Final Analysis
Statistical Analysis Plan (SAP)

Protocol Title:	A Phase II Open-label Multicenter Study to Assess the Efficacy and Safety of AFM13 in Patients with Relapsed or Refractory CD30-positive Peripheral T-cell Lymphoma or Transformed Mycosis Fungoides (REDIRECT)
Protocol Number:	AFM13-202
Protocol Version, Date	V6.0, 12 July 2021
ICON ID:	3499/0003
Document Version, Date:	Final 3.0, 13 September 2022

Prepared by:

ICON Clinical Research Services

On behalf of:

Affimed GmbH

SIGNATURE PAGE

Prepared at ICON Clinical Research by:

[redacted], Senior Statistician I — Date (DDMMMYYYY)

Reviewed at ICON Clinical Research by:

[redacted], Principal Biostatistician — Date (DDMMMYYYY)

Approved at Affimed GmbH by:

[redacted] Senior Statistician — Date (DDMMMYYYY)

[redacted] Senior Director Data Sciences — Date (DDMMMYYYY)

REVISION HISTORY

Version	Date	Section	Changes implemented
1.0	09JAN2020	N/A	Initial version
2.0	01SEP2020	[redacted]	- Addition of subgroups for overall response rate and duration of response. - Addition of new parameters and analyses due to COVID-19. - Addition of new parameters and analyses due to protocol amendment dated 02-Jul-2020.
3.0	13SEP2022	[redacted]	- Clarification on the category names for premedications. - Replace “Most frequent TEAEs” by “Most frequent non-serious TEAEs”. - Addition of COVID-19 vaccination. - Integration of protocol amendment dated 29-Jun-2021. - Modification of Tables 2 and 4. - Modification of section 7.1. - Modification of section 7.2.2. - Modification of section 8.7.3. - Modification of section 8.6.2. - Modification of section 8.8.1. - Modification of section 8.8.5. - Modification of section 9.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

The following abbreviations will be used within this SAP:

Abbreviation or special term	Explanation
ADA	anti-drug antibodies
AE	adverse event
ALP	alkaline phosphatase
ALT	alanine transaminase
AST	aspartate transaminase
ATC	Anatomical, Therapeutic and Chemical
AUC	area under the concentration versus time curve
BLQ	Below limit of quantification
BMI	body mass index
BOR	best overall response
CAILS	Composite Assessment of Index Lesions Severity
CI	confidence interval
CKD-EPI	Chronic Kidney Disease Epidemiology Collaboration
CL	clearance
Cmax	maximum concentration
CR	complete response/remission
CRP	C-reactive protein
CSR	clinical study report
CT	computed tomography
CTCAE	common terminology criteria for adverse events
CV	Coefficient of variation
DOR	duration of response
ECG	electrocardiogram
ECOG	Eastern Cooperative Oncology Group
eCRF	electronic Case Report Form
EQ-5D	European Quality of Life 5-Dimensional
FAS	full analysis set
FDG	fluorodeoxyglucose
ICF	informed consent form
ICH	International Council for Harmonization
IME	important medical event
IRC	Independent Review Committee
IRR	infusion-related reaction

ITT	intent to treat
LDH	lactate dehydrogenase
LLN	lower limit of normal
LLOQ	lower limit of quantitation
MedDRA	Medical Dictionary for Regulatory Activities
mSWAT	modified Severity Weighted Assessment Tool
NCA	non-compartmental analysis
NSAID	nonsteroidal anti-inflammatory drug
ORR	overall (or objective) response rate
OS	overall survival
PD	progressive disease
PET	positron emission tomography
PFS	progression-free survival
PK	pharmacokinetic(s)
PR	partial response/remission
PT	preferred term
PTCL	peripheral T-cell lymphoma
Q2W	every other week
QOL	quality of life
QTc	corrected Q wave to T wave interval
R/R	relapse or refractory
SAE	serious adverse event
SAP	statistical analysis plan
SD	stable disease
SDTM	Study Data Tabulation Model
SMQ	Standardized MedDRA Queries
SOC	system organ class
t1/2	half-life
TEAE	treatment-emergent adverse event
TFL	tables, figures, listings
TMF	transformed mycosis fungoides
TNMB	primary tumor/lymph nodes/metastasis/tumor burden
ULN	upper limit of normal
ULOQ	upper limit of quantitation
Vss	volume of distribution at steady state
WBC	white blood cells
WHO	World Health Organization
WHODDE	WHO Drug Dictionary Enhanced

1 INTRODUCTION

The purpose of this Statistical Analysis Plan (SAP) is to provide detailed descriptions of the statistical methods, data derivations and data displays for study protocol AFM13-202 Version 6.0 "A Phase II Open-label Multicenter Study to Assess the Efficacy and Safety of AFM13 in Patients with Relapsed or Refractory CD30-positive Peripheral T-cell Lymphoma or Transformed Mycosis Fungoides (REDIRECT)" dated 12 July 2021 for interim and final analyses. The table of contents and templates for the TFLs will be produced in a separate document.

Any deviations from this SAP will be described and justified in the Clinical Study Report (CSR).

The preparation of this SAP has been based on International Conference on Harmonization (ICH) E9 and E3 guidelines.

All data analyses and generation of TFLs will be performed using SAS 9.4® or higher.

2 STUDY OBJECTIVES

2.1 Primary objective

- To assess the antitumor activity of AFM13 by an Independent Review Committee (IRC) confirmed -PET-CT based overall response rate (ORR).

2.2 Secondary objectives

The secondary objectives of this study are as follows:

- To assess the antitumor activity of AFM13 by IRC-confirmed complete response (CR) and partial response (PR) rates and computed tomography (CT) scan- based ORR.
- To assess the antitumor activity of AFM13 by Investigator-assessed ORR (defined as ORR-2).
- To assess the duration of response (DOR) to AFM13.
- To assess the safety and tolerability of AFM13.
- To assess the pharmacokinetics (PK) of AFM13.
- To assess the immunogenicity of AFM13.
- To assess the Quality of Life (QOL) of subjects while on treatment with AFM13.

2.3 Exploratory objectives

[redacted]

2.4 Safety objective

Not applicable.

3 STUDY DESIGN

3.1 General study design

This is an open-label, multicenter, Phase 2 study in subjects with R/R CD30-positive PTCL and TMF.

Subjects with PTCL will be enrolled to Cohort A or B based on the centrally determined CD30 expression level detailed below:

- Subjects with ≥10% CD30 expression for Cohort A
- Subjects with ≥1% to <10% CD30 expression for Cohort B

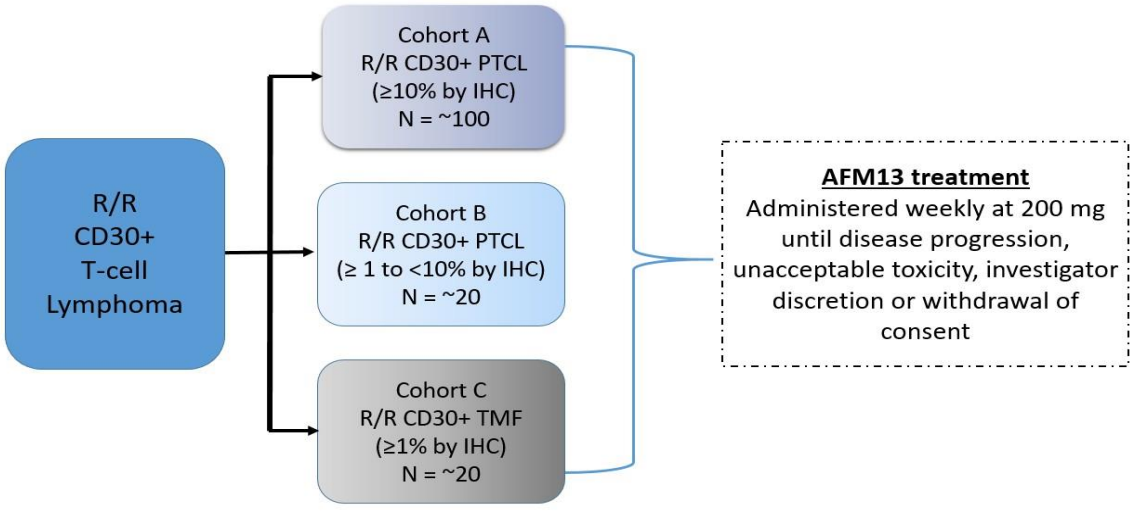
After the planned Interim Analysis (IA), Cohorts A and B may be combined with the CD30 positivity defined as ≥1% by centrally assessed IHC.

Subjects with TMF and centrally determined ≥1% CD30 expression will be enrolled to Cohort C.

The study targets to investigate the antitumor activity and safety of AFM13 dosed weekly at 200 mg until disease progression, unacceptable toxicity, Investigator discretion or withdrawal of consent.

The Study Flow Chart is presented in Figure 1.

Figure 1: Study Flow Chart



Abbreviations: CD = cluster of differentiation; IHC = immunohistochemistry; PTCL = peripheral T-cell lymphoma; R/R = relapsed/refractory; TMF = transformed mycosis fungoides.

All subjects enrolled may participate in the following study periods:

- Pre-screening period (up to 28 days prior to starting the screening period).
- Screening period (up to 21 days prior to first dose of AFM13).
- Treatment period.
- Efficacy and Safety follow-up period (30 to 37 days from last dose of AFM13).
- Survival follow-up (after completion of the study) every three months after Final study visit (+/- 1 month).

Subjects will undergo study assessments from the Screening period to the Efficacy and Safety follow-up period at scheduled site visits. All subjects are considered "on study" until they complete the Efficacy and Safety follow-up period, withdraw consent, are lost to follow up, or die.

A subject is considered to have completed the study if the subject has either a disease progression or died or completed the Efficacy and Safety follow-up or when the subject has at least a partial response/remission and gets transitioned to a transplant.

3.2 Randomization and blinding

Not applicable.

3.3 Study treatments and assessments

AFM13 will be administered intravenously [redacted] Subjects will be treated weekly until disease progression, unacceptable toxicity, Investigator discretion, or withdrawal of consent. After 16 weeks (from Cycle 2 Day 50 onwards), subjects who are experiencing persistent infusion-related reactions (IRRs) and have achieved an objective response (at least PR) may be dosed on an every other week schedule (Q2W), at the discretion of the Investigator, until disease progression, unacceptable toxicity, Investigator discretion or withdrawal of consent.

A detailed description of procedures and assessments to be conducted during this study is presented in Appendix A of the protocol. A cycle in the study is defined as an 8-weeks period.

4 STUDY ENDPOINTS

4.1 Primary efficacy endpoint

The primary efficacy endpoint of this study is:

- ORR (CR + PR) as confirmed by an IRC as assessed by the modified Lugano Classification (Cheson, 2014) for Cohorts A and B (PTCL) based on PET-CT and after at least 8 weeks from the first assessment as assessed by Olsen Criteria (Olsen, 2011) for Cohort C (TMF).

4.2 Secondary efficacy endpoints

The secondary efficacy endpoints of this study are:

- CR rate, PR rate, and CT-based ORR as confirmed by an IRC as assessed by the modified Lugano Classification (Cheson, 2014) for Cohorts A and B (PTCL) based on CT and after at least 8 weeks from the first assessment by Olsen Criteria (Olsen, 2011) for Cohort C (TMF).
- ORR-2 confirmed by Investigator assessment as assessed by the modified Lugano Classification (Cheson, 2014) for Cohorts A and B (PTCL) and after at least 8 weeks from the first assessment by Olsen Criteria (Olsen, 2011) for Cohort C (TMF).
- DOR as assessed by the modified Lugano Classification (Cheson, 2014) for Cohorts A and B (PTCL) and after at least 8 weeks from the first assessment by Olsen Criteria (Olsen, 2011) for Cohort C (TMF) based on IRC assessment.
- Assessment of non-compartmental PK parameters including C_{max}, AUC, volume of distribution at steady state (V_{ss}), half-life ($t_{1/2}$).
- Incidence of subjects who develop anti-drug antibodies (ADA) during treatment in blood and their neutralizing potential.
- QOL as measured by European Quality of Life 5-Dimensional (EQ-5D) for Cohorts A and B, and by Skindex-29 for Cohort C.

4.3 Exploratory endpoints

[redacted]

[redacted]

[redacted]

4.4 Safety endpoints

The safety endpoints of this study are:

- Number and frequency of treatment-related adverse events (AEs).

5 SAMPLE SIZE AND POWER

Refer to protocol section 8.8.1.

6 ANALYSIS POPULATIONS

6.1 Full analysis set (FAS)

The full analysis set (FAS) following the intent to treat (ITT) principle will consist of all subjects who received at least one dose of AFM13.

The FAS will be the primary population for all efficacy related endpoints and the primary objective. A sensitivity analysis for all subjects who received at least one dose of AFM13 and had at least one post-baseline efficacy assessment will be conducted to support the results of the primary analysis.

6.2 COVID analysis set

The COVID analysis set will include all subjects from the FAS excluding subjects who ended the treatment due to COVID-19 before the first post-baseline efficacy assessment. [redacted]

6.3 Safety set (Safety)

The safety set will consist of all subjects who received at least one dose of AFM13 and had at least one post-baseline safety assessment (where the statement that a subject had no AE on the AE eCRF constitutes a safety assessment).

The safety set will be the primary population for all safety related endpoints.

6.4 Per protocol set (PP)

The PP set will consist of all subjects in FAS who did not have any major protocol deviations. To support the results of the primary analysis on the FAS, the primary analysis will be repeated using PP Set.

6.5 Pharmacokinetic set (PK)

The pharmacokinetic set (PK) set consists of all subjects who have at least received one dose of study drug and have at least one post dose PK measurement.

6.6 Protocol deviations and exclusions from analysis sets

Key (defined as deviation that may impact the safety or integrity of the study) and non-key protocol deviations to be collected during the study are listed and described in the Protocol Deviation Criteria Determination document. This document indicates also among all the deviations which ones may be major (a major deviation

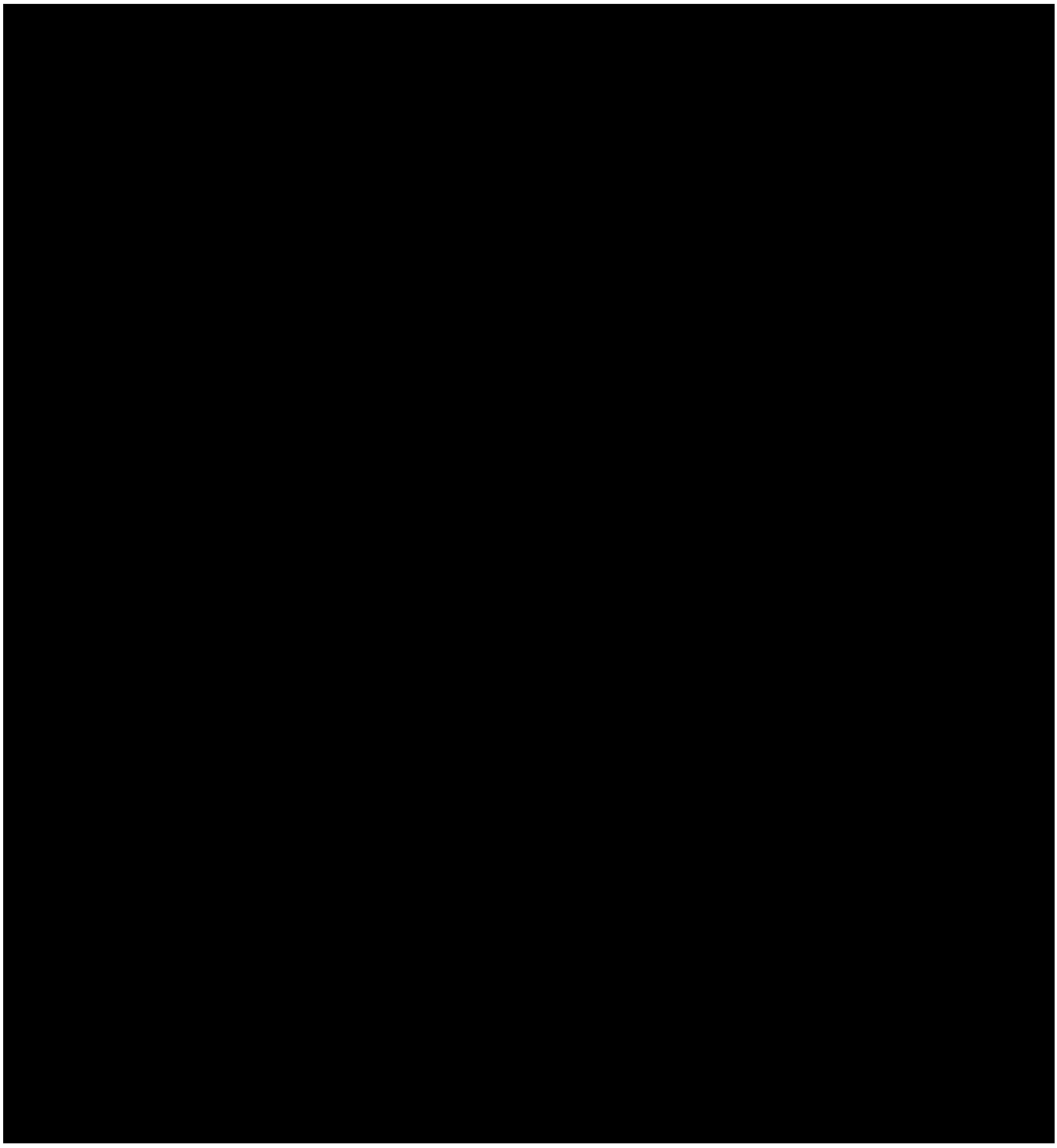
excludes a subject from the PP population). There is no relationship between key/non-key and major/minor deviations as a key deviation can be considered as minor for the statistical analyses.

All deviations will be reviewed at the Classification Meeting through clinical review input provided by Sponsor and will be classified as major or minor. Then, exclusions of subjects from analysis sets will be identified. The following sources of information will be used to support the inclusion or exclusion of subjects in analysis sets:

- Supportive subject listings, provided by the ICON lead statistician ahead of the Classification Meeting, based on data recorded on the eCRF.
- Protocol Deviation Logs, provided by ICON Medical.

[redacted]

[redacted]



[redacted]

8 STATISTICAL METHODS

8.1 General statistical conventions

All statistical procedures will be completed using SAS version 9.4 or higher.

Unless otherwise stated, all statistical testing will be two-sided and will be performed using a significance (alpha) level of 0.05. Two sided 95% confidence intervals (CI) will be provided when relevant.

Continuous variables will be summarized using descriptive statistics, including number of subjects (n), mean, median, standard deviation (SD), minimum (Min) and maximum (Max).

For categorical variables, summaries will include counts of subjects and percentages. Percentages will be rounded to one decimal place.

For summary purposes, baseline will be defined as the last available pre-dose value.

[redacted]

All subject data, including those derived, will be presented in individual subject data listings. Unless otherwise stated, unscheduled visit results will be included in date/time chronological order, within subject listings only. All listings will be sorted by cohort, subject number, assessment date/time and visit. Patient's sex and age will be also stated on each listing. Unless otherwise stated, data listings will be based on FAS.

Subjects who were screened and have signed the informed consent form (ICF) but did not receive any treatment will be listed including reason for screen failure or not being treated and any SAE (this will be recorded in the medical history). These subjects will not be part of any summary table except for summarizing disposition.

8.2 Subject disposition

Subject disposition information will be summarized by cohort and overall. The following number and percent of subjects will be summarized:

- Subjects who are pre-screened.
- Subjects who are screened.
- Subjects who are screen failures.
- Subjects enrolled for the treatment period but not treated.
- Subjects included in the FAS.
- Subjects included in the FAS and have at least one post-baseline efficacy assessment.
- Subjects included in the COVID analysis set.
- Subjects included in the safety set.
- Subjects included in the per protocol set.
- Subjects included in the PK set.
- Subjects who discontinued the treatment phase.
    - Subjects impacted by COVID-19 who discontinued the treatment phase.
    - Subjects affected by COVID-19 who discontinued the treatment phase.
- Subjects who performed the final study visit.
- Subjects who remained in response at the end of the treatment.
- Subjects who did not remain in response at the end of the treatment.
- Subjects who completed the study.
- Subjects who discontinued the study.

The number of subjects in FAS will be used as the denominator for the percentage calculation. Subject disposition and subject classification in each analysis population will be listed.

8.3 Protocol deviations

Major deviations (as defined in section 6.6) will be summarized by cohort.

Protocol deviations identified on protocol deviation Logs from ICON Medical will be listed, including the classification of major and minor deviations. The inclusion/exclusion criteria violated at Pre-screening and at Screening visits will be also listed.

8.4 Demographics and baseline characteristics

8.4.1 Demographics

The following demographic parameters will be summarized descriptively or by using frequencies on FAS:

- Age.
- Sex.
- Race.
- Ethnicity.
- Female of childbearing potential.
- Height at baseline.
- Weight at baseline.
- Body mass index (BMI) at baseline.
- Eastern Cooperative Oncology Group (ECOG) performance status at baseline.
- Electrocardiogram (ECG) interpretation at baseline.
- Systolic blood pressure (mmHg) at baseline.
- Diastolic blood pressure (mmHg) at baseline.
- Heart rate (beats/min) at baseline.
- Respiratory rate (breaths/min) at baseline.
- Temperature (°C) at baseline.

8.4.2 Baseline and disease characteristics

The following baseline and disease characteristics parameters will be summarized descriptively or by using frequencies on FAS:

- Time since initial diagnosis (months).
- Cancer type.
- Cancer sub-type (only for PTCL subjects).
- CD30 expression level (only for PTCL subjects) (using categories: ≥1% to <5%, ≥5% to <10%, ≥10% to <50%, ≥50%).
- Ann Arbor stage and Ann Arbor stage extension at initial diagnosis (only for PTCL subjects).
- Ann Arbor stage and Ann Arbor stage extension at Screening (only for PTCL subjects).
- TNMB staging at baseline (only for TMF subjects).
- Laboratory parameters:
    - Alanine transaminase (ALT).
    - Aspartate transaminase (AST).
    - Lactate dehydrogenase (LDH).
    - Albumin.
    - Total bilirubin.
    - Creatinine.

  - C-reactive protein (CRP).
  - Platelet count.
  - Hemoglobin.
  - White blood cells (WBC).
  - Absolute neutrophils.
  - Absolute monocytes.
  - Absolute lymphocytes.
  - Absolute eosinophils.
  - Absolute basophils.
- Modified Severity Weighted Assessment Tool (mSWAT) total score.
- Composite Assessment of Index Lesion Severity (CAILS) total score.
- Prior systemic therapies (yes/no).
- Prior cancer surgeries (yes/no).
- Prior radiation therapies (yes/no).

These parameters and some other baseline and disease characteristics parameters will be listed for the FAS.

8.4.3 Medical history

A summary of medical history will be presented by system organ class (SOC) and preferred term (PT) coded using Medical Dictionary for Regulatory Activities (MedDRA) version 22.0 or higher on the FAS.

Medical history findings will be also listed.

8.4.4 Prior and concomitant medications

Medications used in this study will be coded by using the latest available version of the World Health Organization Drug Dictionary Enhanced (WHODDE) (March 2019 or later).

Prior medications are defined as those medications with a start date prior to the first dose of study drug.

Concomitant medications are defined as those medications with a start date on or after the first dose of study drug.

Prior and concomitant medications will be summarized separately using frequency tables by ATC class (ATC Level 2) and preferred name by cohort on FAS.

A listing will be also provided for subjects' prior and concomitant medications.

8.4.5 Prior anti-cancer medications

Prior systemic therapies, prior cancer surgeries and prior radiation therapies will be summarized by cohort on FAS.

For prior systemic therapies, the following data will be summarized:

- Subjects with at least one prior systemic therapy.
- Number of treatment lines.
- Time between end of last treatment to study start (months).
- Medication type.
- Best overall response.
- Best response to Brentuximab vedotin (preferred term =’BRENTUXIMAB VEDOTIN’).
- Reason for discontinuing systemic therapy.
- Systemic therapies classified by WHODDE ATC class (ATC Level 2) and preferred name. Transplants and non-transplant therapies will be summarized separately.

Prior cancer surgeries will be summarized by MedDRA SOC and PT.

For prior radiation therapies, the following data will be summarized:

- Subjects with at least one prior radiation therapy.
- Number of prior radiation therapies.
- Time between end of last radiation therapy to study start (months).
- Type of radiation therapy.
- Intent.
- Settings.
- Best overall response to radiation therapy.

8.5 Exposure to study treatment

8.5.1 AFM13 study drug

The following parameters, relative to AFM13 drug administration, will be calculated and summarized by cohort using descriptive statistics for the FAS:

- Number of infusions of AFM13.
- Dose exposure (weeks).
- Cumulative dose (mg).
- Dose intensity (mg/week).
- Relative dose intensity (%).

The number of infusions with interruptions and dose modifications (including the reasons) will be also summarized by cohort. The number of subjects who take the 200 mg dose during the entire study will be also summarized (see Section 8.8.5).

Listing will be provided for AFM13 administration and exposure.

8.5.2 Premedications

The following medications taken before each AFM13 study drug administration will be listed and summarized:

- H1 antihistamine/H2 antagonist/Antihistamine.
- Analgesic/Fever reducer..
- Nonsteroidal anti-inflammatory drugs (NSAIDs).
- Glucocorticoid/Corticosteroid.
- Other.

NSAIDs and "other" medications are selected based on a list of ATCs and/or preferred names that Sponsor will provide. For - H1 antihistamine/H2 antagonist/Antihistamine, Analgesic/Fever reducer and Glucocorticoid/Corticosteroid categories, Sponsor will also provide the list of ATCs and/or preferred names but this will be also based on the study drug administration CRF page questions relative to antihistamine, acetaminophen and dexamethasone (or equivalent steroid) intakes.

These premedications will be summarized by visit. The medications collected during the study (in the prior and concomitant medications CRF pages) are not collected related to a specific visit. Consequently, the following rule will be used to associate the premedications to a visit:

> If premedication start date = Cycle X Day Y drug administration, the medication will be associated to the visit Cycle X Day Y. To be considered as a premedication, the indication of use should be "Other" or "Condition under study" or "Prophylaxis".
>
> For premedications which are recorded only once during the study in the CRF (recorded with frequency = Once weekly), if the medication intake dates include the study drug administration date, the medication the medication will be associated to the study drug administration visit.

To evaluate the relationship between dose intakes and pre-medications, an overall shift table will be generated between the following data to summarize the number of infusions received in each category (ie, number of infusions where the dose administered was > 0 mg):

- Infusion interrupted (Yes/No)

- H1 antihistamine/H2 antagonist/Antihistamine intake (Yes/No)
- Analgesic/Fever reducer intake (Yes/No)
- Glucocorticoid/Corticosteroid intake (Yes/No).

8.6 Efficacy analyses

This section addresses the analyses to be conducted on the primary, secondary and exploratory efficacy variables.

The FAS will be the primary population for all efficacy related endpoints and the primary objective. Sensitivity analyses on all subjects who received at least one dose of AFM13 and had at least one post-baseline efficacy assessment and on PP set will be conducted to support the results of the primary analysis, when deemed necessary (ie, if the number of subjects in these populations is different compared to the FAS). [redacted]

Each efficacy endpoint will be derived separately for IRC and Investigator. Consequently, endpoints using IRC assessments, will be derived using only responses provided by IRC regardless of Investigator responses. Similarly, endpoints using Investigator assessments will be derived using only responses provided by Investigator regardless of IRC responses.

8.6.1 Analysis methods

8.6.1.1 Multiplicity

Not applicable.

8.6.1.2 Treatment by center interaction analysis (multi-center study)

Not applicable.

8.6.1.3 Overall response date of assessment

If the date of assessment of overall response is not recorded/collected in the CRF then this will be assigned to the maximum of all assessments dates which belongs to the disease assessment visit.

8.6.1.4 Derivation of post-baseline disease assessments visits

In the CRF, the post-baseline disease assessments are not linked to study visits. Consequently, this will be performed by programming as per APPENDIX A.

8.6.2 Analysis of efficacy endpoints

The primary endpoint is based on the IRC PET-CT-based metabolic responses. Secondary endpoints are based

on IRC CT based radiologic responses and on Investigator (CT or FDG-PET (=PET-CT)) responses.

Overall response

Within each visit an overall response will be determined by the IRC and Investigator on CT-based radiographic responses and FDG-PET (=PET-CT) based metabolic responses separately utilizing the modified Lugano Classification revised staging system for malignant lymphoma (Cheson 2014) for Cohorts A and B (PTCL) and the Olsen Criteria (Olsen 2011) for Cohort C (TMF).

Subjects who cannot be evaluated for response will be categorized as not evaluable (NE).

Best overall response (BOR)

The best overall response is the best response recorded from the start of the treatment until disease progression/recurrence or death. All disease responses obtained after progressive disease, start of new anti-cancer therapy or a transplant are excluded from the best overall response. Best is defined as the first response when all a subject's response assessments are sorted by complete response/complete radiologic response/complete metabolic response (CR), partial response/partial remission/partial metabolic response (PR), stable disease/no metabolic response (SD), pseudoprogression, progressive disease/progressive metabolic disease (PD), not evaluable (NE).

The subject's best response assignment will depend on the achievement of the measurement criteria (confirmation is not required). BOR will be summarized using a frequency table by cohort. For Cohorts A and B, BOR is derived for radiographic (using only CT-based responses) and metabolic responses (using only FDG-PET (=PET-CT) based responses) separately.

Overall response rate (ORR)

ORR is calculated as the proportion of subjects with a CR or PR. Subjects without at least one post-baseline response assessment will be treated as non-responders.

ORR and Pearson-Clopper 95% confidence interval (CI) will be displayed separately for CT-based response and FDG-PET (=PET-CT) based response. CR and PR rates and their Pearson-Clopper 95% CI will be also displayed separately.

As the study design is based on a Simon's two stage design (See Section 5), an exact binomial test will be performed to analysed the ORR of the Cohorts A and C at the final analysis only. The one-sided p-value will be displayed. The tested hypotheses are as follows for each cohort:

Cohort A: H_0: $p \leq 0.25$

H_1: $p > 0.25$

Cohort C: H_0: p≤0.05

H_1: p>0.05

For all response assessments, swimmer plots will be presented. All response assessments will be also listed.

The above analyses (overall response, BOR and ORR) will be performed for each cohort and for:

Response evaluator	**Population of analysis**	**Analysis order**
IRC	FAS	Primary analysis
IRC	FAS with at least one post-baseline efficacy assessment	Sensitivity analysis, if deemed necessary (based on the number of subjects included in that population after the Classification Meeting)
IRC	COVID analysis set	Sensitivity analysis, if deemed necessary (based on the number of subjects included in that population after the Classification Meeting)
IRC	PP set	Sensitivity analysis, if deemed necessary (based on the number of subjects included in that population after the Classification Meeting)
Investigator	FAS	Secondary analysis
[redacted]	[redacted]	[redacted]

Theses analyses will be also repeated for the subgroups defined in Section 8.8.5 for the FAS.

Duration of response (DOR)

The DOR is defined as time from first assessment of PR or CR to the first assessment of progressive disease or death due to any cause separately for IRC and Investigator using respectively only IRC or Investigator assessments. DOR will be censored for IRC assessments using Table 1 and Investigator assessments using Table 3 when analyses are performed on FAS or PP set. The Table 2 and Table 4 will be used when the analyses are performed on the COVID analysis set:

Table 1: Independent Review Committee Assessments (IRC) Censoring Rules

Event	Event or censored date	Event or censored
Progression (as determined by IRC)	Date of progression (as determined by IRC)	Event
Death due to any cause	Date of death	Event
PD (as determined by IRC) or death after 1 missed assessment (ie, >= 59 days between 2 assessments until Cycle 4 and >= 87 days from Cycle 5)	Date of PD (as determined by IRC) or death	Event
PD (as determined by IRC) or death after 2 or more missed assessments (ie, >= 118 days between 2 assessments until Cycle 4 and >= 174 days from Cycle 5)	Date of last disease assessment before missed assessments	Censored
No baseline tumour assessment	Date of first study drug	Censored
No post-baseline response assessment	Date of first study drug	Censored
Start of new anti-cancer therapy before progression (as determined by IRC) or death	Date of last disease assessment prior to new anti-cancer therapy *	Censored
Transplant before progression (progression as determined by IRC) or death or start of new anti-cancer therapy	Date of last disease assessment prior to date of transplant	Censored
No progression determined by IRC or death	Date of last disease assessment	Censored

Note: Only disease assessments provided to IRC are included in the analysis.

* The new anti-cancer therapy information is only collected on the "Survival follow-up' page but no date is collected. Consequently, it is assumed that the new therapy starts between the follow-up where the question "Did the patient start any new anti-cancer therapies?" is "Yes" and the previous follow-up. The subject will be censored at the last disease assessment prior or on this previous visit.

[redacted]

* The new anti-cancer therapy information is only collected on the "Survival follow-up' page but no date is collected. Consequently, it is assumed that the new therapy starts between the follow-up where the question "Did the patient start any new anti-cancer therapies?" is "Yes" and the previous follow-up. The subject will be censored at the last disease assessment prior or on this previous visit.

Table 3: Investigator Assessments Censoring Rules

Event	**Event or censored date**	**Event or censored**
Progression (as determined by Investigator)	Date of progression (as determined by Investigator)	Event
Death due to any cause	Date of death	Event
PD (as determined by Investigator) or death after 1 missed assessment (ie, >= 59 days between 2 assessments until Cycle 4, >= 87 days from Cycle 5 and >= 122 days from Survival Follow-up)	Date of PD (as determined by Investigator) or death	Event
PD (as determined by Investigator) or death after 2 or more missed assessments (ie, >= 118 days between 2 assessments until Cycle 4, >= 174 days from Cycle 5 and >= 244 days from Survival Follow-up)	Date of last disease assessment or date of last survival contact (for Survival Follow-up) before missed assessments	Censored
No baseline tumour assessment	Date of first study drug	Censored
No post-baseline response assessment	Date of first study drug	Censored
Start of new anti-cancer therapy before progression (as determined by Investigator) or death	Date of last disease assessment or date of last survival contact (for Survival Follow-up) prior to new anti-cancer therapy *	Censored

Event	Event or censored date	Event or censored
Transplant before progression (progression as determined by Investigator) or death or start of new anti-cancer therapy	Date of last disease assessment or date of last survival contact (for Survival Follow-up) prior to date of transplant	Censored
No progression determined by Investigator or death	Date of last disease assessment or date of last survival contact (for Survival Follow-up)	Censored

Note: Only disease assessments recorded in the CRF are included in the analysis.

* The new anti-cancer therapy information is only collected on the "Survival follow-up' page but no date is collected. Consequently, it is assumed that the new therapy starts between the follow-up where the question "Did the patient start any new anti-cancer therapies?" is "Yes" and the previous follow-up. The subject will be censored at the last disease assessment prior or on this previous visit.

[redacted]

[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]

[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]

Note: Only disease assessments collected in the CRF are included in the analysis.

* The new anti-cancer therapy information is only collected on the "Survival follow-up' page but no date is collected. Consequently, it is assumed that the new therapy starts between the follow-up where the question "Did the patient start any new anti-cancer therapies?" is "Yes" and the previous follow-up. The subject will be censored at the last disease assessment prior or on this previous visit.

DOR will be summarized and listed by cohort and for CT and FDG-PET (=PET-CT) based responses separately. Duration of CR and duration of PR will be also reported separately.

DOR will be analyzed using Kaplan-Meier median, its 95% CI (and 25th and 75th percentile) and the estimates at 3, 6, 9, 12, etc months.

Theses analyses will be also repeated for the subgroups defined in Section 8.8.5 for the FAS.

8.6.3 Analysis of exploratory endpoints

[redacted]

[redacted]

8.7 Safety analyses

Safety analyses will be conducted on the safety set and will be performed for all safety variables specified below.

All the data will be summarized by cohort, and cohorts A and B combined.

No statistical test will be performed.

8.7.1 Adverse events

All Adverse events (AEs) will be classified by SOC and PT according to the MedDRA Version 22.0 or higher.

AEs will be graded according to the National Cancer Institute Common Terminology Criteria for Adverse Events (NCI CTCAE) version 5.0.

In summaries by SOC and/or PT, adverse events will be sorted by decreasing frequency within each SOC and/or PT according to the highest subject number in Cohort A.

AEs will be classified as pre-treatment AEs or treatment-emergent adverse events (TEAEs) defined as follows:

Pre-treatment AE: any AE that started after study enrollment (ICF signed) and before the first dose date of study medication.

TEAE: any AE that was newly developed at or after the first dose date of study drug.

Related AE: AE will be defined as related if causality is either 'definitely related', 'probably related', 'possibly related' or missing.

Details for imputing missing or partial start dates of adverse events are described in section 7.2.2.

AE summary tables listed below will be presented:

- All TEAEs.
- All related TEAEs.
- TEAEs by maximum severity.

- Serious TEAEs.
- Non-serious TEAEs.
- Related serious TEAEs.
- Related TEAEs by maximum severity.
- TEAEs with NCI CTCAE grades ≥3.
- Related TEAEs of NCI CTCAE grades ≥3.
- Important medical events (IMEs) (IMEs include only TEAEs):
    - Infusion-related reactions (IRRs).
    - Hepatic TEAEs (selecting events within "Drug related hepatic disorders - comprehensive search" Standardized MedDRA Queries (SMQ)).
    - Other IMEs (ie, SAEs with "Other important medical event" reason).
- Frequency of IRRs by maximum severity and visit.
- TEAEs leading to interruptions or modifications of study drug (action taken as "Drug reduced", "Drug interrupted", Drug delayed", "Treatment held").
- TEAEs leading to premature discontinuation of study drug (action taken as "Drug discontinued").
- AEs leading to death.
- Related AEs leading to death.
- Frequent non-serious TEAEs (SOCs or PTs with at least 5% of subjects within the cohort).
- Glucocorticoid/Corticosteroid premedications (Yes/No) and frequency of IRR by maximum severity and visit.

All AEs will be summarized by MedDRA SOC, PT and cohort using frequency counts and percentages (ie, number and percentage of subjects with an event). In addition, an overall summary for the categories above will be prepared by cohort.

Where a subject has the same adverse event, based on SOC and/or PT, reported multiple times, the subject will only be counted once at the SOC and/or PT level in adverse event frequency tables.

When reporting adverse events by severity, in addition to providing a summary table based on the event selection criteria detailed above, summary table will also be provided based on the most intense event regardless of the relationship to study treatment.

IRR will be summarized by visit and cohort. But IRRs are collected during the study without being linked to a specific visit. Consequently, the following rule will be used to assign the IRRs to a visit:

> If study drug administration at Cycle X Day Y ≤ IRR start date/time ≤ study drug administration at Cycle X Day Y + 24 hours then the visit is assigned to Cycle X Day Y; else visit is assigned to Unscheduled (Unscheduled visit should be also included in the "by visit" table).

8.7.2 Clinical laboratory evaluations

For the purposes of summarization in both the tables and listings, all laboratory values will be presented in standard units.

Laboratory results will be graded by NCI CTCAE version 5.0. For the tests without grading, only the classification according to normal ranges will be done.

Hematology (including coagulation), chemistry and urinalysis will be summarized by cohort, visit and parameters in each category.

If a lab value is reported using a nonnumeric qualifier eg, less than (<) a certain value, or greater than (>) a certain value, the given numeric value will be used in the summary statistics, ignoring the nonnumeric qualifier. This will be explained in the footnotes of the respective outputs.

Descriptive statistics for actual measurement and change from baseline at scheduled visits will be performed.

Urinalysis parameters will be also summarized by frequency counts by cohort and for each scheduled visit for nonnumeric parameters.

Shift tables for CTCAE grades and normal ranges from baseline to each visit and to worst post-baseline visit will be presented.

All laboratory tests will be listed. A separate listing for abnormal lab values (≥Grade 3, and low/high values) will be presented.

Liver and renal events

The subjects with liver events at any post-baseline visit will be summarized and listed for the following categories:

- ALT or AST > 3 × upper limit of normal (ULN).
- Alkaline phosphatase (ALP) > 2 × ULN.
- Total bilirubin > 2 × ULN.
- (ALT or AST > 3 × ULN) and total bilirubin > 2 × ULN and ALP > 2 × ULN.

- LDH > 2 x ULN.
- Albumin < lower limit of normal (LLN).

Glomerular filtration rate (GFR) in mL/min/1.73 m² will be summarized and listed at each visit and to worst post-baseline visit with the following categories: <15, [15-30[, [30-45[, [45-60[, [60-90[and ≥ 90.

GFR values will be derived using Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) equation:

$$GFR = 141 \times MIN\left(\frac{Creat}{K}, 1\right)^{\alpha} \times MAX\left(\frac{Creat}{K}, 1\right)^{-1.209} \times 0.993^{Age} \times$$

$$1.018\,[if\ Female] \times 1.159\,[if\ Black\ or\ African\ American]$$

where

Creat is serum creatinine in µmol/L

K is 61.9 for females and 79.6 for males

α is -0.329 for females and -0.411 for males.

[redacted]

8.7.4 Vital signs

Descriptive statistics for actual measurements of vital signs (systolic/diastolic blood pressure (mmHg), respiratory rate (breaths/min), heart rate (beat/min), temperature (◦C)), weight (kg) and BMI (kg/m²), and changes from baseline for each time point will be presented by cohort.

Listing of vital signs data will be presented.

8.7.5 Physical examinations

A listing will be provided for physical examination by cohort, subject, visit, body system, and description of abnormalities.

8.7.6 Electrocardiograms

Shift tables from baseline (Normal/Abnormal-clinically significant/Abnormal–not clinically significant) to each time point and to worst post-baseline ECG interpretation will be summarized for ECG interpretation data. Descriptive statistics for actual and changes from baseline will be displayed for heart rate (msec), QRS interval (msec), RR interval (msec), PR interval (msec), QT internal (msec), QTc internal (msec).

ECG data will be listed overall and a separate listing for any clinically significant finding in ECG values will be provided.

8.7.7 ECOG performance status

Shift table for performance status (0 to 5 grades) of subjects from baseline to the worst post-baseline visit will be presented by cohort.

A listing of the performance status of subjects will be provided by visit.

8.8 Other analysis

8.8.1 Statistical analysis for PK data

AFM13 serum concentrations will be summarized by cohort, study cycle, and nominal time points. All table summaries and figures for PK concentrations will be compiled using the PK set and will be presented by cohort. For concentration values below the limit of quantification (BLQ), a concentration value of zero will be included for the computation of arithmetic mean and a concentration value of ½ LLOQ will be included for the computation of geometric mean.

Individual subject listings for concentration data will be provided. Plots of serum concentrations (geometric mean (+/- coefficient of variation (CV)) will be presented in both original and semi-logarithmic scales.

PK parameters data will be converted into Study Data Tabulation Model (SDTM) datasets and will be summarized in a separate noncompartmental analysis (NCA) report.

[redacted]

8.8.3 Quality of life analyses

EQ-5D

EQ-5D will be used to assess the subject's 5 dimensions - mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Shift tables for EQ-5D pain/discomfort score ("no pain or discomfort", "moderate pain or discomfort" and "extreme pain and discomfort") from baseline to each visit will be presented for cohorts A and B by visit using FAS. The health state scale actual and change form baseline values will be also summarized.

Skindex-29

The Skindex-29 is composed of 29 items investigating the three domains: Emotions, Symptoms and Functioning. These domains will be derived for each visit as below:

Domains and clusters:

Domains/ Scale	Number of items	Clusters of items	Direction of domains
Emotions	10	3; 6; 9; 12; 13; 15; 21; 23; 26; 28	Higher score = higher impact of skin disease
Symptoms	7	1; 7; 10; 16; 19; 24; 27	
Functioning	12	2; 4; 5; 8; 11; 14; 17; 20; 22; 25; 29; 30	

Scoring procedure of each domain:

1. All responses are transformed to a linear scale of 100.
   Never = 0
   Rarely = 25
   Sometimes = 50
   Often = 75
   All the time = 100

2. A scale score is the mean of non-missing items in a given scale (note: item 18 is a single item, not included in scoring):
   a. If responses to more than 25% of items are missing overall the questionnaire is eliminated (ie, if > 7 items are missing among the 29 items).
   b. If any scale has more than 25% of the responses missing, the scale is missing (ie, if > 2 items for Emotions or > 1 item for Symptoms or > 3 items for Functioning are missing).

c. An item with multiple answers is considered missing.

Sum score:

The sum score is derived as the average of the three scale scores for each subject and visit.

The summary statistics of absolute and change from baseline values for the 3 scale scores and the sum score will be provided at each visit for cohort C using FAS.

Supportive by-subject listings will also be provided for EQ-5D-3L and Skindex-29.

8.8.4 Biomarker analysis

[redacted]

8.8.5 Subgroup analysis

The following subgroup analyses will be performed for the final analysis:

[redacted]

- [redacted]

[redacted]

8.9 Interim analysis

[redacted]

The following analyses will be generated for the interim analysis:

- Disposition.
- Demographic and baseline characteristics.
- Disease characteristics.
- Exposure to treatment (AFM13 and premedications)
- Overall response.
- Duration of response.
- Adverse events.

9 [redacted]

[redacted]

[redacted]

10 REFERENCES

1. ICH Topic E3: Structure and Content of Clinical Study Reports (CPMP/ICH/137/95- adopted December 1995).
2. ICH Topic E9: Statistical Principles for Clinical Trials (CPMP/ICH/363/96 – adopted March 1998).
3. ICH Topic E9(R1): Statistical Principles for Clinical Trials: Addendum: Estimands and Sensitivity Analysis in Clinical Trials (16 June 2017).
4. Cheson BD, Bennett JM, Grever M, et al. National Cancer Institute-sponsored Working Group guidelines for chronic lymphocytic leukemia: Revised guidelines for diagnosis and treatment. Blood 1996;87:4990-4997.
5. Cheson B, Fisher R, Barrington S, et al. Recommendations for Initial Evaluation, Staging, and Response Assessment of Hodgkin and Non-Hodgkin Lymphoma: The Lugano Classification. J. Clin. Oncol. 2014;32(27):3059-3068.
6. Olsen E, Whittaker S, Kim Y, et al. Clinical end points and response criteria in mycosis fungoides and Sézary syndrome: a consensus statement of the International Society for Cutaneous Lymphomas, the United States Cutaneous Lymphoma Consortium, and the Cutaneous Lymphoma Task Force of the European Organisation for Research and Treatment of Cancer. J. Clin Oncol.2011;29(18):2598-2607
7. Simon, R. Optimal Two-Stage Designs for Phase II Clinical Trials. Controlled Clin. Trials.1989;10(1):1-10.
8. Skindex Scaling and scoring, Version 2.0, February 2012, MAPI Research Trust.
9. Protocol Deviation Criteria Determination document: Document type “Protocol Deviation Criteria Template” (document type ID: 103.11) of the eTMF.

11 APPENDIX A

For each disease assessment visit, the visit name will be derived as follows:

Step 1: Determine the minimum and maximum dates per visit as follows:

For Lugano (within the same disease assessment visit):

DateINF = MIN (
TU.TUDTC (Imaging (modified Lugano)),
TR.TRDTC (Imaging (modified Lugano)),
TU.TUDTC (New Lesions (modified Lugano)),
TR.TRDTC (New Lesions (modified Lugano)),
TU.TUDTC (Tissue Sites Involved),
TR.TRDTC (Tissue Sites Involved),
TU.TUDTC (FDG-PET),
TR.TRDTC (FDG-PET),
PR.PRDTC (FDG-PET))

DateSUP = MAX (
TU.TUDTC (Imaging (modified Lugano)),
TR.TRDTC (Imaging (modified Lugano)),
TU.TUDTC (New Lesions (modified Lugano)),
TR.TRDTC (New Lesions (modified Lugano)),
TU.TUDTC (Tissue Sites Involved),
TR.TRDTC (Tissue Sites Involved),
TU.TUDTC (FDG-PET),
TR.TRDTC (FDG-PET),
PR.PRDTC (FDG-PET))

For Olsen (within the same disease assessment visit):

DateINF = MIN (
TU.TUDTC [Imaging (Olsen)],
TR.TRDTC [Imaging (Olsen)],
TU.TUDTC (New Lesions (Olsen)),
TR.TRDTC (New Lesions (Olsen)),
FA.FADTC (TNMB Staging),
QS.QSDTC (mSWAT),
QS.QSDTC (CAILS),
LB.LBDTC (Flow cytometry for Olsen criteria),
PR.PRSTDTC (Lesion Photography),
PR.PRSTDTC (Biospy (Olsen criteria),
MI.MIDTC (Biospy (Olsen criteria))

DateSUP = MAX (
TU.TUDTC [Imaging (Olsen)],
TR.TRDTC [Imaging (Olsen)],
TU.TUDTC (New Lesions (Olsen)),
TR.TRDTC (New Lesions (Olsen)),
FA.FADTC (TNMB Staging),
QS.QSDTC (mSWAT),
QS.QSDTC (CAILS),
LB.LBDTC (Flow cytometry for Olsen criteria),
PR.PRSTDTC (Lesion Photography),
PR.PRSTDTC (Biospy (Olsen criteria),
MI.MIDTC (Biospy (Olsen criteria))

Step 2: Assign the visit name as follows:

If (SV.SVSTDTC – 7 <= DateINF <= SV.SVSTDTC + 7)
or (SV.SVSTDTC – 7 <= DateSUP <= SV.SVSTDTC + 7)
or (DateINF < SV.SVSTDTC – 7 AND DateSUP > SV.SVSTDTC + 7) then AVISIT = SV.VISIT;
Else AVISIT = "Unscheduled".

Notes:

- Visits planned per protocol to have disease assessments are considered from SV.
- If at least one assessment of the disease assessment visit is within the time-window then all the assessments are assigned to the visit.