Hangzhou Fasikl Technology Co., LTD

Fasikl Incorporated

Assessment of a novel NeuroAI-powered transcutaneous neuromodulation device in alleviating tremor symptoms and enhancing quality of life for patients with upper limb essential tremor (TRANQUIL)

FSK-CIP-2

Statistical Analysis Plan

[redacted]

Date: 13Nov2024

Sponsor Approval Page

Assessment of a novel NeuroAI-powered transcutaneous neuromodulation device in alleviating tremor symptoms and enhancing quality of life for patients with upper limb essential tremor (TRANQUIL)

FSK-CIP-2

Statistical Analysis Plan

[redacted]

Project Statistician: [redacted]

Company Name: [redacted]

Signature: _____ Date: _____

Review Statistician: [redacted]

Company Name: [redacted]

Signature: _____ Date: _____

Approved by Sponsor: [redacted]

Company Name: Hangzhou Fasikl Technology Co., LTD
Fasikl Incorporated

Signature: _____ Date: _____

Modification History

Version	Version Date	Author	Description
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]

Table of Contents

Abbreviation

Abbreviation	Specification
ADE	Adverse device effect
AE	Adverse event
ANCOVA	Analysis of covariance
CGI-S	Clinical global impression of severity
CGI-I	Clinical global impression of improvement
CIP	Clinical investigation plan
eCRF	Electronic case report form
ET	Essential tremor
FCS	Fully conditional specification
IRT	Interactive response technology
ITT	Intent-to-treat
LOCF	Last-observation-carried-forward
mADL	Modified activities of daily living
MAR	Missing at random
MI	Multiple imputation
mITT	Modified intent-to-treat
MMRM	Mixed-effect model for repeated measures
MRCT	Multi-region clinical trial
PD	Protocol deviation
PGI-I	Patient global impression of improvement

PGI-S	Patient global impression of severity
PPS	Per-protocol set
PS	Performance subscale
PT	Preferred term
QUEST	Quality of life in essential tremor questionnaire
SADE	Serious adverse device effect
SAE	Serious adverse event
SAP	Statistical analysis plan
SOC	System organ class
TETRAS	Tremor research group essential tremor rating assessment scale
TRANQUIL	Transcutaneous neuromodulation device in alleviating tremor symptoms and enhancing quality of life for patients with upper limb essential tremor
UADE	Unanticipated adverse device effect
US	United States
USADE	Unanticipated serious adverse device effect

Note: not include laboratory test.

1. Introduction

This statistical analysis plan (SAP) was drafted for the "Assessment of a novel NeuroAI-powered transcutaneous neuromodulation device in alleviating tremor symptoms and enhancing quality of life for patients with upper limb essential tremor (TRANQUIL)" (protocol No.: FSK-CIP-2) of Hangzhou Fasikl Technology Co., LTD and Fasikl Incorporated. In this document, the contents and methods of statistical analysis will be described in details.

This analysis specified in this SAP will be carried out at the end of the randomized phase (90-day visit), when the primary endpoint is reached. For this analysis, the primary endpoint analysis, secondary endpoint analysis as well as the safety analysis will be performed. Furthermore, the interim database lock for the analysis described in this SAP will occur after signing the SAP.

This SAP was based on clinical investigation plan (CIP, [redacted], 12Jul2024) and electronic case report form (eCRF, [redacted], 22Jul2024).

2. Study Objective (s)

The objective of this clinical investigation is to evaluate the effectiveness and safety of the Felix™ NeuroAI™ Wristband to aid in the relief of upper limb tremor in adults with essential tremor (ET).

3. Study Design

3.1. Study Type and Control Type

This is a prospective, randomized, sham-controlled, double-blinded, multi-center, multi-region clinical trial (MRCT). Approximately 8 sites in the US and 4 sites in China will participate in the study. Each country will enroll approximately 50% of the study population.

This study will consist of two phases:

- Randomized Phase (baseline to 90 days): Approximately N=120 subjects will be randomized in a 2:1 ratio to either the Felix group or Sham Stimulation group. Data from the randomized phase will be used to support regulatory application.
- Continued Access Phase (90 days to 1 year): All the patients in the randomized phase will be encouraged to participate in the continued access phase. Subjects randomized to the Sham arm will be crossed over to Felix stimulation. Data from the continued access phase will be used to support insurance coverage and medical guideline.

Subjects participating in this clinical investigation will be followed for 1 year (±30 days to allow for flexibility in scheduling the last-day visit). The expected duration of enrollment is 7 months. The total duration of the clinical investigation is expected to be 30 months.

3.2. Randomization Design and Execution

Randomization will be performed in the Interactive Response Technology (IRT) system via the randomization module of the system. Random codes will be generated using stratified blocked randomization by SAS 9.4. Stratification factors include:

[redacted]

Randomization will be managed directly by the eCRF platform. All eligible subjects enrolled in the study will be followed for 90 days with the option to continue up to 1 year. Subjects will be randomly assigned in a 2:1 allocation ratio to either the Felix group or Sham Stimulation group.

3.3. Blinding Design and Execution

All randomized patients will be blinded to the treatment assignment until they have reached the primary endpoint (90 days). All raters will be blinded to the treatment assignment. It is impossible to blind site personnel who will perform device fitting because of the different internal mechanisms between Felix and the sham device. Training on the importance of maintaining blinding will be provided to all site personnel. A blinding assessment will be performed at the end of the randomized phase (90-day visit). Patients will be asked to guess which treatment arm they were assigned to, which will be compared to the actual assignment.

Statistical analysis will be performed by the statistical team, who will be blinded to treatment assignment until the primary endpoint is reached. Sponsor personnel, except for the product support team, will be blinded to treatment assignment. A firewall will be established between the product support team and the rest of the sponsor personnel to ensure blinding.

3.4. Sample Size Consideration

The treatment effect is calculated as the average of the paired difference in the mADL scores between 90-day follow-up (f/u) and baseline pre-stimulation. The lower the value of the difference, the higher the treatment effect. Denote the treatment effect for the Felix arm and the sham arm by D_{Felix} and D_{Sham}, respectively.

D_{Felix} = Mean of (mADL at 90-day f/u – mADL at baseline pre-stimulation) of the Felix arm
D_{Sham} = Mean of (mADL at 90-day f/u – mADL at baseline pre-stimulation) of the sham arm

The statistical hypothesis is defined as:
H0: $D_{Felix} \geqslant D_{Sham}$
H1: $D_{Felix} < D_{Sham}$
Assumptions:

[redacted]

- One-sided alpha of 0.025

Using t-test for two independent samples with equal variances, a total sample size of 120 will provide 92% power.

Test of equal variance will be performed and if the variances are significantly different, then t-test for two independent samples with unequal variances will be used instead.

3.5. Study Flow Chart

Table 1. Schedule of Events

CIP Activity	Day 1/Baseline Pre-Stimulation (In-Clinic)	Day 1/Baseline Post-Stimulation (In-Clinic)	14 Days (In-Clinic /Virtual)	30 Days (In-Clinic)	60 Days (In-Clinic /Virtual)	90 Days[1] (In-Clinic)	180 Days (In-Clinic /Virtual)	1 Year (In-Clinic)
Visit/Visit window	**V1/-7 days**		**V2/±3 days**	**V3/±7 days**	**V4/±7 days**	**V5/±14 days**	**V6/±30 days**	**V7/±30 days**
Informed Consent	X							
Demographics[2]	X							
Physical Examination[3]	X							
Medical History[4]	X							
Medication[5]	X		X	X	X	X	X	X
Alcohol/marijuana consumption[6]	X		X	X	X	X	X	X
Pregnancy Test[7]	X							
Clinical diagnosis of ET	X							
Check against inclusion and exclusion criteria	X							
Randomization	X							
Device Fitting/Programming[8]	X							
TETRAS Performance Subscale (items 1, 4, and 8)	X	X		X		X		X
TETRAS Performance Subscale (items 6 and 7)	X	X	X	X	X	X	X	X

CGI-S	X	X		X		X		X
CGI-I		X		X		X		X
TETRAS ADL Subscale	X		X	X	X	X	X	X
PGI-S	X	X	X	X	X	X	X	X
PGI-I		X	X	X	X	X	X	X
QUEST	X					X		X
Survey[9]			X	X	X	X	X	X
Adverse Event	X		X	X	X	X	X	X
Protocol Deviation	X		X	X	X	X	X	X
Device Deficiency[10]	X		X	X	X	X	X	X
Withdrawal	X		X	X	X	X	X	X

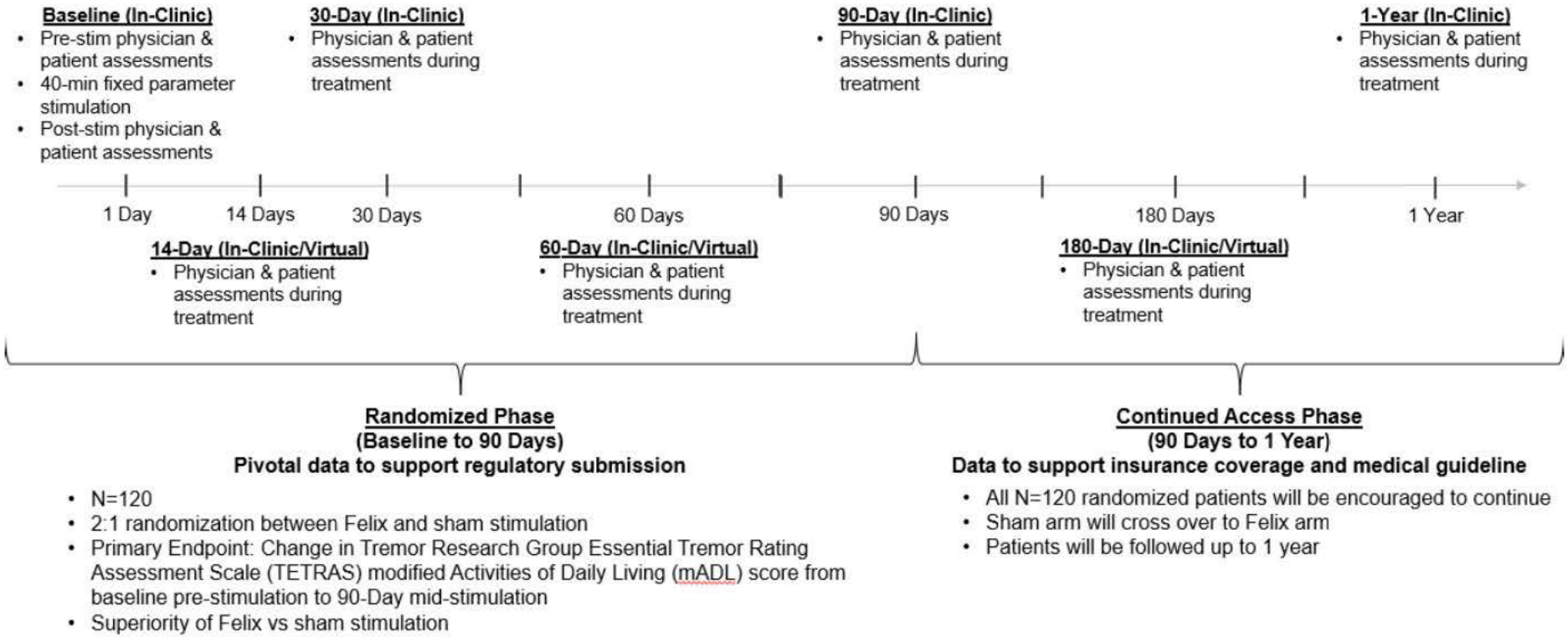
Notes:

1. Sham arm will cross over to Felix arm at 90 days and go through as steps in 9.3.8, and all subjects will be followed up to 1 year;
2. Demographic data: includes age (date of birth), sex, race, ethnicity;
3. Physical examination: includes height, weight, BMI, body part(s) with tremor;
4. Previous and current medical histories, includes ET onset/diagnosis, family history of ET, ET treatment(s), sensitivity of tremor to alcohol and caffeine, AUD as per DSM-5, history of drug abuse, allergic history, history of tumors, history of other serious diseases (such as symptomatic peripheral neuropathy condition of the involved upper, seizure, Parkinson's disease and tuberculosis) and surgical history (such as deep brain stimulation or thalamotomy) et al;
5. Record drugs taken include anti-tremor medications and antidepressant medications at pre-stimulation, and current medications; the drugs taken to treat tremor and concomitant medications for AEs and SAEs will be recorded during follow-up;
6. Record alcohol/marijuana consumption at baseline pre-stimulation and alcohol/marijuana consumption the day before a study visit;

7. Urine pregnancy tests at baseline pre-stimulation for all females of childbearing potential;
8. Sites and patients have the option of unilateral stimulation or bilateral stimulation. For unilateral stimulation, apply the device to the dominant hand;
9. Survey of satisfaction and durability of effect (i.e., how long tremor relieve lasts after stimulation);
10. User error as well as general issues related to the smartphone itself will not be considered as device deficiency.

For monitoring purposes, all subjects will be prompted to complete a brief patient diary on their smart phone daily during the first week of at-home treatment (Day 1 to 1 year), and weekly after the first week. The diary will include the information in section 9.3.4 of CIP.
Except for pre-stim physician & patient assessment, other physician & patient assessment will be measured after at least 40-min of stimulation.

Figure 1. Study Workflow



4. Endpoints

4.1. Efficacy Endpoints

4.1.1. Primary Endpoints

Change in Tremor Research Group Essential Tremor Rating Assessment Scale (TETRAS)14, 15 modified Activities of Daily Living (mADL) score from baseline pre-stimulation to Day 90, Superiority of Felix vs. Sham stimulation.

TETRAS mADL score is a composite sum of items 2 to 11 of the TETRAS ADL subscale and items 6 (bilateral) and 7 (dominant hand) of the TETRAS performance subscales (PS). These items are listed below, and the full TETRAS questionnaire is attached in Appendix 1. TETRAS mADL score is calculated as the sum of all 12 items and ranges from 0 to 52.

TETRAS ADL items (patient rating based on the prior week, each item is rated on a scale from 0 to 4, representing increasing difficulty in doing the activity due to tremor, from normal to severely abnormal/cannot perform the activity):

- Feeding with a spoon;
- Drinking from a glass;
- Hygiene;
- Dressing;
- Pouring;
- Carrying food trays, plates or similar items;
- Using keys;
- Writing;
- Working (if patient is retired, ask as if they were still working. If the patient is a housewife, ask the question as it relates to housework);
- Overall disability with the most affected task (name task, e.g. using computer mouse, writing, etc.).

TETRAS PS items (physician rating, each item is rated on a scale from 0 to 4, representing increasing severity of tremor, from normal to severe):

- Archimedes spirals (both hands, rated separately);
- Handwriting (dominant hand).

4.1.2. Secondary Endpoints

(1) Secondary mADL endpoints

① Change in TETRAS mADL score from baseline pre-stimulation to 14 days, 30 days, and 60 days for Felix vs. sham stimulation. Change from baseline pre-stimulation to 180 days and 1 year for Felix arm, and change from 90 days to 180 days and 1 year for sham arm.

② The proportion of patients responding to the treatment, as defined by change in TETRAS mADL score at or above 10%/15%/20% relative reduction, from baseline pre-stimulation to 14 days, 30 days, 60 days, and 90 days for Felix vs. sham stimulation. Change from baseline pre-stimulation to 180 days and 1 year for Felix arm, and change from 90 days to 180 days and 1 year for sham arm.

(2) Physician-rated assessments (in-clinic)

① TETRAS Performance Subscale (items 1, 4, and 8)

Collect and record the scores of TETRAS Performance Subscale (items 1, 4, and 8) at baseline pre-stimulation, baseline post-stimulation, 30 days, 90 days and 1 year. Items 4 and 8 will be rated separately for each limb.

② TETRAS Performance Subscale (items 6 and 7)

Collect and record the scores of TETRAS Performance Subscale (items 6 and 7) at baseline pre-stimulation, baseline post-stimulation, 14 days, 30 days, 60 days ,90 days, 180 day and 1 year. Item 6 will be rated separately for each hand. Item 7 will be rated for the dominant hand only.

③ TETRAS Dominant Hand Score

Defined as the total scores of TETRAS Performance Subscale for the dominant hand (items 4, 6, 7 and 8) at baseline pre-stimulation, baseline post-stimulation, 30 days, 90 days and 1 year.

④ Clinical Global Impression of Severity (CGI-S, Appendix 2 of Protocol)

Collect and record the scores of CGI-S at baseline pre-stimulation, baseline post-stimulation, 30 days, 90 days and 1 year.

⑤ Clinical Global Impression of Improvement (CGI-I, Appendix 2 of Protocol)

Collect and record the scores of CGI-I at baseline post-stimulation, 30 days, 90 days and 1 year.

(3) Subject-rated assessments (in-clinic and at-home)

① TETRAS ADL Subscale

Collect and record the scores of TETRAS ADL Subscale at baseline pre-stimulation, 14 days, 30 days, 60 days, 90 days, 180 days and 1 year.

② Patient Global Impression of Severity (PGI-S, Appendix 3 of Protocol)

Collect and record the scores of PGI-S at baseline pre-stimulation, 14 days, 30 days, 60 days, 90 days, 180 days and 1 year.

③ Patient Global Impression of Improvement (PGI-I, Appendix 3 of Protocol)

Collect and record the scores of PGI-I at baseline post-stimulation, 14 days, 30 days, 60 days, 90 days and 180 days and 1 year.

④ Quality of Life in Essential Tremor Questionnaire (QUEST, Appendix 4 of Protocol)

Collect and record the QUEST at baseline pre-stimulation, 90 days and 1 year.

⑤ Survey of satisfaction and durability of effect (i.e., how long tremor relieve lasts after stimulation, Appendix 5 of Protocol)

Collect and record the survey of satisfaction and durability of effect at 14 days, 30 days, 60 days, 90 days, 180 days and 1year.

(4) Investigational device-measured inertial data (continuously measured when the device is powered on)

Tremor power, a metric of tremor severity, will be calculated for both pre- and post-stimulation periods, using device-measured inertial data. The pre-stimulation period is defined as the initial phase of device usage each day, allowing for baseline measurements. The post-stimulation period is defined as a phase occurring toward the end of treatment, following a sufficient duration of treatment to assess its effects.

Tremor power will be determined by analyzing short segments of inertial data. For each segment, the dominant tremor frequency will be identified, and its power spectral density (PSD) will be calculated. Symptom changes will be quantified by calculating the mean log10-transformed tremor power for the pre- and post-stimulation periods. The comparison will involve the pre-stimulation periods after the initial visit and the post-stimulation periods before each follow-up.

Statistical significance of symptom changes over time will be evaluated. This analysis is intended to identify potential trends in tremor severity over the course of the study.

4.2. Safety Endpoints

4.2.1. Adverse Events

An adverse event (AE) is any untoward medical occurrence, unintended disease or injury, or untoward clinical signs (including abnormal laboratory findings) in subjects, users or other persons, whether or not related to the medical device under investigation.

As part of ISO14155 Section 3.2, the Adverse Event definition has the following notes:
Note 1: This definition includes events related to the medical device under investigation or the comparator.
Note 2: This definition includes events related to the procedures involved.
Note 3: For users or other persons, this definition is restricted to events related to medical devices under investigation.
If the AE meets any of the criteria below, it is regarded as a **serious adverse event (SAE)**.
a) Led to a death,
b) Led to a serious deterioration in health of the subject, users, or other persons as defined by one or more of the following:
1. a life-threatening illness or injury, or
2. a permanent impairment of a body structure or a body function, or
3. in-patient hospitalization or prolongation of existing hospitalization, or
4. medical or surgical intervention to prevent life threatening illness or injury or permanent impairment to a body structure or a body function.
5. chronic disease
c) Led to fetal distress, fetal death or a congenital abnormality or birth defect.
Note: A planned hospitalization for a pre-existing condition, or a procedure required by the CIP without a serious deterioration in health, is not considered to be an SAE.

An **(serious) adverse device effect (ADE/SADE)** is an AE/SAE related to the use of an investigational medical device. This definition includes AEs resulting from insufficient or inadequate instructions for use, deployment, implantation, installation, or operation, or any malfunction of the investigational medical device. The definition also includes any event resulting from use error or from intentional misuse of the investigational medical device. In conclusion, if the answer of "Is it an Adverse Device Effect?" is "yes", an AE/SAE will be defined as ADE/SADE.

Unanticipated (serious) adverse device effect (UADE/USADE) refers to any serious adverse effect on health or safety or any life-threatening problem or death caused by, or associated with, a device, if that effect, problem or death was not previously identified in nature, severity, or degree of incidence in the investigational plan or application (including a supplementary plan or application), or any other unanticipated serious problem associated with a device that relates to the rights, safety, or welfare of subjects. In conclusion, if the answer of "Is it an Unanticipated Adverse Device Effect?" is "yes", an AE/SAE will be defined as UADE/USADE.

4.2.2. Device Deficiency

Device deficiency is defined as an inadequacy of a medical device related to its identity, quality, durability, reliability, safety or performance, such as malfunction, misuse or use errors and inadequate labeling. This includes the failure of the device to meet its performance specifications or otherwise perform as intended.
Note: performance specifications include all claims made in the labeling of the device.

A device malfunction is the failure of a device to meet its performance specifications or otherwise perform as intended, when used in accordance with the instructions for use or CIP.

4.3. Demographics and Disease Characteristics

Demographics, including age, sex, race, ethnicity and childbearing potential, will be collected.
Essential tremor history, including ET treatment (yes/no), current anti-tremor medications use (yes/no), ongoing non-drug therapy for anti-tremor (yes/no), current antidepressant medications use (yes/no) and ongoing non-drug therapy for antidepressant (yes/no), will be recorded.

Family history of essential tremor (relationship, diagnosis date) will be collected.

Sensitivity of tremor to alcohol and caffeine will be collected.

History of drug abuse, including drug name, frequency, start date/end date and ongoing, will be recorded.

History of tumors, including lesion site, diagnosis date and end date, will be recorded.

Medical history that related to inclusion/exclusion criteria, including bandage allergy, alcohol use disorder, medical condition/event, start date/end date and ongoing, will be collected.

Surgical history (name of surgery, surgery date) that related to inclusion/exclusion criteria will be recorded.

Alcohol/marijuana consumption, including alcohol/marijuana used within 1 day prior to the follow-up visit and amount of alcohol/marijuana consumed, will be collected.

Physical Examination (Height, weight, BMI and body parts with tremor) will be collected.

Urine pregnancy test results (negative/positive) will be recorded when applicable.

4.4. Concomitant Medication

Concomitant medication from screening until study termination, including medication name, dose, unit, frequency, route, start date/end date, ongoing, indication type and indication, will be collected.

Concomitant non-drug therapy from screening until study termination, including therapy/procedure name, start date/end date, ongoing, indication type and indication, will be collected.

4.5. Protocol Deviation

Protocol deviation would be recorded by operation team.

4.6. Device usage

Device usage data will be record by device server and be summarized according to section 7.3.6.

5. Statistical Hypothesis

The null hypothesis is that the mean difference of (mADL at 90-day f/u – mADL at baseline pre-stimulation) of the Felix arm (D_{Felix}) is not superior to that achieved with the sham arm (D_{Sham}), and superiority margin $\Delta=0$:

$$H0: D_{Felix} \geqslant D_{Sham}$$

The alternative hypothesis is that mean difference of (mADL at 90-day f/u – mADL at baseline pre-stimulation) of the Felix arm is superior to that achieved with the sham arm:

$$H1: D_{Felix} < D_{Sham}$$

Significance level is set at 0.025 (one-sided). Because there is only one powered analysis (primary analysis of primary endpoint), study-wide type I error rate is controlled at 0.025 (one-sided).

Study success will be based on the primary analysis of the primary endpoint. Superiority of the Felix group compared to the Sham stimulation group will be assessed using the two-sided 95% confidence level of the difference of the average of the paired difference in the mADL scores between 90-day follow-up (f/u) and baseline pre-stimulation in two groups. The null hypothesis will be rejected if $p<0.025$ using t-test for two independent samples. When the upper limit of two-sided 95% confidence interval of the difference between the Felix arm and the sham arm (D_{Felix} - D_{Sham}) was less than 0, that is, H0 was rejected, the Felix arm could be considered as superior to the sham arm.

6. Analysis Datasets

Intent to Treat (ITT) set: the ITT analysis set is defined as patients who have been randomized. In accordance with the ITT principle, subjects will be kept in their originally assigned treatment group.

Modified Intent to Treat (mITT) set: it is the dataset that includes all the patients who have been randomized and have gone through at least 1 day of at-home treatment.

Per-Protocol Set (PPS): it refers to a subset of the ITT subjects who complete the evaluation of primary performance endpoint without major protocol deviation affected the evaluation of primary performance endpoint.

The primary analysis of the primary endpoint will be based on the ITT population. The primary endpoint analysis will also be performed on the mITT and PPS populations as a sensitivity analysis. All secondary efficacy endpoints will be analyzed for ITT, mITT, and PPS populations. The safety assessments will be analyzed for ITT population.

7. Statistical Methods

7.1. General Statistical Consideration

Demographics, clinical characteristics data, primary endpoint, secondary endpoints and safety endpoints will be analyzed.

Statistical descriptive methods: For continuous data, descriptive statistics will be calculated as number of patients, mean, standard deviation (Std), median, Q1, Q3, minimum, and maximum. The minimum and maximum will keep the same decimal place with the raw data, while mean, median, Q1 and Q3 will keep 1 more decimal place, and Std will keep 2 more decimal places. For categorical data, the frequency and/or percentage of patients in each category will be presented. Each frequency and/or percentage will keep 1 decimal place and a percent sign. Counts that are zero will be displayed as "0".

Statistical test whenever required: For continuous variable, two sample independent t-test or Wilcoxon rank sum test will be used. Levene's F test and Shapiro-Wilk test will be conducted for testing the homogeneity of variances and normality, respectively, and significance level is set at 0.1. For category variable, Chi-square test or Fisher's exact test (when sample size is less than 40 or the expected frequency of a fourfold table is less than 1). For ordinal variable, Wilcoxon rank sum test will be used. P-value will be rounded to 4 decimal places. '<0.0001' will be displayed when the p-value is lesser than 0.0001 and '>0.9999' will be displayed when the p-value is greater than 0.9999.

Unless otherwise indicated, all other tests will be at a 0.05 alpha using two-tailed tests.

All statistical analysis, including tabulations and listings, will be performed using the SAS® version 9.4 or above (SAS Institute, Cary NC, USA) software.

The baseline data is defined as the last available pre-treatment observed value.

Analysis, including demographics, clinical characteristics data, primary endpoint, secondary endpoints and safety endpoints, will be performed on all visits. For safety analysis, all data, including unscheduled visit data, will be used.

Data listings for statistical description and analyses (including subject disposition, efficacy analysis, safety analysis, etc.) will be provided. Subject screening ID will be used as the unique identification of subjects in this study.

7.2. Data Handling

7.2.1. Missing Data

If the earliest date of ET diagnosis is missing, and if the date of diagnosis is incomplete and affects the calculation of the course, it is calculated in the following way without affecting the logic premise:
- If the year, month and day are all missing, no calculation is performed;
- If the day and month are missing and do not contradict other dates, 1 July is assumed;
- If the day is missing and does not contradict other dates, the 15th is assumed.

The dates in the data listing are listed as they were recorded on the CRF.

For the primary endpoint based on ITT and mITT population, three kinds of imputation method will be performed as sensitivity analyses:

- ✓ Last-observation-carried-forward (LOCF) imputation: the former observed value of missing post-treatment visit will be used to impute. If subjects no post-treatment observed value, baseline value will be used.
- ✓ Mixed-effect model for repeated measures (MMRM) using PROC MIXED in SAS: under the assumption that the data is missing at random (MAR), the imputation for missing data is inherent in the statistical models. PROC MIXED uses maximum likelihood methods, data missing at random are ignorable.
- ✓ Multiple imputation (MI): PROC MI in SAS with Fully conditional specification (FCS) method will be conducted.

The results of these sensitivity analysis will be considered supportive in nature.

For all other analyses, no imputation of missing data will be performed.

For categorical variables, the missing data will be defined as "missing".

7.2.2. Derivation and Transformation on Data

Besides the calculation for some endpoints, which are based on data collected via CRF, no derivation or transformation would be done for data. Outliers will be reported to data management function as data issue to be resolved or confirmed before database lock. No special data handling will be done.

The course of ET (month)=(randomized date - the earliest date of ET diagnosis +1)/ 30.4375, and keep 1 decimal place.

1 foot=30.48 cm, 1 inch=2.54 cm, 1lb=0.45359 kg.

7.2.3. Special data handling

①This study plans to conduct the primary analysis after the 90-day visit of the last subject. The data range of the analysis is as follows: all the data generated at the completion of the 90-day visit of the last subject, will be based on all the data obtained to conduct database-lock and unblinding. All data from interim database lock will be used for this planned analysis.

②Some screening number of subject in the EDC system were inconsistent with other subjects. Hard code will be used to keep their formatting consistent, see as follow:

Original screening id	Formatting screening id
101S01	101S001
101S10	101S010
101S0011	101S011
101S0012	101S012
107S01	107S001
107S02	107S002

204S204S003	204S003

③For "On average, how long did tremor relief last after the device was turned off", those data that record as character will be derived as follow:

Original record	Derivation
none really noticed	0
unknown, I always take it off before bed	missing
"unknown, always wear until just before bed"	missing
"n/a" or NA or N/a	missing
varied	missing
?	missing
2-4	3
3-6	4.5
1-2	1.5
2-3	2.5
4-5	4.5

7.3. Subjects in the Study

7.3.1. Subject Disposition

The number of subjects who were enrolled, successful screening, screening failure, randomized and completed the study (primary analysis: completed the 90 days visit) will be provided, as well as the main reason of screening failure and the primary reason for early termination from study.

A descriptive summary of subject disposition by site and by country will be provided.

7.3.2. Protocol Deviation

Protocol deviations (PD) will be classified as major or minor. A descriptive summary of each PD category by site and by country will be presented.

7.3.3. Analysis Datasets

Based on subjects who randomized, the number of subjects who included or excluded from the ITT/mITT/PPS will be provided, as well as the reason of exclusion.

A descriptive summary of subject by site and by country will be presented.

7.3.4. Demographics and Baseline Characteristics

Demographic, essential tremor history, family history of essential tremor, sensitivity of tremor to alcohol and caffeine, history of drug abuse, history of tumors, alcohol/marijuana consumption variables will be tabulated by treatment group and by country. Continuous variables will be summarized by mean, standard deviation (Std), median, Q1, Q3, minimum and maximum, and categorical variables by a count and percentage.

A descriptive summary of subjects' physical examination (height, weight, BMI and body parts with tremor) will be reported. The result of urine pregnancy test (negative/positive) will be reported.

A descriptive summary of the demographics by country, and by site will be presented.

7.3.5. Medical History and Concomitant Medication

Medical history and surgical history will be summarized according to the latest MedDRA coding dictionary.

Concomitant medication will be summarized according to the latest WHODRUG coding dictionary, and concomitant non-drug therapy will be summarized according to the latest MedDRA coding dictionary.

7.3.6. Device usage

Summary of device usage over time, such as cumulative total and average duration of device on time, stimulation time, percentage of patients with unilateral stimulation and bi-lateral stimulation, percentage of patients with dominant hand stimulated and non-dominant hand stimulated and summary of default stimulation intensity set during V1 for each nerve (only Felix arm), will be performed using the methods described in section 7.1.

7.4. Efficacy Analysis

7.4.1. Analysis on Primary Efficacy Endpoint

The analysis of the primary endpoint will be performed on the ITT, mITT and PPS population. The primary analysis of primary endpoint will only be performed on the ITT populations, based on non-missing data.

The mADL score of each scheduled visit will be summarized by treatment group, as well as the change from baseline (pre-stimulation) to each post-stimulation visit.

For primary analysis of the primary endpoint, Levene's F test will be conducted for testing the homogeneity of variances, and t-test for two independent samples with equal variances will be performed if $P \geqslant 0.1$. Otherwise, t-test for two independent samples with unequal variances will be performed. Following sensitivity analysis will be performed:

① sensitivity analysis 1: an analysis of covariance (ANCOVA) model will be performed, the change of mADL score from baseline (pre-stimulation) to Day 90 with non-missing data will be used as the dependent variable, treatment (ref=Sham) as the independent and age (<65(ref), ≥65), gender (female (ref)/male), family history of ET (yes/no (ref)), prior ET treatment (yes/no (ref)), on ET medication (yes/no (ref)), tremor severity (mild/moderate/severe (ref)), site (ref=204), site*treatment (if this interaction effect is significant), country (ref=USA), country*treatment (if this interaction effect is significant), baseline mADL. Additional baseline characteristics may be included as covariates if there is an imbalance between treatment arms.

② sensitivity analysis 2: an analysis of covariance (ANCOVA) model will be performed where missing primary endpoint data will be imputed using LOCF imputation, the change of mADL score from baseline (pre-stimulation) to Day 90 will be used as the dependent variable, treatment (ref=Sham) as the independent and age (<65(ref), ≥65), gender (female (ref)/male), family history of ET (yes/no (ref)), prior ET treatment (yes/no (ref)), on ET medication (yes/no (ref)), tremor severity (mild/moderate/severe (ref)), site (ref=204), site*treatment (if this interaction effect is significant), country (ref=USA), country*treatment (if this interaction effect is significant), baseline mADL. Additional baseline characteristics may be included as covariates if there is an imbalance between treatment arms.

③ sensitivity analysis 3: a mixed-effect model for repeated measures (MMRM) will be performed, the change of mADL score from baseline (pre-stimulation) to Day 90 will be used as the dependent variable, treatment (ref=Sham) as the independent and age (<65(ref), ≥65), gender (female (ref)/male), family history of ET (yes/no (ref)), prior ET treatment (yes/no (ref)), on ET medication (yes/no (ref)), tremor severity (mild/moderate/severe (ref)), site (ref=204), site*treatment (if this interaction effect is significant), country (ref=USA), country*treatment (if this interaction effect is significant), baseline mADL, visit (ref=V2). Additional baseline characteristics may be included as covariates if there is an imbalance between treatment arms. For MMRM model, the unstructured model is preferred for the covariance matrix, and if the unstructured covariance structure does not converge, the following structures are attempted in

turn: Compound Symmetric/Heterogeneous Toeplitz/Heterogeneous First-order Autoregressive/Heterogeneous Compound Symmetric/Toeplitz/First-order Autoregressive.

④ sensitivity analysis 4: an ANCOVA model will be performed where missing primary endpoint data will be imputed using MI imputation, the change of mADL score from baseline (pre-stimulation) to Day 90 will be used as the dependent variable, treatment (ref=Sham) as the independent and age (<65(ref), ≥65), gender (female (ref)/male), family history of ET (yes/no (ref)), prior ET treatment (yes/no (ref)), on ET medication (yes/no (ref)), tremor severity (mild/moderate/severe (ref)), site (ref=204), site*treatment (if this interaction effect is significant), country (ref=USA), country*treatment (if this interaction effect is significant), baseline mADL. Additional baseline characteristics may be included as covariates if there is an imbalance between treatment arms.

7.4.2. Analysis on Secondary Efficacy Endpoint

The analysis of the secondary endpoint will be performed on the ITT, mITT and PPS population.

(1) Secondary mADL endpoints

① Change in TETRAS mADL score from baseline pre-stimulation to 14 days, 30 days, and 60 days for Felix vs. sham stimulation. Change from baseline pre-stimulation to 180 days and 1 year for Felix arm, and change from 90 days to 180 days and 1 year for sham arm.

Summarize the change from baseline pre-stimulation to 14 days, 30 days, and 60 days between two groups. And the t-test for two independent samples with equal/unequal variances or Wilcoxon rank sum test will be performed between treatment groups. Summarize the change from baseline pre-stimulation to 180 days and 1 year for Felix arm, and change from 90 days to 180 days and 1 year for sham arm. No statistical test for 180 days and 1 year visit.

② The proportion of patients responding to the treatment, as defined by change in TETRAS mADL score at or above 10%/15%/20%, from baseline pre-stimulation to 14 days, 30 days, 60 days, and 90 days for Felix vs. sham stimulation. Change from baseline pre-stimulation to 180 days and 1 year for Felix arm, and change from 90 days to 180 days and 1 year for sham arm.

Summarize the proportion of patients responding to the treatment. Chi-square test or Fisher's exact test will be used between treatment groups. No statistical test for 180 days and 1 year visit.

(2) Physician-rated assessments (in-clinic)

① TETRAS Performance Subscale (items 1, 4, and 8)

Summarize the scores of TETRAS Performance Subscale (items 1, 4, and 8) at baseline pre-stimulation, baseline post-stimulation, 30 days, 90 days and 1 year, as well as the change from baseline pre-stimulation to 30 days, 90 days and 1 year (change from 90 days to 1 year for sham arm). And the t-test for two independent samples with equal/unequal variances or Wilcoxon rank sum test will be performed between treatment groups. No statistical test for 1 year visit.

② TETRAS Performance Subscale (items 6 and 7)

Summarize the scores of TETRAS Performance Subscale (items 6 and 7) at baseline pre-stimulation, baseline post-stimulation, 14 days, 30 days, 60 days, 90 days, 180 days and 1 year, as well as the change from baseline pre-stimulation to 14 days, 30 days, 60 days, 90 days, 180 days and 1 year (change from 90 days to 180 days and 1 year for sham arm). And the t-test for two independent samples with equal/unequal variances or Wilcoxon rank sum test will be performed between treatment groups. No statistical test for 180 days and 1 year visit.

③ TETRAS Dominant Hand Score

Summarize the total scores of TETRAS Performance Subscale for the dominant hand (items 4, 6, 7 and 8) at baseline pre-stimulation, baseline post-stimulation, 30 days, 90 days and 1 year, as well as the change

from baseline pre-stimulation to 30 days, 90 days and 1 year (change from 90 days to 1 year for sham arm). And the t-test for two independent samples with equal/unequal variances or Wilcoxon rank sum test will be performed between treatment groups. No statistical test for 1 year visit.

④ Clinical Global Impression of Severity (CGI-S, Appendix 2 of Protocol)

Summarize the scores of CGI-S at baseline pre-stimulation, baseline post-stimulation, 30 days, 90 days and 1 year, as well as the change from baseline pre-stimulation to 30 days, 90 days and 1 year (change from 90 days to 1 year for sham arm). And the t-test for two independent samples with equal/unequal variances or Wilcoxon rank sum test will be performed between treatment groups. For ordinal variable, Wilcoxon rank sum test will be used, and Chi-square test or Fisher's exact test for category variable. No statistical test for 1 year visit.

⑤ Clinical Global Impression of Improvement (CGI-I, Appendix 2 of Protocol)

Summarize the scores of CGI-I at baseline post-stimulation, 30 days, 90 days and 1 year. And the t-test for two independent samples with equal/unequal variances or Wilcoxon rank sum test will be performed between treatment groups. For ordinal variable, Wilcoxon rank sum test will be used, and Chi-square test or Fisher's exact test for category variable. No statistical test for 1 year visit.

(3) Subject-rated assessments (in-clinic and at-home)

① TETRAS ADL Subscale

Summarize the scores of TETRAS ADL Subscale at baseline pre-stimulation, 14 days, 30 days, 60 days, 90 days, 180 days and 1 year, as well as the change from baseline pre-stimulation to 14 days, 30 days, 60 days, 90 days, 180 days and 1 year. And the t-test for two independent samples with equal/unequal variances or Wilcoxon rank sum test will be performed between treatment groups. No statistical test for 180 days and 1 year visit.

② Patient Global Impression of Severity (PGI-S, Appendix 3 of Protocol)

Summarize the scores of TETRAS ADL Subscale at baseline pre-stimulation, 14 days, 30 days, 60 days, 90 days, 180 days and 1 year, as well as the change from baseline pre-stimulation to 14 days, 30 days, 60 days, 90 days, 180 days and 1 year. And the t-test for two independent samples with equal/unequal variances or Wilcoxon rank sum test will be performed between treatment groups. For ordinal variable, Wilcoxon rank sum test will be used, and Chi-square test or Fisher's exact test for category variable. No statistical test for 180 days and 1 year visit.

③ Patient Global Impression of Improvement (PGI-I, Appendix 3 of Protocol)

Summarize the scores of PGI-I at baseline post-stimulation, 14 days, 30 days, 60 days, 90 days, 180 days and 1 year. And the t-test for two independent samples with equal/unequal variances or Wilcoxon rank sum test will be performed between treatment groups. For ordinal variable, Wilcoxon rank sum test will be used, and Chi-square test or Fisher's exact test for category variable. No statistical test for 180 days and 1 year visit.

④ Quality of Life in Essential Tremor Questionnaire (QUEST, Appendix 4 of Protocol)

Summarize the QUEST by each dimension at baseline pre-stimulation, 90 days and 1 year, as well as the change from baseline pre-stimulation to 90 days and 1 year (change from 90 days to 1 year for sham arm, if applicable). And the t-test for two independent samples with equal/unequal variances, Wilcoxon rank sum test, Chi-square test or Fisher's exact test will be performed between treatment groups. No statistical test for 1 year visit.

⑤ Survey of satisfaction and durability of effect (i.e., how long tremor relieve lasts after stimulation, Appendix 5 of Protocol)

Summarize the survey of satisfaction and durability of effect at 14 days, 30 days, 60 days, 90 days and 180 days and 1 year. The t-test for two independent samples with equal/unequal variances, Chi-square test,

Fisher's exact test or Wilcoxon rank sum test will be performed between treatment groups. No statistical test for 180 days and 1 year visit.

(4) Investigational device-measured inertial data (continuously measured when the device is powered on)

Tremor conditions will be characterized by device-measured inertial data, such as tremor power, tremor amplitude, and tremor frequency.

Summarize tremor power, tremor amplitude, and tremor frequency.

7.4.3. Poolability and Subgroup Analyses

(1) Poolability analysis

All poolability assessments will be performed based on the primary endpoint of the change in TETRAS mADL score from baseline pre-stimulation to Day 90 on the ITT population using non-missing data. A descriptive summary of mADL by sites and by country will be presented.

For poolability between countries (US and China), the analysis of the primary endpoint will be performed by an ANOVA model including treatment, country, country*treatment as the covariates. Significance of the interaction effect between country and treatment will be tested at α=0.15.

For poolability between sites, the analysis of the primary endpoint will be performed by an ANOVA model including treatment, site, site*treatment as the covariates. Significance of the interaction effect between site and treatment will be tested at α=0.15. Sites with less than 6 enrolled subjects will be excluded from this analysis.

If the interaction effect is significant, then the analysis of the primary endpoint will be performed on ITT population by including the model above that is significant and the following baseline variables:

[redacted]

(2) Subgroup analysis

Homogeneity of the treatment effect (primary analysis of the primary endpoint) will be assessed across the following subgroups:

1. Age (<65, ≥65)
2. Gender (female/male)
3. Family history of ET (yes/no)
4. Prior ET treatment (yes/no)
5. On ET medication (yes/no)
6. Tremor severity (mild/moderate/severe)

7.5. Safety Analysis

The safety assessments will be analyzed for ITT population.

7.5.1. Adverse Events

Adverse events will be summarized according to the latest MedDRA coding dictionary.

AE, SAE, ADE, SADE, UADE, and USADE will be reported. Events, subjects, and incidence of above adverse events between groups will be summarized. According to the coding dictionary, system organ class (SOC) and preferred term (PT) will be summarized, as well as by severity. If a subject has multiple adverse events of the same SOC/PT, only one subject will be counted under that SOC/PT. The incidence of adverse events will be reported by SOC/PT and severity. If a subject has multiple adverse events with the same PT, only one subject and the most severe degree will be counted under that PT. SOC and/or PT were presented in descending order of incidence in Felix group.

All summary tables will be based on adverse events from the ICF signed until the last visit.

A descriptive summary of AE by site/country will be provided.

7.5.2. Device Deficiency

Cases, subjects, and incidence of device deficiency will be reported between treatment groups, including by reason of the event. The incidence of any adverse events caused by device deficiencies will also be presented.

7.6. Other Analysis

Correlation analysis

Correlation analysis will be performed based on ITT population with non-missing data. The following correlation analyses between different efficacy endpoints will be performed pooling Felix and sham data using Spearman's rank correlation analysis, and correlation coefficients will be calculated:

① CGI-S/PGI-S vs TETRAS PS dominant hand score/TETRAS ADL total score/TETRAS mADL at different visits

② CGI-I/PGI-I vs Change in TETRAS PS dominant hand score/TETRAS ADL total score/TETRAS mADL from baseline pre-stimulation to 90 days

③ TETRAS PS dominant hand score vs TETRAS ADL total score at different visits.

Model fitting

Statistical models will be performed based on ITT population with non-missing data.

① Six mixed-effect models for repeated measures will be performed to analyze different visits of CGI-S/PGI-S vs TETRAS PS dominant hand score/TETRAS ADL total score/TETRAS mADL using PROC MIXED in SAS:

Model 1: the TETRAS PS dominant hand score will be used as the dependent variable, CGI-S as the independent and treatment and visits as the covariates.

Model 2: the TETRAS ADL total score will be used as the dependent variable, CGI-S as the independent and treatment and visits as the covariates.

Model 3: the TETRAS mADL score will be used as the dependent variable, CGI-S as the independent and treatment and visits as the covariates.

Model 4: the TETRAS PS dominant hand score will be used as the dependent variable, PGI-S as the independent and treatment and visits as the covariates.

Model 5: the TETRAS ADL total score will be used as the dependent variable, PGI-S as the independent and treatment and visits as the covariates.

Model 6: the TETRAS mADL score will be used as the dependent variable, PGI-S as the independent and treatment and visits as the covariates.

For MMRM model, the unstructured model is preferred for the covariance matrix, and if the unstructured covariance structure does not converge, the following structures are attempted in turn:

Compound Symmetric/Heterogeneous Toeplitz/Heterogeneous First-order Autoregressive/Heterogeneous Compound Symmetric/Toeplitz/First-order Autoregressive.

② Six multiple linear regression models will be performed to analyze CGI-I/PGI-I vs Change in TETRAS PS dominant hand score/TETRAS ADL total score/TETRAS mADL from baseline pre-stimulation to 90 days:

Model 1: the Change in TETRAS PS dominant hand score from baseline pre-stimulation to 90 days will be used as the dependent variable, CGI-I as the independent and treatment as the covariate.

Model 2: the Change in TETRAS ADL total score from baseline pre-stimulation to 90 days will be used as the dependent variable, CGI-I as the independent and treatment as the covariate.

Model 3: the Change in TETRAS mADL from baseline pre-stimulation to 90 days will be used as the dependent variable, CGI-I as the independent and treatment as the covariate.

Model 4: the Change in TETRAS PS dominant hand score from baseline pre-stimulation to 90 days will be used as the dependent variable, PGI-I as the independent and treatment as the covariate.

Model 5: the Change in TETRAS ADL total score from baseline pre-stimulation to 90 days will be used as the dependent variable, PGI-I as the independent and treatment as the covariate.

Model 6: the Change in TETRAS mADL from baseline pre-stimulation to 90 days will be used as the dependent variable, PGI-I as the independent and treatment as the covariate.

③ A mixed-effect model for repeated measures will be performed to analyze different visits of TETRAS PS dominant hand score vs TETRAS ADL total score using PROC MIXED in SAS: the TETRAS ADL total score will be used as the dependent variable, TETRAS PS dominant hand score as the independent and treatment and visits as the covariates. For MMRM model, the unstructured model is preferred for the covariance matrix, and if the unstructured covariance structure does not converge, the following structures are attempted in turn: Compound Symmetric/Heterogeneous Toeplitz/Heterogeneous First-order Autoregressive/Heterogeneous Compound Symmetric/Toeplitz/First-order Autoregressive.

7.7. Change from the Analysis Plan in Protocol

Add correlation analyses and model fitting to explore the associates between efficacy endpoints in section 7.6.

8. Reference

1. National Medical Products Administration. Biostatistical Principles for Drug Clinical Trials. 2016 Jul.
2. National Medical Products Administration. Data Management and Statistical Analysis Plan for Drug Clinical Trials. 2021; 63.
3. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. E9 Statistical Principles for Clinical Trials.

9. TFL Shell

Referred to attachment FSK-CIP-2_TFL mock shell.

10. Dataset Specification

Referred to attachment FSK-CIP-2_ADS SPECS.