FRESENIUS KABI

STATISTICAL ANALYSIS PLAN

Investigational Product:	Aurora Xi with Investigational Software 2.0
CSP ID:	PAND-004-CMD
Study Title	Evaluation of the Aurora Xi New Nomogram Investigational Software 2.0
Version:	02 Oct 2023 Final

Contents

PAND-004-CMD

Evaluation of the Aurora Xi New Nomogram Investigational Software 2.0

Signature Page

This Statistical Analysis Plan has been reviewed and approved by the following Fresenius Kabi personnel:

_____ _____

[redacted]
Sr. Statistical Specialist

Date

_____ _____

[redacted]
Manager, Clinical Affairs

Date

_____ _____

[redacted]
Sr. Director, Clinical Affairs

Date

1 Abbreviations

CSP	Clinical Study Protocol
CSP ID	Fresenius Kabi Study Identifier
DAE	Donor Adverse Event
DSMB	Data Safety Monitoring Board
IQPP	International Quality Plasma Program
ITT	Intention-to-Treat, Intent-to-Treat
PP	Per-Protocol
SHAE	Significant Hypotensive Adverse Event

2 Study Design

This is a multicenter randomized controlled clinical trial to evaluate the Aurora Xi New Nomogram for the collection of plasma. The study is to determine whether the overall rate of significant hypotensive adverse events (SHAEs) in donors using the Aurora Xi New Nomogram is less than double the rate of SHAEs in donors using the Aurora Xi Optimized Nomogram. Randomization will be carried out within blocks of variables size (2, 4, or 6) within centers. Donors will be recruited until outcomes are available for a minimum of 51,720 plasmapheresis procedures (approximately 11,519 individual donors) from three BioLife plasma centers. Donors will be randomized in a 1:1 ratio to donate in procedures under the New Nomogram (Test) or the Optimized Nomogram (Control). Donors will provide donations over an 8-week follow-up period. Subjects will be retained in the assigned study arm (Test/Control) throughout their participation in the study period.

2.1 Study Objectives

Primary Objective:

The primary objective of this study is to demonstrate that the overall rate of significant hypotensive adverse events (SHAEs, IQPP DAE Classification 1.2-1.6) in donors using the Aurora Xi New Nomogram algorithm is less than double the SHAE rate in donors using the Aurora Xi Optimized Nomogram algorithm.

Secondary Objectives:

- To determine if the incidence rate of SHAEs per donor status (first-time or repeat donor) observed with the New Nomogram (Test arm) is non-inferior to the incidence rate observed with the Optimized Nomogram (Control arm).
- To determine if the incidence rate of SHAEs for female donors observed with the New Nomogram (Test arm) is non-inferior to the incidence rate observed with the Optimized Nomogram (Control arm).
- To determine if the incidence rate of SHAEs for donors ≤20 years of age observed with the New Nomogram (Test arm) is non-inferior to the incidence rate observed with the Optimized Nomogram (Control arm).
- To determine if the incidence rate of SHAEs for donors weighing ≤124 lbs observed with the New Nomogram (Test arm) is non-inferior to the incidence rate observed with the Optimized Nomogram (Control arm).
- To determine if the incidence rate of hypotensive severe/injury adverse events (IQPP DAE Classification 1.5 or 1.6) observed with the New Nomogram (Test arm) is non-inferior to the incidence rate observed with the Optimized Nomogram (Control arm).
- To determine if the time from start of plasmapheresis procedure to the first SHAE observed with the New Nomogram (Test arm) is non-inferior to the time observed with the Optimized Nomogram (Control arm).

Exploratory Objectives:

- To compare the plasma volume collected with the New Nomogram (Test arm) versus the Optimized Nomogram (Control arm)
- To determine the collection time with the New Nomogram (Test arm)

3 Statistical Methodology

3.1 Sample Size Determination

The sample size requirement is calculated based on a model for correlated binary data in which the marginal mean is modelled with a log-link function and set by

$$log\ \mu_{ij} = \beta_0 + \beta_1 X_i$$

where X_i= 1 for an individual assigned to the experimental arm and X_i= 0 otherwise. Note that center is not controlled for in the sample size calculation. Controlling for center is anticipated to improve power by explaining variation in the response, so this is a conservative approach to estimating the sample size since the power is anticipated to be at least the desired power. For the sample size derivation, the fitted model uses the identity variance function and a working independence assumption with a robust variance derived from a joint model to account for the association in the responses from serial donations from the same donor. The calculations require specification of the mean $\mu_x = E(Y_{ij}|X_i = x)$ and the variance $V_x = \text{var}(Y_{ij}|X_i = x)$ of the response Y_{ij} to procedure j from donor i as well as the within-donor correlation. If κ_x and μ_x is the mean and variance for the number of procedures for a donor in arm x and n is the total number of donors to be recruited, then the asymptotic (large sample) variance (*asvar*) of the estimator of the log relative rate satisfies

$$asvar\left\{\sqrt{n}\left(\hat{\beta}_1 - \beta_1\right)\right\} \simeq \frac{2V_0}{\kappa_0\mu_0^2}\left(1 + \frac{\delta_0\rho}{\kappa_0}\right) + \frac{2V_1}{\kappa_1\mu_1^2}\left(1 + \frac{\delta_1\rho}{\kappa_1}\right)$$

where $\delta_x = v_x + \kappa_x^2 - \kappa_x$ and ρ is the within-donor correlation in the responses over successive procedures from the same donor. The sample size involves computation of an asymptotic robust variance estimate for the estimator based on a multivariate binary probability mass function with specified marginal event rates of SHAE for the Test and Control arms and a specified exchangeable correlation for responses over successive procedures within the same donor, using the conditional linear family (Preisser and Qaqish, 2014) of multivariate binary models. In addition to the response model, a zero-truncated negative binomial model is used for the number of donation procedures over time from each donor which accommodates between-donor heterogeneity in the procedure rate. The number of required donors is computed based on the constraint

$$P\left(\hat{\beta}_1 + 1.96\sqrt{asvar_A\left(\hat{\beta}_1\right)} < \beta_{1M} | \beta_1 = \beta_{1A}\right) = 0.80$$

which ensures that the probability that the upper limit of a two-sided 95% confidence interval for the relative risk is below $RR_M = \exp(\beta_{1M}) = 2$ with 80% power if the true log relative rate is determined by β_{1A}. Based on historical data, the event rate in the control arm is expected to be 0.221% thus setting $p_C = 0.00221$ or equivalently $\log \mu_0 = \beta_0 = \log(0.00221)$. The margin of non-inferiority is specified as $\exp(\beta_{1M}) = 2$. To accommodate a slight elevation in the SHAE rate under the alternative hypothesis of non-inferiority, the sample size calculations will be carried out with $\exp(\beta_{1A}) = 1.2$, corresponding to a 20% relative increase in the SHAE rate (still within the region of non-inferiority). Accommodating a slight increase helps ensure that the power objectives are met to demonstrate non-inferiority even if there is a 20% relative increase in the SHAE rate using

the Aurora Xi New Nomogram. In order to achieve 80% power when $\exp(\beta_{1A})$ equals 1.2, when donations are taken over an 8-week follow-up period, a sample size of n = 11,519 donors may be required. The total number of procedures across all trial participants of $n \times \kappa$ will target 51,720 for the primary endpoint evaluation.

3.2 Analysis Populations

3.2.1 Intent-to-Treat (ITT) Analysis (All Subjects Enrolled)

The Intent-to-Treat Analysis (All Subjects Enrolled) will be based on all subjects who provided informed consent, were randomized, and had at least one procedure initiated during the study. Outcomes for all procedures will be analyzed and responses attributed to the arm to which donors were randomized.

3.2.2 Per-Protocol (PP) Analysis

The Per-Protocol Analysis will restrict attention to procedures in which the device used (Test or Control) matched the arm to which the donor was randomized.

3.3 Methods of Statistical Analysis

Summary statistics (N, mean, standard deviation, median, minimum, and maximum) will be reported on continuous (numeric) parameters for each randomization group for the Intent-to-Treat and Per-Protocol analyses unless otherwise indicated. All summary statistics will be presented at a rounded precision. Summary statistics (e.g., counts and percentages) will be reported on categorical (non-numeric) parameters for each randomization group for the Intent-to-Treat and Per-Protocol analyses unless otherwise indicated. In addition to the Secondary Endpoint analyses, summary statistics will be shown for illustrative purposes for subgroups defined by donor status, gender, age, and weight. All calculations will be performed using SAS 9.4 (Windows platform).

4 Analysis Parameters

4.1 Primary Parameter

The primary objective is to demonstrate non-inferiority of the Aurora Xi New Nomogram compared to the Aurora Xi Optimized Nomogram with respect to the occurrence of a significant hypotensive adverse event (SHAE). Let P_E denote the probability of a SHAE for a procedure using the Aurora Xi New Nomogram algorithm (where E stands for experimental) and let P_C denote the probability of a SHAE for a procedure using the Aurora Xi Optimized Nomogram algorithm (where C stands for Control). The effect of the Aurora Xi New Nomogram compared to the Aurora Xi Optimized Nomogram is expressed as a relative rate $RR = P_E/P_C$ for a SHAE. A non-inferiority margin of $RR_M = 2$ is set to define a non-inferiority region of $[0, RR_M)$ so that any relative rate less than RR_M corresponds to settings where the Aurora Xi New Nomogram is considered non-inferior to the Aurora Xi Optimized Nomogram algorithm. The corresponding hypotheses are

$$H_0 : RR \geq RR_M \text{ (inferiority)} \quad vs. \quad H_A : RR < RR_M \text{ (non-inferiority)}$$

with $RR_M = 2$ so if the null hypothesis (H_0) is rejected in favor of the alternative hypothesis (H_A) then non-inferiority is demonstrated. The test will be carried out based on the confidence interval

approach. If the upper limit of a two-sided 95% confidence interval for RR is below $RR_M = 2$, then non-inferiority will be demonstrated.

4.2 Secondary Parameters

Suppose individual i in a sample of size n provides K_i procedures, $i = 1, ..., n$. Let $Y_{ij} = 1$ if the jth procedure from individual i leads to a SHAE and let $Y_{ij} = 0$ otherwise, $j = 1,..., K_i$, $i = 1,..., n$. Let $X_i = 1$ if individual i is randomized to the experimental arm and $X_i = 0$ otherwise and let V_i be an additional covariate vector which includes indicators for the sites, and let $Z_i = (X_i, V_i)$. Let $\mu_{ij} = E(Y_{ij})$ denote the mean response for the marginal probability of a SHAE given the full covariate vector Z_i. The event rate is expected to be very low, so in order to estimate the relative risk, the log link function in a generalized linear model (McCullagh and Nelder, 2019) is used in the form of

$$\log \mu_{ij} = \beta_0 + \beta_1 X_i + \beta_2 V_i$$

where $\beta = (\beta_0, \beta_1, \beta_2)$ and the relative rate of a SHAE is $\exp(\beta_1)$ when adjusting for the covariates V_i. Analysis will be carried out based on quasi-likelihood with the identity variance function (Lumley et al, 2006).

The variance function is not the binomial variance function and the responses Y_{ij}, $j = 1,..., K_i$ corresponding to the SHAE responses for successive procedures from the same donor are not independent. Therefore, in order to provide protection against a) misspecification of the variance function and b) a working independence assumption for procedures from the same donor, analyses will be based on the generalized estimating equation approach (Liang and Zeger, 1986). A robust variance estimate will be used for inference by defining clusters based on a donor identification variable. Once the above log link model is fitted, the confidence interval approach will be used such that non-inferiority will be claimed if

$$exp\left(\hat{\beta}_1 + 1.96 s.e.(\hat{\beta}_1)\right) < RR_M$$

where $s.e.\left(\hat{\beta}_1\right)$ is the square root of the robust variance estimate.

The following is the list of secondary parameters and how they will be analyzed. If any subgroup target number of donations has not been achieved after the overall sample size has been met, sites will be asked to continue to collect donations over an extended period of follow-up specifically for donors in the target subgroups who are already on-study, and to focus recruitment of new individuals to those of each target subgroup(s) until the minimum number of procedures is met for each subgroup. Each subject in any target subgroup may donate only during their 8-week participation period. If enrollment during the extension period for any of the subgroups is not sufficient to complete the study in a reasonable time period (defined as an extension period of up to 3 months), then the extension period could be stopped at the discretion of the sponsor.

1. <u>Incidence rate of SHAEs per donor status (first-time)</u>
   Analysis will be carried out to estimate the relative risk of significant hypotensive events for first-time donors. First-time donors are defined as individuals who have not donated plasma in the preceding six months. Individuals retain their first-time status only for their first procedure during the study period. Since there is no clustering of responses, a simple robust variance estimate will be used to provide protection against misspecification of the

identity variance function. Once the model is fitted, a point estimate and a 95% confidence interval will be computed for the relative rate of SHAE for first time donors. With a relative rate for a SHAE = 1.2 and the non-inferiority margin of $RR_M = 4$, a sample size of n = 1,329 procedures is required for this subgroup to have 80% power.

2. Incidence rate of SHAEs for female donors
   The log linear model stated above will be fitted for female donors with the same specifications as the primary analysis with a binary procedure-specific response: a log-link, identity variance function, and working independence assumption for procedures from the same donor. Once the model is fitted, a point estimate and a 95% confidence interval based on a robust variance estimate will be computed for the relative rate of SHAE for female donors. With a relative rate for a SHAE = 1.2 and the non-inferiority margin of $RR_M = 2.5$, a sample size of n = 26,366 procedures is required for this subgroup to have 90% power.

3. Incidence rate of SHAEs for donors ≤20 years of age
   The log linear model stated above will be fitted for the subgroup of donors ≤20 years of age with the same specifications as the primary parameter analysis with binary procedure-specific response: log-link, identity variance function, and working independence assumption for procedures from the same donor. Once the model is fitted, a point estimate and a 95% confidence interval based on a robust variance estimate will be computed for the relative of SHAE for donors ≤20 years of age. With a relative rate for a SHAE = 1.2 and the non-inferiority margin of $RR_M = 4$, a sample size of n = 2,514 procedures is required for this subgroup to have 80% power.

4. Incidence rate of SHAEs for donors ≤124 lbs
   The log linear model stated above will be fitted for the subgroup of donors weighing ≤124 lbs with the same specifications as the primary parameter analysis with binary procedure-specific response, log-link, identity variance function, and working independence assumption for procedures from the same donor. Once the model is fitted, a point estimate and a 95% confidence interval based on a robust variance estimate will be computed for the relative of SHAE for donors ≤124 lbs. With a relative rate for a SHAE = 1.2 and the non-inferiority margin of $RR_M = 4$, a sample size of n = 4,827 procedures is required for this subgroup to have 80% power.

5. Incidence rate of hypotensive severe/injury adverse events (IQPP DAE Classification 1.5 or 1.6)
   The log linear model stated above will be fitted for the subgroup of donors to find the incidence rate of severe/injury hypotensive adverse events with the same specifications as the primary parameter analysis with binary procedure-specific response, log-link, identity variance function, and working independence assumption for procedures from the same donor. Once the model is fitted, a point estimate and a 95% confidence interval based on a robust variance estimate will be computed for the relative incidence rate of severe/injury hypotensive adverse events.

6. Time from start of plasmapheresis procedure to the first SHAE

For each procedure, the time (in minutes) from the start of plasmapheresis to the first SHAE will be recorded. For procedures not leading to a SHAE, the time will be censored at the time the plasmapheresis is completed. The cumulative incidence function for the time to the first SHAE for a procedure will be estimated for each arm of the study with a gamma distributed random effect used to accommodate a dependence in the time to a SHAE across procedures within the same donor. To estimate the effect of the device on the time to the first SHAE, a stratified Cox regression model will be fitted with a donor-specific gamma distributed random effect (with a mean of 1 and variance φ to be estimated) and treatment arm (Aurora Xi New Nomogram or the Aurora Xi Optimized Nomogram), with stratification by center. The exponentiated regression coefficient will give the hazard ratio for the time to the first SHAE per procedure.

4.3 Exploratory Parameters

1. Plasma volume collected
   Statistical summaries (mean, median, and range) will be presented for plasma volumes. Additionally, the collected plasma volumes per procedure and over time for the New Nomogram (Test arm) will be compared to the collected plasma volumes per procedure for the Optimized Nomogram (Control arm). The comparison between the Test arm and the Control arm will be based on a t-test.

2. Collection time
   Statistical summaries (mean, median, and range) will be presented for collection time. Additionally, the collection time for the New Nomogram (Test arm) will be compared to the collection time for the Optimized Nomogram (Control arm). The comparison between the Test arm and the Control arm will be based on a t-test.

5 Data Conventions

5.1 Data Collection

See Protocol Section 11.

5.2 Data Verification

See Protocol Section 11.

5.3 Data Monitoring

See Protocol Section 11.

5.4 Data Handling

5.4.1 Handling of Missing Data

No missing or incomplete dates will be imputed for this study. No missing or incomplete parameter results will be imputed for this study.

If missing data arises that affects whether an individual donation can contribute to the analyses described that procedure will be excluded in a complete case analysis.

6 Trial Termination Procedures

See Protocol Section 12.

7 Roll-In Analyses

Upon completion of the roll-in procedures at each site (minimum 150 procedures at each clinical site), summary statistics regarding the incidence of adverse events (category and sub-category) will be evaluated by the DSMB. The summary statistics listing will have the randomization category, collection volume at time of adverse event, total plasma volume, and total collection volume fields hidden from reviewers. If deemed acceptable, the study enrollment will remain unaffected and will continue. Roll-in procedures will be included in the evaluable population.

8 Statistical References

Base SAS 9.4 Procedures Guide, SAS Publishing: Cary, NC

9 Randomization Schedule

See Protocol Section 8.

10 Statistical Code for Hypothesis Testing

SAS code will be written to perform hypothesis testing described in Section 4.1. R code is available to perform the sample size estimation described in Sections 3.1 and 4.2. SAS and R codes will be written to generate output. The SAS and R code will be stored on Fresenius Kabi Medical Affairs shared drive.

11 References

1. Lumley T, Kronmal R and Ma S. Relative risk regression in medical research: models, contrasts, estimators, and algorithms. University of Washington Biostatistics Working Paper Series, 7-19-2006.

2. McCullagh P and Nelder JA. *Generalized Linear Models*. Routledge, 2019.

3. Preisser JS and Qaqish BF. A comparison of methods for simulating correlated binary variables with specified marginal means and correlations. *Journal of Statistical Computation and Simulation*, 84(11), 2441-2452, 2014.

4. Seaman S, Pavlou M and Copas A. Review of methods for handling confounding by cluster and informative cluster size in clustered data. *Statistics in Medicine*, 33(30), 5371-5387, 2014.