure for routine clinical use. Clin Exp Dermatol. 1994;19(3):210-6.

[7] Rabin R, de CF. EQ-5D: a measure of health status from the EuroQol Group. Ann Med. 2001;33(5):337-43.

[8] Kratz A, Ferraro M, Sluss PM et al. (2004): Laboratory reference values. N Engl J Med 351: 1548-63.

[9] SAS Institute Inc. SAS/STAT® 14.1 User's Guide. SAS Institute Inc. 2016.

[10] O'Kelly M, Ratitch B (2014): Clinical Trials with Missing Data: A Guide for Practitioners. Wiley, Chichester, UK.

[11] Rubin DB. Multiple Imputation for Nonresponse in Surveys. New York, NY, USA: Wiley & Sons; 1987.

7 Deliverables

Based on a TFL shells document, the following tables, figures and listings will be produced.

7.1 In-text Tables

In-text tables will be produced identical to selected post-text tables by content and in the layout of the report.

7.2 In-text Figures

In-text figures will be copied from selected post-text figures.

7.3 Post-text Tables

The following post-text tables will be produced. All tables will summarize results by treatment group.

Table 6: Table of Content of Post-Text Tables

14.1	**Disposition, Baseline Characteristics, Exposure and Concomitant Therapies**
Table 14.1.1.1	Subject Disposition – Summary
Table 14.1.1.2	Subject Disposition – Summary by Country
Table 14.1.1.3	Subject Disposition – Summary by Site
Table 14.1.2	Analysis Sets – Summary
Table 14.1.3.1	Inclusion and Exclusion Criteria – Summary
Table 14.1.3.2	Protocol Deviations – Summary
Table 14.1.3.3	Major Protocol Deviations – Summary
Table 14.1.4.1	Demographics – Summary on the FAS
Table 14.1.4.2	Demographics – Summary on the Safety Set
Table 14.1.4.3	Demographics – Summary on the PPS
Table 14.1.5.1	Baseline Disease Characteristics – Summary on the FAS
Table 14.1.5.2	Baseline Disease Characteristics – Summary on the Safety Set
Table 14.1.5.3	Baseline Disease Characteristics – Summary on the PPS
Table 14.1.5.4	Baseline Disease Characteristics – Summary by Country on the FAS
Table 14.1.5.5	Baseline Disease Characteristics – Summary by Site on the FAS
Table 14.1.6	Vital Signs, Height and Weight at Baseline – Summary on the FAS
Table 14.1.7.1	Medical History – Summary by SOC and PT on the FAS
Table 14.1.7.2	Concurrent Diagnoses at Screening – Summary by SOC and PT on the FAS
Table 14.1.8.1	Previous Drug Therapies – Summary by ATC Class and PT on the FAS
Table 14.1.8.2	Concomitant Drug Therapies – Summary by ATC Class and PT on the FAS
Table 14.1.8.3	Post-Treatment Drug Therapies – Summary by ATC Class and PT on the FAS
Table 14.1.9	Exposure and Treatment Compliance - Summary on the FAS
14.2	**Efficacy Data**
14.2.1	**Primary Efficacy Endpoint**
Table 14.2.1.1	mPASI – Summary by Visit on the FAS, Observed Cases
Table 14.2.1.2	mPASI Score – Summary by Visit on the FAS, Averaged Multiple Imputations, Treatment-policy Strategy
Table 14.2.1.3	mPASI Score – Summary by Visit on the FAS, Averaged Multiple Imputations, Hypothetical Strategy
Table 14.2.1.4	mPASI Score – Summary by Visit on the PPS, Observed Cases
Table 14.2.1.5	%Change in mPASI Score at Week 8 – Primary Analysis on the FAS using MI, Treatment-policy Strategy

Table 14.2.3.6	DLQI Score – Summary by Visit on the FAS, Averaged Multiple Imputations, Treatment-policy Strategy
Table 14.2.3.7	DLQI Score at Week 4 and Week 8 – Analysis on the FAS using MI, Treatment-policy Strategy
Table 14.2.3.8	EQ-5D – Summary by Visit on the FAS, Observed Cases
Table 14.2.3.9	EQ-VAS at Week 4 and Week 8 – Analysis on the FAS using MI, Treatment-policy Strategy
Table 14.2.3.10	BSA Involvementon the Body and the Scalp – Summary by Visit on the FAS, Observed Cases
14.3	**Safety Data**
14.3.1	**Adverse Events**
Table 14.3.1.1	Adverse Events – Overview on the Safety Set
Table 14.3.1.2	Treatment-emergent Adverse Events – Summary by SOC and PT on the Safety Set
Table 14.3.1.3	Treatment-emergent Adverse Events – Summary by Maximum Severity, SOC and PT on the Safety Set
Table 14.3.1.4	Treatment-emergent Drug-related Adverse Events – Summary by SOC and PT on the Safety Set
Table 14.3.1.5	Treatment-emergent Drug-related Adverse Events – Summary by Maximum Severity, SOC and PT on the Safety Set
Table 14.3.1.6	Treatment-emergent Serious Adverse Events – Summary by SOC and PT on the Safety Set
Table 14.3.1.7	Treatment-emergent Drug-related Serious Adverse Events – Summary by SOC and PT on the Safety Set
Table 14.3.1.8	Treatment-emergent Adverse Events Leading to Drug Withdrawal – Summary by SOC and PT on the Safety Set
Table 14.3.1.9	Treatment-emergent Drug-related Adverse Events Located in the Treatment Area – Summary by SOC and PT on the Safety Set
14.3.2	**Local Skin Reaction**
Table 14.3.2.1	Local Skin Reaction Assessments – Summary by Visit on the Safety Set
Table 14.3.2.2	Most Intense Local Skin Reaction Assessments – Summary on the Safety Set
14.3.3	**Laboratory Parameters**
Table 14.3.3.1.1	Serum Biochemistry – Summary by Visit on the Safety Set
Table 14.3.3.1.2	Serum Biochemistry – Shifts from Baseline by Visit on the Safety Set
Table 14.3.3.2.1	Urinalysis – Summary by Visit on the Safety Set
Table 14.3.3.2.2	Urinalysis – Shifts from Baseline by Visit on the Safety Set
Table 14.3.3.3	Urine Pregnancy Tests – Summary by Visit on the Safety Set
14.3.4	**Vital Signs**
Table 14.3.4	Vital Signs - Summary by Visit on the Safety Set
14.3.5	**Physical Examination**
Table 14.3.5	Physical and Dermatological Examination – Summary by Visit on the Safety Set

7.4 Post-text Figures

The following post-text figures will be produced.

Table 7: Table of Content of Post-Text Figures

Figure 14.2.1.1	Percentage Change from Baseline in mPASI – Box Plot by Treatment and Visit, Treatment Policy Strategy, FAS Using MI
Figure 14.2.1.2	Percentage Change from Baseline in mPASI – Box Plot by Treatment and Visit, Treatment Policy Strategy, PPS
Figure 14.2.2.1	PGA Success Rate on the Body – Bar Chart by Treatment and Visit, Treatment Policy Strategy, FAS Using MI
Figure 14.2.3.1	PTCS Total Score - Box Plot by Treatment and Visit, While on Treatment Strategy, FAS Using LOCF

7.5 Statistical Output Documentation

Original SAS® outputs of inferential statistical procedures will be available upon request.

7.6 Data Listings

The following individual patient data listings will be produced.

Table 8: Table of Content of Data Listings

16.1	
Listing 16.1.7	Subject Enrollment and Randomization
16.2	**Patient Data Listings**
16.2.1	**End of Study**
Listing 16.2.1.1	Screen Failure
Listing 16.2.1.2	Subject Disposition
Listing 16.2.1.3	Follow-up Visit
16.2.2	**Protocol Deviations**
Listing 16.2.2.1	Protocol Deviations
16.2.3	**Analysis Populations**
Listing 16.2.3.1	Analyses Populations
16.2.4	**Baseline Characteristics**
16.2.4.1	**Demographics**
Listing 16.2.4.1.1	Demographic Data, Height and Weight
16.2.4.2	**Diagnosis**
Listing 16.2.4.2.1	Details of Diagnosis and Psoriasis Treatment History
16.2.4.3	**Prior and Concomitant Therapy**
Listing 16.2.4.3.1	Prior and Concomitant Therapy – Details
Listing 16.2.4.3.2	Prior and Concomitant Therapy – Coding
16.2.4.4	**Prior and Concomitant Procedures**
Listing 16.2.4.4.1	Concomitant Procedures - Details
Listing 16.2.4.4.2	Concomitant Procedures - Coding
16.2.4.5	**Medical History**
Listing 16.2.4.5.1	Medical History - Details
Listing 16.2.4.5.2	Medical History - Coding
16.2.5	**Other Collected Data**
Listing 16.2.5.1	Inclusion Criteria

Listing 16.2.5.2	Exclusion Criteria
Listing 16.2.5.3	Dates of Visits and Times Between Visits
Listing 16.2.5.4	End of Treatment
Listing 16.2.5.5	Study Drug Assignment per Visit
Listing 16.2.5.6	Study Drug Return Weight per Visit
Listing 16.2.5.7	Study Drug Accountability per Visit
Listing 16.2.5.8	Subject Treatment Diary Card & Dose Documentation
Listing 16.2.5.9	Overall Study Drug Exposure and Compliance
Listing 16.2.5.10	General Comments
16.2.6	**Individual Efficacy Data**
Listing 16.2.6.1.1	mPASI Score by Visit
Listing 16.2.6.1.2	mPASI Assessments by Visit
Listing 16.2.6.2.1	PGA on the Body by Visit
Listing 16.2.6.2.2	PGA on the Scalp by Visit
Listing 16.2.6.3	PTCS by Visit
Listing 16.2.6.4	SGA by Visit
Listing 16.2.6.5	DLQI by Visit
Listing 16.2.6.6	EQ-5D by Visit
Listing 16.2.6.7	BSA Involvement Assessment by Visit
16.2.7	**Safety & Tolerability Listings**
16.2.7.1	**Adverse Events Listings**
Listing 16.2.7.1.1	Adverse Events - Timing
Listing 16.2.7.1.2	Adverse Events – Details
Listing 16.2.7.1.3	Adverse Events – Coding
Listing 16.2.7.1.4	Serious Adverse Events - Timing
Listing 16.2.7.1.5	Serious Adverse Events – Details
Listing 16.2.7.1.6	Serious Adverse Events – Coding
Listing 16.2.7.1.7	Adverse Events Leading to Drug Withdrawal - Timing
Listing 16.2.7.1.8	Adverse Events Leading to Drug Withdrawal – Details
Listing 16.2.7.1.9	Adverse Events Leading to Drug Withdrawal – Coding
16.2.7.2	**Local Skin Reactions**
Listing 16.2.7.2	Local Skin Reactions Assessment by Visit
16.2.7.3	**Laboratory Parameters**
Listing 16.2.7.3.1	Serum Biochemistry by Visit
Listing 16.2.7.3.2	Urinalysis by Visit
Listing 16.2.7.3.3	Urine Pregnancy Test by Visit
16.2.7.4	**Vital Signs**
Listing 16.2.7.4	Vital Signs by Visit
16.2.7.5	**Physical Examination**
Listing 16.2.7.5	Physical and Dermatological Examination by Visit