Statistical Analysis Plan

Version 1.0, 04-July-2022

(Based on Study Protocol version 2.0 of 28 June 2022)

For

SAFETY STUDY OF ACOZIBOROLE IN g-HAT SEROPOSITIVE NON-PARASITOLOGICALLY CONFIRMED SUBJECTS

A multicenter unbalanced randomized double blind versus placebo-controlled study

Study Code:
DNDi-OXA-04-HAT

Sponsor:

DND*i*, Chemin Camille-Vidart, 15, CH-1202 GENEVA Switzerland

Phone : +41 22 906 9260

Prepared by (Trial Statistician):	
Name:	
Role:	Statistician
Signature: 5108D35D4D38434...	**Date:** 27/7/2022 | 05:05 PDT
Reviewed by (Clinical Trial Manager):	
Name:	
Role:	Clinical Program Officer
Signature:	**Date:** 3/8/2022 | 01:04 PDT
Approved by (Head of Disease Program):	
Name:	
Role:	Head of HAT Clinical Program
Signature:	**Date:** 4/8/2022 | 07:32 PDT

Version History

Version#	Version Date	Affected section(s)	Changes Summary
1.0	04-July-2022	First approved version	NA

Table of contents

Abbreviations

ABS	Absolute
AE	Adverse Event
ATC	The Anatomical Therapeutic Chemicals
BMI	Body Mass Index
CATT	Card agglutination test for trypanosomiasis
CI	Confidence Interval
CMH	Cochran-Mantel-Haenszel
CONSORT	Consolidated Standards of Reporting Trials
CSR	Clinical Study Report
DBS	Dry Blood Spot
DNDi	Drug for Neglected Diseases *initiative*
MedDRA	Medical Dictionary for Regulatory Activities
DRC	Democratic Republic of Congo
ECG	Electrocardiogram
ER	Excess Rate
HAT	Human African Trypanosomiasis
ICF	Informed Consent Form
ICH	The International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use
IHC	Immunohistochemistry
ITT	Intention-to-Treat
NNT	Number Needed to Treat
NSSCP	National Sleeping Sickness Control Programmes
OR	Odds Ratio
PCR	Polymerase chain reaction
PD	Pharmacodynamic
PK	Pharmacokinetic
PT	Preferred Term
Q1	Lower Quartile
Q3	Upper Quartile
QQ	Quantile-Quantile
RT-qPCR	Quantitative Reverse Transcription PCR
QTc	Corrected QT interval
RDT	Rapid Diagnostic Test
RR	Risk Ratio
SAP	Statistical Analysis Plan
SAS	Statistical Analysis System
SHERLOCK	Specific High-sensitivity Enzymatic Reporter unLOCKing
SOC	System Organ Class

TEAE	Treatment Emergent Adverse Event
TLF	Tables Listings and Figures
WHO	World Health Organization
WMW	Wilcoxon-Mann-Whitney

1. Introduction

This Statistical Analysis Plan (SAP) describes the statistical analysis methods and data presentations for the DNDi-OXA-04-HAT study. This SAP is based on the Study Protocol version 1.0, dated 19 July 2021. International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH): *Guidance on Statistical Principles in Clinical Trials*[1] and the ICH E3 Guideline: *Structure and Content of Clinical Study Reports*[2] were used as a guide for the preparation of this SAP. Additional analysis not described in the SAP but deemed necessary will be clearly described as a post-hoc analysis in the Clinical Study Report (CSR).

1.1. Background and rationale

Human African trypanosomiasis (HAT) is a neglected tropical disease, transmitted by the bite of a tsetse fly, affecting sub-Saharan African countries. Without prompt diagnosis and treatment, it is usually fatal. Few treatments are available and, although the treatment options have advanced with the recent approval of fexinidazole, a ten-days oral treatment for the parasitologically confirmed g-HAT cases, there are still gaps with regards to management of the challenges of integrated treatment not only of confirmed but also the suspects individuals when foreseeing the 2030 Public Health objective of sustained HAT elimination.

Whilst the delivery of fexinidazole has improved the management of g-HAT cases and facilitated the integration of HAT treatment into the general health system, it is expected that the current investment in acoziborole will help boost such integration and the elimination efforts.

In the single ascending dose study in healthy volunteers (randomized, double-blind, placebo-controlled study), acoziborole was well tolerated and the following pharmacological parameters were observed[3]:

- Tmax around 72h, stable for at least 4 days
- Long half live (T1/2 of about 400h) allowing single administration
- Low metabolite detection
- Low inter-subject variability
- Unbound fraction around 2.2%
- Adverse events observed were not dose dependent

Acoziborole as an oral, single-dose treatment envisioned for all stages of g-HAT, was studied in an open-label pivotal Phase II/III trial (DNDi-OXA-02-HAT) in DRC and Guinea. The safety and efficacy results from the pivotal study have provided data that envision the treatment of parasitologically confirmed g-HAT cases.

The standard g-HAT case definition implies the demonstration of the parasite in any body fluid via microscopy. However, there are factors such as low parasitemia and the complexity and low sensitivity of parasitological methods that make such demonstration difficult. It is known that

there are some "indirect" signs of g-HAT that could indicate the presence of the parasite. The most known signs are biological (positive antibody response to either available serological tests CATT and RDT, or to molecular tests) and clinical. In some circumstances it has been suggested (WHO Technical Report Series 984) to consider subjects presenting "indirect" signs but unable to demonstrate the parasite as g-HAT "suspected" individuals as cases "eligible" for treatment. Only in special conditions, "suspected" g-HAT subjects have been treated[4–6] since the complexity and toxicity of available treatment options typically precludes systematic treatment of these "suspects" due to an individual negative risk-benefit ratio. It has, however, been demonstrated that a variable proportion (mainly depending on the prevalence) of such "suspects" are real cases and, therefore, potential reservoirs of the parasite and a source of new infections hindering the efforts to eliminate the disease [7].

Given the encouraging safety and efficacy signs observed during the pivotal study conducted with acoziborole, DNDi opened a debate on 24th January 2019 in the frame of the ninth meeting of the sub-group "Integration of new tools into national and global policies" of the WHO g-HAT elimination network in order to define the hypothetical "preventive" use of acoziborole in the population of "suspected" g-HAT subjects. The meeting concluded that for such expanded use, the need to plan an additional trial to enlarge the safety data to complement the safety profile obtained in the pivotal trial to allow a better understanding on the risk benefit to apply treatment of acoziborole in these population of g-HAT suspected but non parasitologically confirmed subjects.

In addition, a sub-study *'TrypSkin'* is planned to assess the presence of dermal *T. b. gambiense* in this population. In an observational cohort study recently conducted in Guinea, the presence of extravascular dermal trypanosomes has been observed in individuals presenting Card Agglutination Test for Trypanosomiasis (CATT) seropositive in plasma dilution ≥ ¼ but non-confirmed by parasitological examination of blood and/or lymph[8]. If these dermal trypanosomes correspond to *T. b. gambiense* subspecies and are able to infect vectors, these human carriers could act as reservoirs to the transmission of g-HAT, hampering the sustained elimination goal.

1.2. Trial objectives and endpoints

1.2.1. Trial objectives

1.1.1.1. Primary objective

- To assess the safety and tolerability of a single dose of acoziborole compared with placebo during a follow-up period of 4 months in seropositive individuals who are not confirmed parasitologically.

1.1.1.2. Secondary objectives

- To assess safety of acoziborole versus placebo in:

  - Haemato-biochemistry
  - Electrocardiogram (ECG) data at Day 5 (double delta versus baseline and time, age and sex matched comparison)
- To assess blood PK at Day 5 and Month 1
- To assess correlation between ΔQTc measurements and acoziborole concentrations in blood at Day 5

1.1.1.3. Exploratory objectives (TrypSkin sub-study)

- To estimate the prevalence of extravascular dermal *T.b. gambiense* in seropositive individuals non-confirmed parasitologically.
- To compare the properties (relative sensitivity, relative specificity, positive predictive value, relative negative predictive value, and repeatability) of different methods of detection of *T. brucei s.l. and T.b. gambiense* (polymerase chain reactions [PCRs], quantitative PCRs [qPCRs], multiplex quantitative Reverse Transcription PCRs [RT-qPCRs], Specific High-sensitivity Enzymatic Reporter unlocking [SHERLOCK], immuno-histochemistry [IHC]) in skin biopsies, blood samples and DBS with gold standard PCR.
- To determine the presence/absence of extravascular *T.b. gambiense* in skin biopsies of all subjects at enrolment, and at the end of the study.
- To assess the effectiveness of acoziborole on elimination of extravascular *T.b. gambiense* parasites in dermis with the methods or combination of methods, contingent on having better validity than *Trypanosoma brucei gambiense*-specific glycoprotein PCR (TgSGP-PCR) and provided that the prevalence of dermal trypanosomes by any method at baseline is greater than 2%.

1.2.2. Trial endpoints

1.1.1.4. Primary endpoint

- Occurrence of Treatment Emergent Adverse Events (TEAEs)

1.1.1.5. Secondary endpoints

- Occurrence of adverse events (AEs) from ICF signature to end of study (EoS)
- Change from baseline in ECG, biochemistry, haematology
- Blood concentration of acoziborole at D5 and M1
- Placebo-corrected change from baseline in ECG parameters (double delta [$\Delta\Delta$] heart rate [HR], $\Delta\Delta$RR, $\Delta\Delta$PR, $\Delta\Delta$QRS, $\Delta\Delta$QT and $\Delta\Delta$ QTc)
- ΔQTc measurements at baseline and Day 5
- Incidence of abnormal values and morphological findings in ECG at D5

1.1.1.6. Exploratory endpoints

Primary exploratory endpoints:

- Occurrence of extravascular dermal *T.b. gambiense* at enrolment and at 4 months by TgSGP-PCR in skin.

Secondary exploratory endpoints:

- Clinical parameters: occurrence of dermatitis and/or pruritus.
- Diagnostic parameters: occurrence of IHC+, TBR-PCR+ and/or 18S-PCR+ in skin, blood, and DBS.
- Novel diagnostic parameters: occurrence of positive results by multiplex qPCRs, multiplex RT-qPCRs and SHERLOCK in skin, blood, and DBS.
- Occurrence of dermatitis, pruritus, IHC+, multiplex qPCRs+, multiplex RT-qPCRs+ and SHERLOCK+ in skin, blood and DBS at enrolment and 4 months.
- Occurrence comparisons between all diagnostic methods alone or in combination.

2. Trial methods

2.1. Trial design

The trial is a randomized, multicenter, double-blind, placebo-controlled, parallel-arm phase II/III study to assess the safety and tolerability of a single dose of acoziborole compared with placebo in adolescent and adult subjects seropositive to human African trypanosomiasis due to *T.b. gambiense* (g-HAT positive; not confirmed parasitologically [no demonstration of the parasite in any body fluid via microscopy]) in DRC and Guinea.

Subjects will be hospitalized at least 1 day before study treatment commences and will be randomized in a 3:1 ratio to acoziborole or matched placebo on Day 1 of the study. Following hospitalization, subjects will attend follow-up visits at the study center at 1- and 4-months post-treatment. At M2 and M3, investigator contacts will be established between the site and study subjects to ensure close follow-up of subjects and encourage study adherence. An onsite unscheduled visit will be proposed if deemed necessary by the investigator. Any subject with a positive parasitological test result during the study will end their participation in the study and referred to the National Sleeping Sickness Control Programmes (NSSCP) for appropriate standard treatment.

In addition, a sub-study 'TrypSkin' to assess the presence of extravascular dermal *T. b. gambiense* in this population, will be proposed in clinical sites participating in the sub-study to subjects having given their consent to participate in the seropositive safety study. Subjects can decide whether or not to participate in the sub-study independently of his/her participation in the main safety study.

2.2. Randomization

Unbalanced randomization of 3:1 will be performed by blocks of size 16 to keep the allocation ratio as close as possible to the planned allocation ratio, and to get the number of randomized subjects exposed to acoziborole as close to 900 as possible.

2.3. Schedule of events

The schedule of events is provided in Table 2.3.1 and for more details refer to sections on Study Assessment and Study Schedule of the protocol.

Table 2.3.1: Study Schedule of Events

	Hospitalisation							**Follow-up**	
	Screening (before dosing)	**Dosing**	**Post dosing period**					**Site visit**	
	D-2 to D1H0	**D1H0**	**D1**	**D2**	**D3**	**D4**	**D5**	**M1**	**M4 (EoS)**
Informed consent (Main study + TrypSkin sub-study[1])	X								
Serological test (CATT or HAT sero-K-set RDT)	X								X[2]
Parasitological test(s)[3]	X							X	X
Pregnancy test[4]	X							X	X
Demographics	X								
Medical History	X								
Concomitant medications record	X	X	X	X	X	X	X	X	X
Karnofsky index	X							X	X
Clinical examination (including vital signs)	X						X	X	X
Dermatological examination	X							X	X
Blood pressure	X		X[5]				X	X	X
Haematology /Biochemistry	X						X	X	X
Inclusion/Exclusion review	X								
Randomisation/Treatment		X							
ECG[6]	X						X		
Trypanolysis test[7]	X								
Blood PK (on DBS)							X	X	
Collection of AEs	X	X	X	X	X	X	X	X	X
TrypSkin sub-study procedures									
Blood sampling for molecular biology	X								X
2 mm-skin-punch biopsies for IHC and molecular biology	X								X

AE=adverse event; CATT=card agglutination test for trypanosomiasis; D=day; DBS=Dry Blood Spot; ECG=electrocardiogram; EoS=End of Study; H= Hour; IHC=immunohistochemistry; M=month; PK=Pharmacokinetics; RDT=rapid diagnostic test.

1. For sites participating in the TrypSkin sub-study only
2. M4, same serological test than baseline to be repeated
3. mAECT or mAECT-BC and lymph node puncture if suspicious enlarged cervical lymph nodes are detected.
4. For women of childbearing potential only.
5. At D1, blood pressure will be taken 1 hour before dosing, at D1H4 and at D1H9
6. ECG triplicate recording at pre-dose on D1 and D5 according to acceptable time window described in section **Error! Reference source not found.** .
7. Not for inclusion but as a reference test for follow-up.

2.4. Sample size and ratio determination

2.4.1. Sample size for the main study

The sample size was determined by calculation based on statistical probabilities to detect uncommon TEAEs in active group. The definition for uncommon event is a relative frequency between 0.1% and 1%. The median relative frequency is therefore 5/1000 (0.5%). The probability to detect at least 1 event if the true incidence rate is 0.5% is as high than 98% with 900 subjects and the probability is 83.5% for a true incidence rate of 2/1000 (0.2%). According to the calculations, 900 subjects exposed to acoziborole shows a high probability (>99%) of detecting at least one event (if the incidence rate is 1%) with a similar probability when exposing 1200 subjects. Therefore, 900 subjects will be administered with acoziborole and 300 subjects with placebo, for an overall sample size of 1200 subjects. The calculations were performed using StatXact 9 software and presented in Table 2.4.1.

A subject declared as lost to follow-up will be replaced in order to maintain the planned sample size. From 5 months after the start of the study, the rate of subjects lost to follow-up will be regularly assessed and recruitment will be adjusted accordingly to ensure 1200 subjects will complete the 4 months of follow-up.

Table 2.4.1.1: Statistical probabilities for detecting uncommon treatment emergent adverse events (TEAEs) in the acoziborole group

Objectives/Sample size	**N=900**	**N=1200**
Chance of detecting at least one event if the incidence rate is 1%	99.9%	99.99%
Chance of detecting at least one event if the incidence rate is 0.1%	59.36%	68.89%
Precision: 95% exact CI if one event is observed	0.003 – 0.6%	0.002 – 0.46%
Precision: 95% exact CI if three events are observed	0.068-0.97%	0.052-0.73%

CI=confidence interval; TEAEs=treatment-emergent adverse events
Uncommon TEAEs are those with frequency ≥1/1000 and <1/100 as per regulatory guidelines

2.4.2. Sample size for the TrypSkin study

For TrypSkin sub-study, the primary objective is to estimate the prevalence of extravascular dermal *T.b. gambiense* trypanosomes in a population of seropositive subjects non-confirmed parasitologically. Preliminary results from Guinea, on a limited number of subjects, have shown that all seropositive individuals presenting a positive CATT test in plasma dilution ≥1:4, but non-confirmed by parasitological examination of blood and lymph, had parasites in skin biopsies[8]. Knowing that approximately 25% of the RDT positive subjects non-confirmed parasitologically

present a positive CATT test in plasma dilution ≥1:4, it is hypothesized that up to 20-25% of the population of RDT-positive subjects non-confirmed parasitologically enrolled in the TrypSkin study could possibly carry dermal extravascular parasites. Bearing in mind this higher limit of the skin-dwelling trypanosome prevalence, we have estimated the minimum sample size required to be able to evaluate the real prevalence with a satisfactory precision (a narrow CI). Under the assumption that skin-dwelling trypanosomes could be observed in up to 20-25% of the 1,200 subjects, to obtain a 95% confidence interval with a precision of 3%, a minimum of 690 evaluable subjects is required. A provision of 5% (35 subjects) is planned in case of skin samplings would be non-analyzable. Overall, the number of subjects to be enrolled in the TrypSkin ancillary study is therefore 725. The determined sample size is guided by the expected precision as presented in Table 2.4.2.1.

Table 2.4.2.1: Sample sizes and corresponding precision for TrypSkin exploratory study

	Estimated Prevalence					
Precision	**0.05**	**0.10**	**0.15**	**0.20**	**0.25**	**0.30**
0.025	300	550	780	980	1160	1300
0.03**Error! Reference source not found.**	200	390	540	**690**	800	900
0.05	80	130	200	245	290	320

CI=confidence interval

[a] Width of 95% CI approximately 0.006 (6%)

2.4.3. Ratio determination

Operational properties of various allocation ratios for the following assumptions are as follows:

1. Total sample size (for safety main study): n=1200, (900 acoziborole and 300 placebo)
2. Excess rate of 1% (acoziborole = 2% of AEs of interest, Placebo = 1%) for safety purposes
3. Probability to detect an infrequent AE (≤ 1% upper limit of class)
4. Probability to detect an infrequent AE (0.5% mid-class)
5. Total sample size (for efficacy – TrypSkin exploratory sub-study): n=700
6. Binary endpoint for efficacy: Success = no tryps detected in punch skin at month 4 (failure = at least one tryp-positive result in punch skin at Month 4)
7. Probability to detect at least one trypanosome in derma at baseline: 20%
8. No effect of placebo on occurrence of tryps in derma at month 4: failure rate at 4 months = 20%
9. Effect of 95% of acoziborole on occurrence of tryps in derma at Month 4 (efficacy): Failure rate at Month 4 = 0.05 × 0.20 = 0.01 or 1%.
10. Excess rate of success in acoziborole group = 19% assuming the same efficacy as stage 2 patients**.**

Table 2.4.3.1 presents the summary of the results for different assumptions on precision and power for the study to determine the required ratio.

Table 2.4.3.1: Summary on ratio determination for different assumptions

	Safety			**Efficacy (for TrypSkin sub-study)**		
Allocation ratio	Sample size for safety	Exact 95% Confidence interval limits of excess rate	Probability to detect an unfrequent AE if occurrence rate = 1% (upper limit), = 0.5% (mid-class) and = 0.1% (lower limit of class)	Sample size for efficacy	Exact 95% CI limits of excess rate Estimate = 19%	Power (chance to detect a difference vs pbo) if acoziborole efficacy is 95%
3:1	PBO = 300 ACO = 900	Lower = -0.92% Upper = +2.41%	Upper: 99.99% Mid: 99.90% Lower: 59.36%	PBO = 175 ACO = 525	Lower = 13.38% Upper = 25.55%	99.99%
2:1	PBO = 400 ACO = 800	Lower = -0.64 Upper = +2.42%	Upper: 99.96% Mid: 98.18% Lower: 55.06%	PBO = 233 ACO = 467	Lower =14.02% Upper = 24.62%	99.99%
3:2	PBO = 480 ACO = 720	Lower = -2.30%* Upper = +0.66%*	Upper: 99.93% Mid: 97.29% Lower: 51.34%	PBO = 280 ACO = 420	Lower = 14.51% Upper = 24.15%	99.99%
1:1	PBO = 600 ACO = 600	Lower = -0.44% Upper = +2.51%	Upper: 99.75% Mid: 95.06% Lower: 45.23%	PBO = 350 ACO = 350	Lower: 14.81% Upper: 23.48%	99.99%

**Not comparable with other allocation ratio due to decimal digits in the number of subjects requiring rounding to the closest entire number*

Conclusions:

- The four allocation ratios allow detection, with sufficient probability, of an infrequent AE except for the lower limit of the class (0.1%) of infrequent events. In that case, the larger the sample in the acoziborole group the better the probability of detection.
- The upper limit of the exact 95% CI of the excess rate of AEs is similar irrespective of the allocation ratio.
- The lower limit of the exact 95% CI is closer to the estimate (shorter 95% CI) with a balanced (1:1) allocation ratio.
- The power of a test of success rate (success = no trypanosome observed in skin punch sample at Month 4) is excellent irrespective of the allocation ratio if the effect of acoziborole is similar to the efficacy in stage 2 (efficacy ≥95%).
- The advantage of an allocation of 3:1 is to maximize the number of subjects seropositive subjects exposed to acoziborole. There is a small disadvantage in the exact lower limit of the 95% CI (larger CI), but no real advantage in the power of the efficacy test.
- The main disadvantage of the balanced allocation ratio (1:1) is a smaller number of exposed subjects

2.5. Interim Analyses

For the TrypSkin sub-study, an interim analysis for futility will be performed once baseline data of 50% of subjects will be collected. If the occurrence rate of positivity by any diagnostic method

in any sample at baseline lead to less than 2%, then the recruitment in the TrypSkin sub-study will be stopped for futility because there would be no possibility to assess the effectiveness of Acoziborole in elimination of extravascular *T.b. gambiense* parasites in dermis.

2.6. Timing of Final Analysis

The main final analysis is planned when all subjects have concluded 4 months of follow up, all data up to 4 months have been entered, verified, and validated and the primary database has been locked.

2.7. Timing of Outcome Assessments

2.7.1. Definition of baseline

The baseline value for analysis will be the last value collected during screening period, prior to randomization. If multiple validated measurements occur at the same time or time is not available, the least favorable among these measurements will be considered as the baseline value.

2.7.2. Analysis Windows

For all clinically planned assessments, visits should occur within the window of the scheduled visit. Visits outside visit window will be regarded as a minor protocol deviation. The target day and visits window are defined from first day (D1), corresponding to the day of study drug administration. For analysis and tabulation purposes, the study time points are defined as in Table 2.7.2.1. If more than one visit fall into the same time point interval, the visit with a relative day closest to the target scheduled visit will be selected for the analyses. The visit(s) not selected will only be listed and flagged as use=no in the analysis dataset. If there are multiple visits at the same distance of the scheduled visit day (meaning: equal ABS(DY –target day)), then the one latest in time is selected.

Table 2.7.2.1: Definition of Analysis Windows

Scheduled visits	Ideal visit day	Analysis Window*
End-of-hospitalisation (EoH) visit	D5	D5 - D8
1 month	1 months (D31)	D21 – D41
4 months	4 months (D121)	D111 - D151

* The analysis window for the visit starts on the first day of the period mentioned and ends on the last day of the period mentioned.

3. Statistical Principles

3.1. Confidence Intervals and p-values

All calculated p-values will be two-sided and compared to a 5% significance level. If a p-value is less than 0.05, the corresponding treatment group difference will be denoted as statistically significant. All estimates will be presented with two-sided 95% confidence intervals. No

adjustments of type one error for multiplicity testing are planned because the goal of the trial is to detect any potential safety issues instead of confirming and investigating a safety concern.

3.2. Treatment Emergent AE definition

The information on adverse events will be collected from the time of signing ICF. The treatment-emergent adverse events (TEAEs) will be defined as events occurring from administration of study drug or events starting before study drug administration but worsening after study drug administration to 4 months post-treatment.

3.3. Handling of AE missing dates or severity

Partial adverse event dates will be imputed with reference to dose intake date and follow up end date. If complete (imputed) AE end date is available and the imputed AE start date is greater than the (imputed) AE end date, then imputed AE start date should be set to the (imputed) AE end date. Completely missing AE start dates will be imputed with the dose intake date unless the end date suggests it could have started prior to the dose intake. For end date that is completely missing, if the AE has stopped and start date is prior to the dose intake then impute with the date of dose intake, if AE started on or after the dose intake then impute with the last follow up date. A conservative approach will be used in the missing date imputation for such AEs to be defined as treatment emergent. More details on the imputation procedure will be provided in the programming specification.

For the missing severity grade, the highest severity will be used though it is not expected to have AE severity missing.

3.4. Analysis Populations

The analysis population are described separately for subject disposition, main study analysis population for the primary objective and sub-study for exploratory objective analysis population.

3.4.1. Subject Disposition populations

Screened set includes all subjects who were screened and provided informed consent by signing the ICF.

Randomized set (ITT) includes subjects who signed the informed consent and were considered eligible by the investigator and randomized into the study, i.e., randomized to receive acoziborole or placebo which corresponds to intention-to-treat (ITT) set population.

Treated set includes all randomized subjects who received three tablets of IP (placebo or acoziborole).

3.4.2. Main study analysis populations

Safety set of subjects (SS) is composed of all randomized subjects who received at least one tablet of IP (placebo or acoziborole). This analysis population will be the primary set of subjects for the main study analyzed per the actual treatment received.

3.4.3. Sub-study analysis populations

The primary set of subjects for exploratory efficacy analyses is the *Safety set of subjects*, who participated in the TrypSkin sub-study.

The secondary set of subjects for exploratory efficacy analyses is the set of all seropositive subjects who were positive for the presence of trypanosome by at least one method in skin, blood and/or DBS at baseline.

4. Trial Population

4.1. Trial Subject Disposition

All subjects who provide informed consent will be accounted for in this study. Subject disposition and withdrawals will be presented for the screening analysis set. The total number of screened subjects, reasons for screening failure, number randomized, and reasons not randomized will be summarized and tabulated by site. A Consolidated Standards of Reporting Trials (CONSORT) flow diagram will be used to summarize the number of subjects who were screened, eligible/ineligible, randomized, lost to follow up, discontinued and those completed study.

The status of eligible and randomized subjects at trial end will be tabulated by treatment group according to study follow up status. Time from randomization to withdrawal/lost to follow-up will be presented graphically using the Kaplan-Meier estimator. A summary of the analysis population sets by treatment group and subjects excluded from analysis population will be provided. The data listing for subjects excluded from the population analysis sets will be provided including the reasons for exclusion. A summary of protocol deviations (PDs) will be provided including minor/major classification and PDs categorizations. The PD categorization and classification will be determined before the database lock and unblinding. PD individual data listing will also be provided.

4.2. Baseline Subject Characteristics

Subject demographic characteristics to be summarized will include gender, age in years by treatment group and overall, for the ITT set. The subject demographic will also be summarized for the Safety Set (SS) if there is a difference of at least 5% with ITT set. Subject demographics will be summarized using descriptive statistics (N, mean, standard deviation, median, 25^{th}/75^{th} percentiles, minimum, and maximum) for age variable, and number and percentages of subjects for gender and categorized age (adolescent 15-17 years and adults ≥18 years) variables. Subject demographic data listing will be provided.

4.3. Medical history

Medical history data will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) version 21.0 or higher. System organ classes (SOC) and preferred terms (PT) will be attached to the clinical database. The numbers and percentages of subjects in the safety set with a medical history will be summarized by SOC and PT and presented overall and by treatment group for the ITT set. Subjects will appear in the table only once per SOC and PT. The table will be sorted alphabetically by SOC and then by descending frequency of PT in the "Overall" column. A listing of all medical history data will be provided.

4.4. Previous and concomitant medications

Previous medications, defined as any medications taken prior to study treatment intake, and concomitant medications, defined as any medications other than study drugs that are taken between first study treatment intake and the end-of-study visit, will be coded using the WHO Drug Dictionary – March 2019 version or higher. Medication could be both previous and concomitant, if started before first study drug intake and continuing after first study drug intake. The Anatomical Therapeutic Chemicals (ATC) preferred names and codes will be attached to the clinical database. The numbers and percentages of subjects in the safety set with a previous medication and those with a concomitant medication will be summarized by ATC level 2 and ATC level 3 and presented overall and by treatment for the safety set (mITT). Subjects will appear in the summaries only once per ATC level 2 and ATC level 3. The summaries will be sorted alphabetically by ATC level 2 and then by decreasing frequency of ATC levels 3 in the "Overall" column. A listing of all previous and concomitant medications will be provided.

4.5. Treatment Compliance

Information regarding the number of subjects who have fully taken the three tablets during the single dose of study treatment and the number of subjects who vomit in the hour following the administration will be presented descriptively and data listing provided for the safety set (SS). The treatment compliance will also be summarized for the exploratory TrypSkin sub study.

5. Analysis

5.1. Primary Objective Analysis

The assessment of the safety of a single dose of acoziborole compared to placebo during a follow up period of 4 months will be conducted using excess rate (ER) of occurrence of any TEAE over the first 4 months in the acoziborole group compared to placebo. The numerator of the occurrence rate will be the number of subjects with at least one TEAE and the denominator will be the number of subjects receiving at least one tablet of IP. Loss of subjects to follow-up and withdrawal for a reason unrelated to safety or tolerance will not be considered as AEs. The 95%

confidence interval (CI) of the excess rate will be calculated with an exact binomial distribution method (Clopper-Pearson estimation). The excess rate (ER) of occurrence of any TEAE will be transformed to the number of subjects needed to treat (NNT) to observe one more subject with a TEAE in the acoziborole group compared to placebo (NNT = 1 / ER). The 95% CI will be calculated as inverted CI of the excess rate i.e., if the CI for the excess rate is given as (CI_L: CI_U) then the confidence interval for the NNT will be ($1/CI_U$:$1/CI_L$).

The analysis will be performed by age groups (defined as adolescents aged 15-17 years and adults aged ≥18 years), gender, TEAE severity category and period of occurrence. Four categorizations of TEAEs per severity will be derived as 1) Serious AEs regardless of the relationship to drug, 2) severe and possibly related to drug TEAEs and 3) mild and/or moderate regardless of relationship to drug, and 4) severe but not possibly related to drug. TEAEs period of occurrence will be categorized as: 1) Occurrence during the hospitalization period and 2) Occurrence after the hospitalization period (from EoH to month 4). The same event for the same subject can occur at both periods (i.e., the event is still present at the second period). However, for the analysis the AE start date will be used to identify AEs in each period. The denominator of occurrence rates will be the number of subjects receiving at least one tablet of IP for the first period and the number of subjects attending the 4-month visit or have reported a TEAE during the second period.

The occurrence rate analysis will be performed on safety set (mITT). MedDRA preferred terms of TEAE with occurrence rates ≥1% in one treatment group will be compared between treatment groups on the excess rates and NNT for any TEAE, by TEAE severity to assess whether the more severe events are not “diluted” in non-severe events and by period of occurrence to assess whether the late events are not “diluted” in all events.

A Cochran Mantel-Haenszel test stratified on the site of investigation will be used to compare the occurrence rate of both treatment groups. The same test will be applied to the common TEAE (>1% in any group) based on preferred terms. Generalizability of the treatment effect will be assessed using a test of homogeneity of the odds ratio (OR) across study centers. If the number of randomized subjects in each site reaches 100, then each site will be a stratum. If one or several sites had fewer than 100 subjects, then the smallest sites will be pooled together to obtain strata of at least 100 subjects. If the test of homogeneity is not significant (i.e., p-value ≥ 0.05), the combined odds ratio will be estimated as well as the 95% CI. If the hypothesis of homogeneity is rejected (i.e., p-value<0.05), then the odds ratio per site will be provided.

Time course treatment effect will be analyzed using a log rank test stratified by site and treatment group on the time to occurrence of TEAEs, considering the attrition rate of followed subjects. A loss to follow-up or a withdrawal will be censored at the time of the last attended visit. Kaplan Meier estimates of the proportion of subjects without any TEAE will be presented graphically to assess the time course of the response including results from log rank test.

The rate of occurrence of AEs will be analyzed by the true positive status (subjects with parasites detected during the follow up period) in either treatment or placebo groups.

The rate of conversion to parasite positive and the rate of return to seronegative over 4 months will be analyzed descriptively using frequency and proportion by treatment group.

There is a multiplicity of testing issue and an inflation of type one error that cannot be discarded. However, no adjustments of type one error are planned because the goal of the trial is to detect any potential safety issues instead of confirming and investigating a safety concern.

The listings for the TEAEs resulting in death, Serious TEAEs and TEAEs of grade 3 or higher will be provided.

5.2. Secondary Objectives Analysis

5.2.1. All adverse events

The occurrence rate of adverse events from ICF signature to end of study will be compared between the treatments groups.

5.2.2. Adverse Event Summary

An overview table of TEAEs will be provided summarizing the number (%) of subjects and number of events by any, serious, severity, possibly drug related TEAEs. The numbers and percentages of subjects who experienced the events and the number of events will also be summarized by SOC and PT and by PT only. Subjects will appear in the tables only once per SOC and PT. Tables with events summarized by SOC and PT will be sorted alphabetically by SOC and then by descending frequency of PT in the “Overall” column, and tables with events summarized by PT only will be sorted by descending frequency of PT in the “Overall” column. The summary will be presented for all TEAEs, serious TEAE, severity and possibly treatment related TEAE.

5.2.3. Laboratory Safety results

Changes from baseline at D5, M1 and M4 of haematology and blood chemistry parameters will be presented by treatment group as well as shift tables from baseline. The changes will be summarized using mean, standard deviation, median, interquartile range (Q1-Q2) and range (minimum – maximum). WMW test will be used when the data is highly skewed where assumption of normality seems not valid. Boxplots on the change from baseline and scatterplots of baseline values by time point and treatment group will be generated for each laboratory parameter. The boxplot of each treatment group will be plotted side-by-side at each time point.

5.2.4. Vital Signs results

The baseline summary of the vital signs will be provided using descriptive summaries of actual (absolute) values and changes from baseline will be presented for Body Mass Index (BMI), axillary temperature, respiratory rate, heart rate, systolic and diastolic blood pressure and Karnofsky performance status. These tables will be grouped by treatment group, and by visit. Individual vital signs data will also be presented in a subject listing.

5.2.5. Electrocardiogram (ECG)

The analysis of electrocardiogram (ECG) data will be described and reported separately.

5.2.6. Acoziborole Blood Concentration

The analysis of PK data will be described and reported separately

5.2.7. Exposure-Response Analysis

The analysis of PK data will be described and reported separately

5.2.8. Change in Signs and Symptoms

The binary variable of presence or absence of signs and symptoms will be derived from the clinical examination data. The change from baseline in number (proportion) of subjects with signs and symptoms during D1-D5, D5-M1, M1-M4 and D5-M4 will be compared between treatment and placebo group using contingency table with chi-square test (χ^2). The results will be presented in and physical examination data listing will be presented.

5.3. Exploratory Objectives (Sub-study) Analysis

5.3.1. Exploratory Analysis

The primary set of subjects for exploratory efficacy analyses is the modified ITT set composed of subjects who participated in the TrypSkin sub-study and who received at least on tablet of the randomized treatment (acoziborole or placebo). The primary exploratory objective is to establish whether acoziborole reduce the occurrence at 4 months of dermal trypanosomes in comparison with placebo. The exploratory efficacy will be evaluated using diagnostic and clinical endpoints. The primary diagnostic endpoint will be dermal trypanosome positive if TgSGP-PCR in skin is positive and the clinical endpoint will be the occurrence/presence of dermatitis or pruritus symptoms at 4 months. The occurrence rate of positive dermal trypanosome at 4 months based on the primary diagnostic parameters will be the primary measure of interest because the sample size is under control, and it corresponds to a measure of the level of potential transmission of reservoirs. The magnitude of the treatment effect will be the odds ratio of positive dermal trypanosomes. A Cochran Mantel Haenszel test using sites as strata will be used to compare acoziborole and placebo occurrence rate at 4 months of positive dermal trypanosome as well as

occurrence of dermatitis and/or pruritus symptoms. In case of small proportions, the exact Cochran-Mantel-Haenszel (CMH) test will be performed. The Breslow-Day test of homogeneity of the odds ratio across sites will be performed to assess the generalizability of the treatment effect. If the test is significant, a forest plot will be provided to display the heterogeneity of the magnitude of treatment effect across sites and to detect sites explaining the heterogeneity.

As the odds ratio is not easy to interpret when the rate of positive subjects in the placebo group is not small ($P_0 > 0.05$), the odds ratio (OR) will be converted into the risk ratio (RR) using the following formula:

$$RR = \frac{P_1}{P_0} = \frac{OR}{(1 - P_0) + (P_0 \times OR)}$$

where P_1 and P_0 are the estimated proportions of positive subjects in the acoziborole group and placebo group respectively. The excess rate ($P_1 - P_0$) will also be estimated as a secondary measure of the magnitude of the treatment as well as the NTT (number of subjects needed to treat to get one less positive subject at 4 months in acoziborole group. The exploratory efficacy analysis results will be presented and individual data listing.

5.3.2. Sensitivity Efficacy (sub-group) analysis

The secondary set of subjects for exploratory efficacy is the set of all seropositive subjects who were positive for the presence of trypanosome (Tb spp.) by at least one method in skin, blood and/or DBS at baseline. Secondary diagnostic parameters are occurrence of positive results by IHC in skin, by TBR-PCR, 18S-PCR in skin, blood and/or DBS, as well as by any novel diagnostic parameters corresponding to occurrence of positive results by multiplex qPCRs, multiplex RT-qPCRs and/or SHERLOCK in skin, blood and/or DBS. The secondary analyses will be exactly the same as the primary diagnostic analysis, as the main sensitivity analysis will be performed on the subgroup of positive dermal trypanosome subjects at entry (Secondary exploratory diagnostic Analysis). Missing data at month 4 will not be imputed to avoid loss of sensitivity of test (exploratory analysis of observed cases). However, a sensitivity analysis will assess the robustness of the primary analyses after imputation of a failure (positive diagnosis and presence of symptoms accordingly) in case of missing information.

5.3.3. Comparability of treatment groups

To assess whether randomization provided comparable treatment groups, randomized groups will be compared using descriptive statistics at baseline and more precisely the proportion of dermal trypanosome positive subjects at baseline and proportion of subjects with dermatitis or pruritus at baseline will be compared using chi-square (χ^2) test as a measure of deviation from perfect comparability.

5.3.4. Interpretability of results

The knowledge of the change of status between baseline and 4 months is helpful to interpret results. If in the placebo group a large proportion of subjects change their status from negative to positive and positive to negative, then the predictive validity of the diagnostic test is questionable (see secondary analysis of the primary endpoint).

Moreover, the positive association at baseline between dermatitis or pruritus symptoms and positive dermal trypanosome test will be assessed using the Cramer V coefficient and both treatment groups. In the presence of a significant association, both the clinical effect of acoziborole and the clinical validity of the dermal tryp test will be strengthen. The Cramer V coefficient measure is defined as:

$$V = \sqrt{\frac{\chi^2}{nt}}$$

Where *t* is the smaller of the number of rows minus one or the number of columns minus one. If *r* is the number of rows, and *c* is the number of columns, then

$$\mathrm{t} = \text{ Minimum } (\mathrm{r}-1, \mathrm{c}-1).$$

Cramer's V equals 0 when there is no relationship between the two variables, and generally has a maximum value of 1, regardless of the dimension of the table or the sample size. This makes it possible to use Cramer's V to compare the strength of association between any two cross classification tables. Tables which have a larger value for Cramer's V can be considered to have a strong relationship between the variables, with a smaller value for V indicating a weaker relationship.

5.3.5. Multiplicity of testing

Due to the exploratory nature of the sub-study, no adjustment for multiplicity of testing will be done. As a matter of fact, the level of the p value will be used to assess the strength of the finding.

5.3.6. Properties of diagnostic tests Analysis

Considering successively the results of (1) TgSGP-PCR in skin (primary diagnostic parameter), and (2) IHC, TBR-PCR and/or 18S-PCR in skin (secondary diagnostic parameters), as gold standard methods, the relative specificity, relative sensitivity, positive and relative negative predictive values of multiplex qPCRs, multiplex RT-qPCRs and SHERLOCK on skin, blood and DBS will be analyzed at enrolment. Four analyses will be performed for skin, blood, DBS and considering the result of any of the three samples.

Based on the placebo group data, the association between baseline and 4 months diagnosis results will be estimated through Cramer V coefficient for each diagnostic test and combination of diagnostic tests. The test or combination of tests providing the best association (reproducibility

of baseline result) will be used for ranking tests. Other properties: relative sensitivity, relative specificity, predictive values, Cohen Kappa coefficient, cost and rapidity for obtaining a response will also be used for ranking tests. Multicriteria ranking (sum of ranks) will be performed and the primary analysis will be redone on this new reference (best multicriteria rank) and each diagnostic test property will be computed with respect to this new reference.

6. Statistical software

All statistical analyses and generation of TLFs will be done using SAS software. Copyright © 2016 SAS Institute Inc[9].

7. Shell Tables, Listings and Figures (Attached separately)

8. References

1. *International Conference on Harmonisation; Guidance on Statistical Principles for Clinical Trials E9; Availability--FDA. Notice.* Vol 63.; 1998.

2. *International Conference on Harmonisation: Structure and Content of Clinical Study Reports E3, Japan and USA*.; 1995.

3. *Investigator Brochure Acoziborole V4.0 12 November 2019*.

4. Chappuis F, Stivanello E, Adams K, Kidane S, Pittet A, Bovier PA. Card agglutination test for trypanosomiasis (CATT) end-dilution titer and cerebrospinal fluid cell count as predictors of human African trypanosomiasis (Trypanosoma brucei gambiense) among serological suspected individuals in southern Sudan. *Am J Trop Med Hyg*. 2004;71:313–317.

5. Simarro PP, Ruiz JA, Franco JR, Josenando T. Attitude towards CATT-positive individuals without parasitological confirmation in the African trypanosomiasis (T. b. gambiense) focus of Quicama (Angola). *Trop Med Int Heal*. 1999;4:858–861.

6. Simarro PP, Franco JR, Ndongo P, Nguema E, Louis FJ, Jannin J. The elimination of Trypanosoma brucei gambiense sleeping sickness in the focus of Luba, Bioko Island, Equatorial Guinea. *Trop Med Int Heal*. 2006;11:636–646.

7. World Health Organization (WHO); *Control and Surveillance of Human African Trypanosomiasis, Technical Reports Series 984, Geneva*.; 2013.

8. Mariame Camara, Soumah AM, Ilbouldo H, et al. Dermal trypanosomes in suspected and confirmed cases of gambiense Human AfricanTrypanosomiasis. *Clin Infect Dis Infect Dis*. 2020. https://pubmed.ncbi.nlm.nih.gov/32638003.

9. SAS Institute Inc. 2016. SAS/STAT ® 9.4. Cary, NC: SAS Institute Inc.