STATISTICAL ANALYSIS PLAN (SAP)

Study of Abobotulinum Toxin Versus Neubotulinum Toxin Injection For Hemifacial Spasm in Thai Patients

ClinicalTrials.gov Identifier	NCT04589364
Study No.	A-01-2020
Document	DNHFS Randomized Double-Blind Cross-over Trial - Statistical Analysis Plan
Analysis	GEE and MMRM Longitudinal Analysis
Version	Draft v0.5
Document Date	09 August 2026
Status	Pre-unblinding analytical specification

This cover page is included to identify the official registered study title, ClinicalTrials.gov identifier, and document date. The statistical analysis specifications that follow are unchanged from SAP Draft v0.4 dated 09 August 2026.

STATISTICAL ANALYSIS PLAN (SAP)

DNHFS Randomized Double-Blind Cross-over Trial

GEE and MMRM Longitudinal Analysis

Draft v0.4 | 09 August 2026 | Pre-unblinding analytical specification

Protocol	A-01-2020, Phase III, multicenter, randomized, double-blind cross-over study
Treatments	Abobotulinum toxin type A (Dysport) 100 U vs Neubotulinum toxin type A (Neuronox) 33.33 U
Randomized structure	Two treatment sequences (A→B or B→A); treatment-period effect and sequence effect modeled explicitly
Study schedule	Repeated 12-week injection cycles within two treatment blocks, with protocol-specified washout before crossover
Primary data source	24-hour HFS diary with hourly symptom/disability/sleep coding
Primary analysis population	ITT/FAS; PP as key sensitivity for non-inferiority
Core principle	Preserve protocol-defined 4-week total-intensity non-inferiority question while using cross-over GEE/MMRM to exploit repeated diary data

1. Purpose

This SAP specifies the confirmatory and supportive statistical analyses for the DNHFS cross-over trial before treatment-code unblinding. It incorporates the randomized sequence, treatment period, within-patient cross-over comparison, repeated injection cycles, dense daily diary observations, baseline severity, and possible carryover. The SAP also defines complementary GEE and MMRM analyses so that treatment conclusions are not driven by a single modeling framework.

2. Trial structure to be represented in the analysis

The protocol schedules injections and assessments across a long-term cross-over design. For statistical purposes, the data will be organized into two treatment blocks. Treatment Block 1 consists of the participant's initially randomized formulation and its repeated injection cycle(s); Treatment Block 2 consists of the alternate formulation after the protocol-specified washout/crossover interval. Each injection creates a new post-injection diary cycle with Day 0 defined as the injection date and Days 1–84 (or the available protocol-defined 12-week window) as follow-up.

Analysis level	Variable	Examples	Role in model	Notes
Participant	ID, site, randomized sequence	AB / BA	Clustering / random subject	Between-subject unit
Treatment block	Period 1 / Period 2	Before and after crossover	Fixed period effect	Controls secular/period differences
Injection cycle	Cycle 1–4	V1, V2, V3, V4 injections	Repeated cycle within participant	Cycle nested within period
Post-injection time	Day 1–84 or Week 1–12	Daily diary	Repeated time	Pharmacodynamic trajectory
Treatment	Dysport / Neuronox	100 U / 33.33 U	Primary fixed effect	Within-subject comparison after crossover

3. Protocol-defined endpoints and estimands

The protocol identifies the 24-hour HFS diary as a primary efficacy source and includes time to peak improvement, response duration, total intensity score per day, facial-spasm duration, and functional-impairment duration. The protocol also states a non-inferiority objective for the total intensity score at the 4-week period. Accordingly, this SAP preserves the 4-week total-intensity non-inferiority estimand as the protocol-defined confirmatory question and treats the full longitudinal diary analysis as a prespecified supportive/estimand-enriching analysis.

Endpoint	Primary time frame	Estimand	Direction	Analysis status
Total intensity score/day	Weeks 1–4 after each relevant injection	Within-patient treatment difference in baseline-adjusted 4-week mean total intensity	Lower better	Confirmatory NI + supportive superiority
Total HFS duration (h/day)	Weeks 1–4 and full 12-week cycle	Within-patient treatment difference in change from cycle baseline	Lower better	Key supportive
HFS with disability (h/day)	Weeks 1–4 and full cycle	Treatment difference in disability burden	Lower better	Key supportive; zero-heavy

Endpoint	Primary time frame	Estimand	Direction	Analysis status
HFS without disability (h/day)	Weeks 1–4 and full cycle	Treatment difference in non-disabling spasm burden	Lower better	Supportive
Sleep duration (h/day)	Weeks 1–4 and full cycle	Treatment difference in change from baseline	Higher better	Supportive
Time to peak improvement	Per injection cycle	Within-patient treatment difference in days	Earlier preferred if efficacy comparable	Supportive
Response duration	Per injection cycle	Within-patient treatment difference in days to loss of effect	Longer better	Key supportive
HFS-30 / AIMS / SF-36 / CES-D / PHQ-9	Protocol visits	Within-patient treatment/period adjusted difference	Scale-specific	Secondary

4. Analysis populations

- ITT/FAS: all randomized participants who receive at least one randomized treatment and have at least one post-treatment efficacy assessment. Participants are analyzed according to randomized sequence/treatment assignment.
- Per-Protocol (PP): participants without major protocol deviations affecting treatment exposure, crossover integrity, primary endpoint measurement, or washout. PP is a key analysis population for non-inferiority.
- Safety population: all treated participants, analyzed according to treatment actually received in each treatment period.
- Cross-over evaluable set: participants with evaluable data under both formulations. This set will be supportive and will not replace ITT/FAS.

5. Baseline definitions

Because the study contains repeated injections and crossover, a single study-entry baseline is insufficient for all treatment-effect questions. Three baseline concepts will therefore be retained: (1) study baseline before the first randomized injection; (2) treatment-period baseline immediately before the first injection in each treatment block; and (3) cycle baseline immediately before each injection. The baseline used in a model will correspond to the outcome and estimand being analyzed. For cycle-level diary outcomes, the cycle-specific pretreatment baseline is preferred if available. No post-injection measurement will be used to construct baseline.

6. Data structure and derived diary variables

- One long-format record per participant × treatment block × injection cycle × post-injection day.
- Daily total HFS duration = hours coded as HFS with or without disability.
- Daily disability duration = hours coded as HFS with disability.
- Daily non-disability HFS duration = total HFS hours − disability hours.
- Daily sleep duration = hours coded as sleep.
- Daily total intensity score = sum/mean according to the protocol diary scoring algorithm; the exact derivation will be locked before unblinding.
- Change from baseline = post-treatment value − corresponding baseline value; for outcomes where lower is better, negative change indicates improvement.
- Reduction from baseline = baseline − post-treatment value; positive values indicate improvement.
- Weekly summaries = participant-level mean of valid days within Days 1–7, 8–14, 15–21, and 22–28 for the 4-week MMRM analysis; analogous weeks will be created for the 12-week supportive trajectory.

7. Confirmatory non-inferiority analysis at 4 weeks

The protocol-defined confirmatory analysis will compare Neubotulinum toxin A 33.33 U with Abobotulinum toxin A 100 U for total intensity score during the first 4 weeks after treatment. The non-inferiority margin specified in the protocol is $\delta = 0.36$ units. The treatment effect will be oriented as Neuronox minus Dysport, with lower total-intensity scores being favorable. Non-inferiority will be concluded if the upper bound of the prespecified confidence interval for the adjusted treatment difference is below +0.36.

Because non-inferiority trials are sensitive to protocol deviations, the primary NI conclusion will require concordant interpretation in both ITT/FAS and PP populations. If ITT and PP differ materially, both will be reported without choosing the more favorable result. If non-inferiority is established, superiority will be evaluated only as a prespecified hierarchical supportive analysis using the same adjusted model and a two-sided 95% CI. No additional confirmatory superiority hypothesis, multiplicity rule, endpoint elevation, or selective repeated-time contrast will be

introduced after treatment-code access; any analysis outside this hierarchy will be labeled exploratory.

8. Cross-over GEE model

GEE will be the principal population-average longitudinal model for dense daily diary outcomes. It will exploit all valid repeated diary observations while accounting for within-participant correlation and the cross-over structure.

Component	Primary GEE specification	Sensitivity / fallback	Purpose
Outcome	Daily total intensity or daily HFS burden/change from cycle baseline	Absolute post-treatment outcome + cycle baseline covariate	Preserves clinical change while enabling adjusted analysis
Cluster	Participant	Participant × treatment block if software requires nested implementation	Accounts for repeated within-person measurements
Fixed effects	Treatment + period + sequence + injection-cycle-within-period + day + treatment×day + baseline	Add site and prespecified prognostic covariates in sensitivity models	Separates treatment from period/sequence/time effects
Time	Categorical day for primary 4-week trajectory	Restricted cubic spline / week grouping	Avoids assuming linear response
Working correlation	AR(1) preferred for daily data	Exchangeable, then independence	Adjacent diary days expected to correlate more strongly
SE	Robust sandwich	Small-sample corrected sandwich if available	Robust to correlation misspecification
Primary 4-week contrast	Model-based treatment difference averaged over Days 1–28	Week-specific contrasts	Head-to-head early efficacy
Full-cycle contrast	Average treatment difference over Days 1–84	AUC-like model-based marginal mean difference	Treatment-cycle burden
Interaction	Treatment×day	Treatment×week	Describes onset/waning, not sole efficacy criterion

Model form (conceptual): $E(Y) = \beta0 + \beta1(\text{Treatment}) + \beta2(\text{Period}) + \beta3(\text{Sequence}) + \beta4(\text{Cycle}) + \beta5(\text{Time}) + \beta6(\text{Treatment}\times\text{Time}) + \beta7(\text{Baseline})$ [+ Site]. Treatment is the principal estimand; period and sequence are retained regardless of statistical significance because they are design variables.

9. Cross-over MMRM model

MMRM will provide a complementary likelihood-based analysis for continuous diary-derived outcomes and visit-based scales. For the 4-week diary analysis, daily observations will be aggregated into four weekly repeated measurements to stabilize covariance estimation. For the full 12-week cycle, weekly aggregation will produce up to 12 repeated measurements per injection cycle.

Component	Primary MMRM specification	Fallback	Interpretation
Outcome	Weekly mean total intensity/HFS burden	Weekly change from baseline or absolute outcome + baseline	Continuous repeated efficacy
Fixed effects	Treatment + period + sequence + cycle-within-period + week + treatment×week + baseline	Add site as sensitivity	Cross-over adjusted trajectory
Participant effect	Repeated measures within participant	Random intercept only if software implementation requires mixed-model form	Captures within-subject dependence
Covariance	Unstructured for 4 weekly values	Heterogeneous AR(1), AR(1), then compound symmetry	Flexible covariance with prespecified hierarchy
12-week covariance	Heterogeneous AR(1) preferred	AR(1), Toeplitz, then CS	UN may be unstable with 12 weeks
Estimation	REML	ML for covariance/model comparison only	Standard MMRM
Degrees of freedom	Kenward–Roger	Satterthwaite; asymptotic only if unavailable	Small-sample inference
Primary contrast	LS-mean treatment difference averaged over Weeks 1–4	Week-4 and week-specific contrasts	Sensitivity to GEE/NI result
Missing data	Uses all observed weekly outcomes under MAR	Pattern-mixture / tipping-point if material missingness	No LOCF

10. Treatment-period, sequence, and carryover assessment

Cross-over inference requires separation of treatment effect from period and sequence. Period and randomized sequence will therefore be included in all principal cross-over models independent of their p values. Carryover will not be tested by a low-powered preliminary significance test and then ignored. Instead, carryover risk will be

assessed using the protocol-defined washout, diary return toward pretreatment severity before crossover, response-duration data, and sensitivity analyses.

- Primary assumption: the protocol-specified washout is adequate and clinically meaningful residual treatment effect is absent at the start of the crossover block.
- Diagnostic: summarize cycle/period baseline HFS burden by sequence and treatment block; display standardized differences and within-subject return toward baseline.
- Carryover sensitivity 1: add previous-treatment/carryover indicator to the model.
- Carryover sensitivity 2: analyze first treatment block only as a parallel-group comparison.
- Carryover sensitivity 3: exclude participants whose response-duration diary indicates ongoing effect at the crossover baseline.
- If clinically important carryover is evident, the full crossover estimate will be interpreted cautiously and the first-period analysis will become the key supportive estimate rather than silently assuming washout.

11. Repeated injection cycles within each formulation

The protocol contains repeated 12-week injections within each treatment block. Therefore, injection cycle is not interchangeable with treatment period. Cycle will be modeled explicitly (cycle within treatment block/period), allowing estimation of treatment effects while accounting for possible second-injection or habituation effects. Treatment×cycle interaction will be exploratory unless prespecified before unblinding. The principal treatment effect will average across comparable injection cycles within each formulation.

12. Distribution-specific analysis rules

- Total intensity score / total HFS hours: Gaussian identity-link GEE/MMRM may be used if participant-level or weekly summaries are reasonably continuous; robust SE and distributional diagnostics remain mandatory.
- Disability hours: expected to be zero-heavy. First quantify proportion of zero days. If substantial, use a two-part model: (a) probability of any disability, and (b) positive disability duration conditional on disability. Gaussian MMRM is not primary for this endpoint unless diagnostics support it.
- Counts, if any diary-derived endpoint is represented as a true count: assess overdispersion; use Poisson or negative-binomial GEE as appropriate.
- Time to peak improvement and response duration: summarize Kaplan–Meier-like distributions when censoring is present; compare within crossover using models accounting for participant, period, and treatment where feasible.
- All outcomes: report mean±SD and median [IQR], because skewness and zero inflation are clinically informative.

13. Missing data and diary completeness

Diary completeness will be summarized by participant, treatment, period, cycle, week, and day. A valid-day rule and minimum-valid-days-per-week rule must be locked before unblinding. GEE will use all available valid daily observations without single imputation. MMRM will use all available valid weekly observations under the MAR assumption. No LOCF will be used. If missingness is material or differs by treatment/period, sensitivity analyses will include complete/near-complete diary analysis and, where appropriate, pattern-mixture or delta-adjustment/tipping-point analyses.

14. Baseline imbalance and covariate adjustment

Baseline comparability will be described using means/SDs, medians/IQRs, and standardized mean differences rather than baseline hypothesis testing. The corresponding pretreatment baseline for each outcome is mandatory in adjusted continuous-outcome models. Randomized sequence and treatment period are design variables and will always be retained. Site will be included as a fixed-effect sensitivity covariate if cell sizes permit stable estimation.

15. Non-inferiority, superiority, multiplicity, and interpretation

The protocol-defined non-inferiority comparison of 4-week total intensity is the confirmatory efficacy question. Other diary outcomes and longitudinal trajectories are supportive unless formally elevated in a signed SAP amendment before unblinding. For non-inferiority, emphasis will be placed on the treatment-effect confidence interval relative to $\delta=0.36$, not on a conventional $p<0.05$ superiority test. Any superiority claim following NI will follow the prespecified hierarchical rule and will be clearly separated from exploratory comparisons. Daily or weekly repeated contrasts will not each be treated as independent confirmatory tests.

16. GEE–MMRM concordance rule

The principal interpretation will consider effect direction, magnitude, clinical relevance, and 95% CI across GEE and MMRM. A result will be regarded as statistically robust when both models support materially consistent treatment conclusions. If they disagree, both results will be retained and the discrepancy evaluated in terms of missingness,

covariance assumptions, outcome distribution, baseline/carryover, and estimand differences. No model will be discarded because its p value is less favorable.

17. Preliminary sponsor-unblinded 1-month analysis

A sponsor-unblinded analysis after the first month is prespecified as a preliminary, descriptive/supportive data cut and is not the final cross-over efficacy analysis. It will use only the first 28 post-injection days available at that data cut and will preserve randomized treatment, baseline adjustment, site, period/sequence information, and the analysis populations defined in this SAP where applicable. Access to treatment codes for this preliminary analysis will not permit modification of the final estimand, confirmatory 4-week non-inferiority hypothesis or margin, endpoint hierarchy, derivations, valid-day/completeness rules, covariate hierarchy, carryover rules, GEE/MMRM specifications, covariance/working-correlation hierarchy, or multiplicity strategy on the basis of observed treatment effects. Any later change must be documented as a dated SAP amendment with rationale and explicit identification of whether treatment-code information was available. If the 1-month analysis is used for formal adaptation, stopping, or confirmatory inference rather than descriptive/supportive review, a separate prospective interim-analysis procedure, including appropriate type-I-error control, must be finalized before code opening.

18. Safety analyses

- AE and SAE incidence by treatment and treatment period, using the safety population.
- Treatment-emergent AEs attributed to each injection cycle; multiple events in one participant summarized both as participant incidence and event counts.
- Exact confidence intervals/risk differences for sparse toxin-related AEs.
- Period/sequence summaries to identify whether AE risk differs after crossover or repeated exposure.
- No efficacy-driven exclusion of safety data.

19. Planned tables and figures

- Table 1. Participant disposition by randomized sequence and analysis population.
- Table 2. Baseline characteristics at study entry and crossover baseline, including SMDs.
- Table 3. Diary completeness by treatment, period, cycle, and week.
- Table 4. Protocol-defined 4-week total-intensity NI analysis: ITT and PP.
- Table 5. Cross-over GEE treatment effects for total HFS burden, disability burden, and sleep.
- Table 6. Cross-over MMRM LS-means and treatment contrasts by week.
- Table 7. Time to peak improvement and response-duration comparisons.
- Table 8. Carryover/first-period sensitivity analyses.
- Table 9. Safety by treatment and treatment period.
- Figure 1. Daily baseline-adjusted HFS burden trajectories by treatment.
- Figure 2. Weekly MMRM estimated marginal means with 95% CIs.
- Figure 3. Within-participant paired treatment-effect plot for 4-week total intensity.
- Figure 4. Response-duration curves by treatment.
- Figure 5. Disability zero-proportion and positive-duration components by treatment.

20. Data lock and unblinding rules

Before final treatment-code merge, the analysis dataset, treatment-period mapping, injection dates, crossover dates, period/cycle baselines, diary coding, valid-day rules, protocol deviations, carryover flags, and AE coding must be frozen as far as possible. Programs should be developed using masked treatment labels (e.g., Treatment A/B). Any post-unblinding change to an endpoint definition, derivation, population, covariance hierarchy, carryover rule, or model must be documented with date, rationale, and whether it was data-driven.

21. Pre-unblinding protocol reconciliation and analysis lock

The protocol contains internal inconsistencies in visit timing/dosing language. These are operational data-reconciliation issues and will not be resolved by examining comparative efficacy results. Before the relevant treatment-code access and before final database lock, the randomization list, pharmacy/accountability records, CRF, injection dates, and treatment-administration records will be reconciled under masked treatment labels whenever feasible. The statistical dataset will use the verified randomized/administered treatment, period, injection cycle, and dose from source/accountability records rather than an unverified textual protocol entry. The reconciliation decisions below must be documented in a dated data-definition/analysis decision log before the corresponding

comparative analysis is run; they may not be selected according to which option produces a more favorable treatment result.

Randomized sequence coding (AB/BA) will be taken directly from the final randomization list and mapped to masked Treatment A/B labels before comparative efficacy analysis.

Treatment Period 1 and Treatment Period 2 boundaries will be defined from the verified crossover and treatment-administration dates; the mapping rule will be frozen before treatment-effect estimation.

Injection-cycle doses at V1-V4 will be derived from pharmacy/accountability and administration records. Any discrepancy with narrative protocol text will be listed and resolved by source-document hierarchy without reference to efficacy outcomes.

The actual washout/crossover interval will be calculated from verified administration dates. Protocol deviations from the planned interval will be flagged prospectively for PP classification and carryover sensitivity analyses.

The confirmatory estimand remains the protocol-defined baseline-adjusted mean total-intensity score over Days 1-28 after the relevant injection. The exact rule identifying the relevant injection(s) in repeated cycles will be frozen from protocol/CRF/source chronology before comparative treatment results are examined; alternative cycle-level summaries will be supportive only.

Total-intensity diary derivation will follow the protocol diary scoring algorithm. The NI contrast is oriented as Neuronox minus Dysport, lower scores are favorable, and non-inferiority requires the upper confidence-limit bound to be below +0.36.

The valid-day and minimum-valid-days-per-week rules will be finalized from diary/CRF conventions and documented before treatment-effect modeling. No threshold will be chosen after examining treatment-group efficacy differences.

The sponsor's planned 1-month code opening is classified in this SAP as preliminary descriptive/supportive analysis unless a separate prospective formal interim procedure is finalized before code opening. It will not alter the final cross-over estimand or confirmatory hierarchy.

22. Statistical software and reproducibility

Analyses may be implemented in R, SAS, Stata, or Python using validated procedures. The final report will document software/version, procedure/package, link function, covariance or working-correlation structure, convergence status, degrees-of-freedom method, and exact contrasts. Reproducible analysis code and a treatment-blinded output shell will be archived before unblinding.

SAP summary. The final DNHFS analysis will treat the trial as a true randomized cross-over experiment, not as two independent parallel groups. Treatment, period, sequence, repeated injection cycle, time, and baseline will be modeled explicitly. GEE will quantify population-average longitudinal diary effects; MMRM will provide likelihood-based repeated-measures confirmation; and the protocol-defined 4-week total-intensity non-inferiority question will remain the confirmatory anchor.