STATISTICAL ANALYSIS PLAN

Multicenter External Validation of the APTTO Predictive Model Based on Clot Waveform Analysis for the Assessment of Prolonged Activated Partial Thromboplastin Time

NCT Number: NCT07418099

Version: 1.0

Document date: 9 September 2026

Status: Final

Purpose

This SAP prospectively specifies the principal statistical analyses for the external validation of APTTO (Paper 2) and the perioperative implementation and clinical utility program (Paper 3). It distinguishes primary, key secondary, supportive, sensitivity, and exploratory analyses before final database lock and outcome analysis.

Ethics approval: CEIm-FJD EO012-26_FJD; Protocol Version 2.0, 21 January 2026; favorable opinion signed 4 February 2026.

1. Document control and analysis governance

This SAP is version-controlled. Any substantive modification after Version 1.0 will be documented in a revision history stating the date, rationale, affected analyses, and whether the change was made before or after database lock or examination of outcome results.

This SAP is finalized before the corresponding final analyses. Routine data-quality reviews may be conducted during recruitment to identify missing, inconsistent, or incorrectly derived variables. Such reviews are operational quality-control activities and are not formal interim efficacy or performance analyses. Data-quality corrections made to restore variables to their prespecified definitions will be source-verified, traceable, and will not be selected according to their effect on model performance.

1.1 Revision history

Version	Date	Status	Summary
1.0	9 Sep 2026	Final	Finalized external-validation estimands, reference standards, sample-size/recruitment framework, FJD implementation dates, perioperative endpoints, safety outcomes, and analysis populations.

2. Study program and objectives

APTTO is a sequential clot waveform analysis (CWA)-based decision-support algorithm developed for diagnostic triage of isolated prolonged APTT. The statistical program comprises two publication-level analysis plans with distinct primary questions.

2.1 Paper 2 - External validation

Primary objective: externally validate the prespecified APTTO models without refitting their original coefficients or cut-offs. The principal focus is discrimination, calibration, classification performance at frozen thresholds, clinical utility, and robustness across centers, clinical case-mix, and technical/pre-analytical conditions.

2.2 Paper 3 - Perioperative implementation and clinical utility

Primary conceptual objective: determine whether an APTTO-supported pathway can improve elective perioperative flow by decoupling immediate perioperative decision-making from extensive etiological hemostatic work-up when clinically appropriate, without compromising safety.

Paper 3 comprises three complementary modules: (A) a dedicated prospective multicenter preoperative pathway cohort beginning 1 January 2027; (B) an ambispective FJD pre/post implementation study, with POST implementation beginning 1 October 2026; and (C) an analysis of the incremental functional information provided by non-atypical CWA morphology. A pragmatic micro-costing analysis is a secondary cross-cutting component. FJD participants recruited between 1 October and 31 December 2026 may contribute to both the Paper 2 external-validation cohort and the Paper 3 FJD POST implementation cohort, addressing distinct prespecified questions.

3. General statistical principles

- Two-sided tests will be used with a nominal alpha of 0.05 unless otherwise specified.
- Effect estimates will be reported with 95% confidence intervals. Clinical interpretation will prioritize effect size and precision over isolated P values.
- Continuous variables will be summarized using mean (SD) or median (IQR), as appropriate; categorical variables using n (%).
- No variable selection for primary adjusted models will be based solely on univariable P values.
- Model assumptions will be assessed using appropriate diagnostic procedures. When assumptions are materially violated, prespecified alternative parameterizations or robust methods will be used and documented.

- Center clustering will be accounted for when relevant using cluster-robust variance estimation or mixed-effects models.
- The primary analysis populations and endpoints will not be redefined after examination of final outcome results.

3.1 Multiplicity

Paper 2 has a prespecified registered primary external-validation outcome comprising two model-specific estimands: the AUC of APTTO1 and the AUC of APTTO2 for their respective reference-standard outcomes. No joint hypothesis test or formal multiplicity-adjusted success criterion is defined; interpretation will be based on point estimates and 95% confidence intervals. Secondary and subgroup analyses will be interpreted as supportive unless explicitly designated otherwise. Paper 3 has one primary implementation endpoint; other pathway, safety, morphology, subgroup, and economic analyses are secondary or exploratory. No formal family-wise multiplicity correction is planned for exploratory analyses; exact P values and confidence intervals will be reported and interpreted accordingly.

3.2 Software

All analyses will be performed using R (R Foundation for Statistical Computing, Vienna, Austria). The R version and package versions will be recorded at the time of final analysis and archived with the reproducible analysis scripts.

PART I. PAPER 2 - EXTERNAL VALIDATION OF APTTO

4. Paper 2 analysis populations

4.1 Full external validation population

The full external validation population will include all consecutively recruited participants fulfilling the external-validation protocol eligibility criteria and with a completed, classifiable reference-standard etiological assessment. One participant will contribute one index sample to the primary analysis. Participants with an indeterminate or non-conclusive reference-standard diagnosis will not contribute to primary model-performance estimates but will be counted, characterized, and included in the participant flow.

4.2 APTTO1 evaluable population

Participants with all variables required to calculate the frozen APTTO1 probability and a classifiable APTTO1 reference-standard outcome. No participant will contribute more than one index observation.

4.3 APTTO2 evaluable population

For validation of APTTO2 as a prediction model, the evaluable population will comprise participants with a definitive diagnosis of lupus anticoagulant (LA) or intrinsic-pathway factor deficiency/von Willebrand factor abnormality, with all variables required to calculate the frozen APTTO2 probability. In this model-validation population, factor deficiency/von Willebrand factor abnormality is the positive outcome and LA is the negative outcome. Separately, the clinical performance of the locked sequential APTTO1-to-APTTO2 strategy will be evaluated by applying APTTO2 only to participants routed to the second step by APTTO1 >0.104. These two analyses answer distinct questions and will not be conflated.

4.4 Per-center populations

Center-specific performance will be reported when the number of evaluable participants and outcome events is sufficient to avoid unstable or potentially misleading estimates. Small-center results may be pooled descriptively rather than presented as stand-alone performance estimates.

4.5 Reference-standard definitions

APTTO1: the positive reference-standard outcome is a confirmed LA, an intrinsic-pathway factor deficiency, or von Willebrand disease/reduced von Willebrand factor (VWF). The negative outcome comprises a completed etiological investigation with no abnormality identified and prolonged APTT attributable to anticoagulant treatment. LA will be diagnosed according to applicable ISTH criteria. Intrinsic-pathway factor deficiency will be defined by FVIII, FIX, FXI, or FXII activity below the local laboratory reference interval, irrespective of severity or bleeding phenotype. The VWF category will include established von Willebrand disease, including VWF levels <30%, and reduced VWF levels of

30-50%. These analytical reference-standard definitions will be kept distinct from subsequent clinical characterization of bleeding relevance.

APTTO2: the model-validation reference standard is factor deficiency/VWF abnormality versus LA, using the same diagnostic definitions above. Participants with a normal etiological work-up or anticoagulant-associated APTT prolongation are not part of the APTTO2 model-validation classification problem. Their disposition under the complete sequential algorithm will nevertheless be described.

4.6 Paper 2 recruitment cut-off

The prespecified recruitment cut-off for the Paper 2 external-validation cohort is 31 December 2026. Eligibility for Paper 2 will be determined by the date of the index sample/inclusion. Reference-standard testing, source verification, query resolution, and database cleaning may be completed after this date. Final database lock will occur only after outstanding queries relevant to eligibility, model calculation, and reference-standard classification have been resolved. From 1 January 2027, continued multicenter recruitment for Paper 3 will be restricted to the dedicated prospective preoperative cohort.

5. Frozen APTTO model specification

5.1 APTTO1

Second-derivative CWA morphology is classified using the prespecified four-level system: (A) strictly normal, (B) minimal deviation, (C) shoulder-like, and (D) biphasic/double-peaked. For the frozen binary morphology classification, patterns A-B are non-atypical and patterns C-D are atypical. APTTO1 will be calculated using the originally developed coefficients, variable coding, scaling, and transformations, without coefficient re-estimation in the external cohort:

$L1 = 1.595 + 1.982 \times X1 + 1.207 \times X2 + 0.074 \times X3 - 0.081 \times X4.$

$p1 = 1/[1 + \exp(-L1)].$

X1: non-atypical CWA morphology indicator, coded 1 for patterns A-B (strictly normal or minimal deviation) and 0 for patterns C-D (shoulder-like or biphasic/double-peaked).

X2: strictly normal CWA morphology indicator, coded 1 for pattern A and 0 for patterns B-D.

X3: maximum deceleration parameter, entered per 10,000-unit increment (equivalently, the raw analyzer value is divided by 10,000 before applying the coefficient).

X4: acceleration time, expressed in seconds.

The morphology indicators are hierarchical rather than mutually exclusive: pattern A gives X1 = 1 and X2 = 1; pattern B gives X1 = 1 and X2 = 0; patterns C-D give X1 = 0 and X2 = 0. The prespecified APTTO1 cut-off ($p1 \leq 0.104$) is applied to the final predicted probability p1, not to the linear predictor L1.

5.2 APTTO2

APTTO2 will be calculated using the originally developed equation and its prespecified recalibration, without coefficient re-estimation:

$L2 = 5.472 - 1.519 \times X1 - 0.535 \times X2.$

$R2 = -0.449 + 1.475 \times L2.$

$p2 = 1/[1 + \exp(-R2)].$

X1: non-atypical CWA morphology indicator, coded 1 for patterns A-B and 0 for patterns C-D. This is the same binary morphology variable used as X1 in APTTO1.

X2: acceleration time, expressed in seconds.

The prespecified APTTO2 cut-off (p2 <= 0.112) is applied to the final predicted probability p2 calculated from the recalibrated logit R2; it is not applied to either the original logit L2 or the recalibrated logit R2 itself.

The original coefficients, variable coding, scaling, transformations, recalibration equation, and prespecified cut-offs will be applied without re-estimation or modification during external validation.

5.3 CWA variables and quality control

CWA morphology and quantitative parameters will be derived using the protocol-defined procedures. The four-level morphology classification will retain the categories A = strictly normal, B = minimal deviation, C = shoulder-like, and D = biphasic/double-peaked. The binary morphology classification will retain A-B = non-atypical and C-D = atypical. These representations will not be treated as interchangeable, and model inputs will be coded exactly as specified in Sections 5.1 and 5.2.

Known data-entry or derivation errors will be corrected only when the source record permits reconstruction according to the prespecified definition. In particular, acceleration time must represent the intended interval and not the x-coordinate of the plotted acceleration point. Corrected values will be source-verified and traceable through the query/audit trail. Primary and secondary performance analyses will use the corrected protocol-conformant values; known erroneous values will not be used as an alternative analytic dataset.

6. Paper 2 endpoints

6.1 Primary endpoint

The registered primary external-validation outcome is discrimination of the frozen APTTO models for their prespecified reference-standard outcomes, quantified by model-specific areas under the receiver operating characteristic curve (AUC) with 95% confidence intervals. APTTO1 and APTTO2 AUCs are two prespecified estimands within this registered primary outcome; no joint hypothesis test or multiplicity-adjusted success criterion is defined.

6.2 Secondary validation endpoints

- Calibration-in-the-large and calibration slope.
- Brier score.
- Sensitivity, specificity, negative predictive value, positive predictive value, and corresponding 95% confidence intervals at the frozen cut-offs.
- Observed outcome prevalence within the low-probability and selected-for-further-testing groups.
- Decision-curve analysis across clinically relevant threshold probabilities.
- Interobserver agreement for waveform morphology when replicate readings are available.
- Performance of the locked sequential APTTO1-to-APTTO2 strategy, including routing proportions, final branch distribution, false-negative cases across the complete sequence, and potential reduction or reordering of etiological testing.

6.3 Safety-oriented diagnostic endpoints

Because the clinical purpose is rule-out/triage rather than detection of every analytical abnormality, false-negative cases will be clinically characterized in addition to their reference-standard classification. Analyses will report diagnosis, factor/VWF level where applicable, bleeding phenotype, preoperative context, magnitude of APTT prolongation, and center. Particular attention will be paid to bleeding-associated abnormalities versus abnormalities that prolong APTT without conferring bleeding risk (e.g., FXII deficiency). These analyses will not redefine the primary reference standard.

6.4 Sample-size and recruitment considerations

The external-validation study is designed primarily to estimate performance of the frozen APTTO models with useful precision rather than to test superiority against a null model. The registered primary endpoint remains discrimination (AUC). Sample-size planning also gives particular weight to precision of sensitivity at the frozen APTTO2 threshold because factor deficiency/VWF abnormality is the less frequent target outcome and false-negative classification is clinically important.

In the development cohort, factor deficiency/VWF abnormality represented approximately 26.6% of participants and APTTO2 sensitivity at the frozen threshold was approximately 97.8%. Approximately 200 factor deficiency/VWF outcomes are considered desirable for precise estimation of threshold sensitivity. Assuming the development-cohort prevalence, approximately 760 participants evaluable for etiological classification are expected to yield approximately 200 factor deficiency/VWF outcomes. This constitutes a minimum accrual target based on event-precision considerations rather than a formal minimum sample size for the APTTO2 model-validation population. At this accrual, precision of the registered AUC estimands is also expected to be adequate for external validation. Recruitment is planned to continue through 31 December 2026, with an operational expectation of approximately 1,000 or more participants if accrual permits. Recruitment beyond the event-based target is intended to improve precision of discrimination and calibration estimates and strengthen prespecified heterogeneity analyses; it will not be used for model refitting, coefficient updating, variable selection, or data-driven threshold optimization.

7. Paper 2 statistical methods

7.1 Discrimination

AUCs will be estimated with 95% confidence intervals for APTTO1 and APTTO2 in their respective prespecified model-validation populations. No data-driven re-estimation of model coefficients or cut-offs will be undertaken. Heterogeneity analyses will emphasize model-specific estimates and confidence intervals rather than data-driven optimization or formal claims of subgroup superiority.

7.2 Calibration

Calibration will be assessed using calibration-in-the-large, calibration slope, graphical calibration plots, and Brier score. Flexible calibration curves may be added as supportive analyses if sample size permits. Any post-hoc recalibration will be clearly labeled exploratory and will not replace performance of the frozen model.

7.3 Classification metrics

Sensitivity, specificity, PPV and NPV at the prespecified thresholds will be calculated with 95% confidence intervals. Exact or Wilson-type binomial intervals will be used as appropriate. Counts and clinical details of false-negative cases will be reported.

7.4 Decision-curve analysis

Net benefit of the frozen model will be evaluated across threshold ranges considered clinically interpretable for the intended triage use and compared with default strategies such as investigating all and investigating none. The displayed range and its clinical rationale will be reported transparently and will not be selected by maximizing observed net benefit.

7.5 Interobserver morphology agreement

Agreement for categorical waveform morphology will be assessed using Cohen's kappa for two raters or an appropriate multi-rater kappa when applicable. Weighted kappa will be used for ordered detailed morphology categories if the ordering is clinically defensible. Percent agreement will also be reported.

7.6 Prespecified heterogeneity and transportability analyses

- Clinical/case-mix heterogeneity: participating center.
- Clinical/case-mix heterogeneity: preoperative versus non-preoperative context, where applicable.
- Clinical/case-mix heterogeneity: anticoagulation status where allowed by the Paper 2 protocol.
- Clinical/case-mix heterogeneity: magnitude of APTT prolongation and major etiological category, where estimable.
- Technical/pre-analytical heterogeneity: ACL TOP analyzer model, SynthASil reagent lot, and sample processing time (≤3 hours versus >3 hours from collection), where denominators and outcome counts permit stable estimation.
- These analyses are prespecified supportive assessments of transportability and technical robustness. They are not powered for formal subgroup comparisons, and statistical significance or non-significance will not be interpreted as evidence of equivalence or differential model validity. Estimates will be presented with confidence intervals when sample size and outcome counts permit. Reagent lot, analyzer model, and center may be

correlated; observed heterogeneity will therefore be interpreted cautiously and not attributed causally to a technical factor without supporting evidence.

No additional subgroup will be promoted to prespecified status after final database lock. Any additional subgroup analysis will be explicitly labeled exploratory/post hoc.

8. Paper 2 missing data and protocol deviations

The amount and pattern of missingness will be reported for each model variable and reference-standard outcome, overall and by center where informative. Missing or indeterminate reference-standard outcomes will not be imputed. If a variable required by the frozen model cannot be reconstructed according to its prespecified definition, that participant will be non-evaluable for the corresponding model-performance analysis. The primary validation will therefore use complete model inputs rather than introducing a new imputation rule not present in the locked model. Characteristics of evaluable and non-evaluable participants will be compared descriptively when numbers justify it.

Major protocol deviations affecting eligibility, reference-standard ascertainment, or model calculation will be listed before database lock. A protocol-conformant sensitivity population may be analyzed in addition to the primary pragmatic validation population.

PART II. PAPER 3 - PERIOPERATIVE IMPLEMENTATION AND CLINICAL UTILITY

9. Paper 3 Module A - Prospective multicenter perioperative pathway cohort

9.1 Population

From 1 January 2027, participating centers will continue prospective recruitment restricted to preoperative patients with isolated prolonged APTT for the dedicated Paper 3 multicenter pathway cohort. This cohort is specifically designed for Paper 3 and is analytically distinct from the Paper 2 external-validation cohort defined by an index-sample cut-off of 31 December 2026.

Eligibility will require a preoperative surgical assessment, prolonged APTT, normal PT, and absence of a clinically relevant personal bleeding history. The same eligibility framework and the same standardized perioperative variables and pathway dates will be used across participating centers.

9.2 Objective

The principal objective is to quantify the real-world time and resource burden of resolving isolated prolonged APTT in elective perioperative care and to characterize between-center organizational variability.

9.3 Outcomes

The dedicated Paper 3 multicenter CRF will prospectively standardize the same pathway dates used at FJD: T0 index prolonged APTT; T1 referral/activation of the hematological pathway; T2 first Hematology assessment; T3 diagnostic completion sufficient for perioperative decision-making; T4 hematological clearance; and T5 surgery. The local mechanism of pathway activation/referral will additionally be documented so that T1 has a functionally comparable meaning across centers. Planned surgical or invasive procedures will be mapped to the three-tier ISTH SSC procedural bleeding-risk framework described by Shaw et al. (minimal, low/moderate, or high bleeding risk) to standardize procedure-related risk classification across centers. This use of the framework is for procedural classification and does not imply that study participants are receiving anticoagulant therapy.

9.4 Analysis

Times will be summarized overall and by center. Mixed-effects regression or cluster-robust methods will account for patients nested within centers. When appropriate, the intraclass correlation coefficient will quantify the proportion of variability attributable to center-level differences. Center-specific estimates will be displayed only when denominators are adequate.

10. Paper 3 Module B - FJD pre/post implementation study

10.1 Design and populations

The FJD implementation analysis is an ambispective before-and-after study. The PRE cohort comprises eligible patients managed conventionally from 1 January 2022 through 30 September 2026. The POST cohort will prospectively include consecutive eligible patients from 1 October 2026, the formal date of incorporation of APTTO into routine FJD perioperative clinical decision-making. The primary implementation estimand follows an intention-to-implement principle: eligible patients managed after implementation will remain in the POST group irrespective of individual adherence to the APTTO-supported recommendation. Adherence and reasons for deviation will be analyzed separately.

Historical cases are identified through systematic review of outpatient schedules of hematologists dedicated to hemostasis, followed by individual electronic medical-record confirmation. The same clinical eligibility framework will be applied to PRE and POST cohorts.

10.2 Eligibility

- Adult patient undergoing elective surgery.
- Isolated prolonged APTT identified during preoperative assessment.
- Normal PT.
- Referral from Anesthesiology to Hematology specifically because of prolonged APTT.
- No clinically relevant personal bleeding history.
- Other protocol-defined conditions that make the APTTO perioperative pathway non-applicable will be documented and handled according to the prespecified eligibility rules; exclusions will not be introduced on the basis of observed outcomes.

10.3 Baseline variables

- Age and sex.
- ASA physical status.
- APTT in seconds and APTT ratio.
- PT/ratio.
- Surgical specialty.
- Exact planned surgical procedure.
- Procedural bleeding risk: minimal, low/moderate, or high, according to the prespecified ISTH SSC procedural bleeding-risk framework of Shaw et al.
- CMA (ambulatory major surgery) status.
- Calendar year/period.

Final etiological diagnosis will be recorded for characterization but will not be treated as a baseline confounder in the primary causal model because it is determined after entry into the pathway.

11. FJD pathway time points and endpoint definitions

- T0 - Index prolonged APTT: date of the result prompting perioperative hematological evaluation.
- T1 - Referral: date of referral from Anesthesiology to Hematology.
- T2 - First Hematology assessment: first assessment specifically addressing prolonged APTT.
- T3 - Diagnostic completion: first date on which sufficient hemostatic information was available for perioperative decision-making.
- T4 - Hematological clearance: date on which the Hematologist signs the clinical report documenting hematological clearance for the planned surgery, without requirement to await further immediate etiological work-up.
- T5 - Surgery: date on which the planned procedure was actually performed.

11.1 Operational definition of T3

PRE: T3 is the date on which the final result required by the treating hematologist to complete the conventional perioperative assessment became available.

POST low-probability APTTO pathway: T3 is the date on which the APTTO result became available and interpretable for perioperative decision-making, allowing hematological clearance without waiting for completion of deferred etiological investigation when clinically appropriate.

POST requiring further work-up: T3 is the date on which the last additional result required for perioperative decision-making became available.

Etiological investigations intentionally deferred until after hematological clearance do not modify T3.

11.2 Primary implementation endpoint

Time from Anesthesiology referral to hematological clearance (T1-T4).

11.3 Key secondary endpoint

Time from Anesthesiology referral to actual surgery (T1-T5). Because the originally planned surgical date cannot be reliably recovered in the historical cohort, this endpoint measures overall perioperative flow and will not be interpreted as a direct measure of delay attributable exclusively to Hematology.

11.4 Secondary process endpoints

- T0-T1, T1-T2, T1-T3, T2-T3, T3-T4, and T4-T5 intervals.
- Number and type of hemostatic investigations.
- Number of Hematology assessments.
- Repeat laboratory testing and repeat sampling when reliably ascertainable.
- Surgical cancellation/reprogramming associated with unresolved hemostatic assessment when reliably documented.
- Direct diagnostic laboratory/resource cost.
- Length of stay if clinically meaningful and reliably available.

T3-T4 will be specifically analyzed as an internal process measure to distinguish faster diagnostic completion from faster clinician issuance of clearance.

12. Perioperative management and safety outcomes

12.1 Management

- Prophylactic plasma.
- Desmopressin.
- Factor concentrate.
- Other prophylactic hemostatic treatment.
- Red-cell or other blood-component transfusion related to perioperative bleeding/hemostatic management, with timing categorized as intraoperative, 0-24 h postoperatively, and >24 h to day 30.

12.2 Safety

- Perioperative major bleeding according to the ISTH SSC surgical definition, assessed from the start of surgery through postoperative day 30.
- Unplanned transfusion attributable to bleeding, with timing categorized as intraoperative, 0-24 h, and >24 h to day 30.
- Reintervention because of bleeding.
- Postoperative Hematology assessment prompted by unexpected bleeding.
- Subsequent diagnosis of a clinically relevant bleeding disorder established during the deferred etiological work-up associated with the index preoperative episode and not recognized at the time of preoperative hematological clearance.

Symptomatic objectively confirmed thrombotic event through postoperative day 30, including venous or arterial thrombosis.

A prespecified composite safety outcome will be reported as the occurrence of any major bleeding, symptomatic objectively confirmed thrombosis, reintervention for bleeding, unplanned transfusion attributable to bleeding, or

unplanned therapeutic hemostatic treatment prompted by perioperative bleeding. Each component will also be reported separately. The composite is descriptive/exploratory, is not intended as a validated composite clinical endpoint, and will not replace component-specific safety interpretation.

Perioperative major bleeding will be defined according to the ISTH SSC definition for surgical patients (Schulman et al., J Thromb Haemost. 2010;8:202-204) and assessed from the start of surgery through postoperative day 30. Exact event dates will be retained. Hemostatic treatment and transfusion will be recorded with timing (intraoperative, 0-24 h, and >24 h to day 30). Thrombotic events will require symptoms and objective confirmation and will be assessed through day 30. APTTO will be interpreted as a clinical decision-support strategy, not as a guarantee of absence of bleeding or thrombotic risk. No formal non-inferiority claim will be made unless sample size and event rates provide adequate power.

13. APTTO-specific POST variables

- APTTO1 probability and classification.
- APTTO2 probability and classification when applicable.
- Recommended pathway.
- Actual pathway followed.
- Additional testing performed despite low-probability classification.
- Testing avoided or deferred.
- Concordance between recommendation and management.
- Reason for deviation.

APTTO will not be retrospectively reconstructed or imputed in PRE patients because the required CWA parameters were not systematically available.

14. FJD primary statistical analysis

14.1 Descriptive comparison and balance

PRE and POST baseline characteristics will be described using standardized mean differences. Conventional P values may be shown descriptively but will not determine covariate inclusion.

14.2 Primary implementation estimand and analysis

The primary estimand is the adjusted between-period difference in time from T1 to T4, expressed in days. Absolute adjusted differences will be prioritized for clinical interpretation. The statistical model will be selected according to the observed distribution and censoring structure using a prespecified outcome-blinded assessment of model assumptions. When censoring is material, appropriate time-to-event methods will be used; alternatives may include accelerated failure-time models or restricted mean time methods. Kaplan-Meier curves may be used descriptively when appropriate.

Median times, interquartile ranges, and adjusted absolute differences in days with 95% confidence intervals will be reported. Relative time-to-event estimates may be reported as supportive measures when clinically and statistically appropriate.

14.3 Primary adjusted model

Adjustment will use prespecified variables available before clinical decision-making. Candidate variables are age, sex, ASA, baseline APTT ratio, surgical specialty/procedure, procedural bleeding risk according to the prespecified ISTH SSC three-tier classification, CMA status, and calendar time. Final diagnosis, number of tests and T3 will not be adjusted for as baseline confounders because they are post-exposure variables and may mediate the effect of implementation.

15. Propensity-score strategy

Because PRE/POST assignment is non-randomized, stabilized inverse probability of treatment weighting (IPTW) will be the primary adjustment strategy provided that prespecified diagnostics demonstrate adequate overlap, positivity,

weight behavior, and effective sample size. The planned approach is to inspect stabilized weights and, when needed, truncate at the 1st and 99th percentiles to limit undue influence of extreme observations.

- The propensity model will contain prespecified baseline variables and will not include post-exposure variables.
- Propensity-score distributions and common support will be inspected.
- Stabilized weights will be used where appropriate.
- Extreme weights and positivity violations will be assessed before outcome-model interpretation. Untruncated weights will be retained when weight behavior is acceptable; 1st/99th percentile truncation is the planned corrective strategy when needed. Alternative truncation (e.g., 5th/95th percentiles) may be examined as a sensitivity analysis and will be reported transparently.
- Post-weighting balance will be assessed using standardized mean differences; |SMD| <0.10 will generally indicate acceptable balance. Effective sample size will also be reported.
- A Love plot will display covariate balance before and after weighting.
- Propensity-score matching may be performed as a sensitivity analysis.

If severe lack of common support, marked positivity violations, or unacceptable effective sample size make IPTW unreliable, the prespecified multivariable adjusted regression analysis will provide the principal adjusted estimate and the reason for not relying on IPTW will be documented. Propensity-score matching may be used as a sensitivity analysis.

16. Difference-in-differences analysis

If sufficient contemporaneous longitudinal data are available from centers that do not implement APTTO and outcomes/pathway definitions are comparable, a Difference-in-Differences (DiD) analysis will be performed as a supportive quasi-experimental analysis to assess implementation-associated changes beyond secular trends.

The key estimand will be the interaction between intervention group (FJD versus non-implementing controls) and POST period. Appropriate generalized linear, survival, or transformed-time models will be selected according to the outcome distribution.

The parallel-trends assumption will be evaluated using available pre-implementation data and graphical approaches when the number and distribution of time points permit. DiD results will be interpreted as supportive and only when comparability and parallel-trend assumptions are reasonably credible.

17. Temporal trends, CMA and sensitivity analyses

- PRE outcomes will be examined across 2022, 2023, 2024, 2025, and January-September 2026 to identify pre-existing secular trends before the 1 October 2026 implementation date.
- No interrupted time-series analysis is prespecified. Pre-implementation temporal patterns will instead be described to assess secular change and to inform interpretation of the primary PRE/POST analysis and any supportive DiD analysis.
- CMA status will be included as a prespecified baseline covariate where appropriate. Effect heterogeneity by CMA status may be explored using a period-by-CMA interaction if sample size and outcome distribution are adequate.
- Effect heterogeneity by procedural bleeding risk (minimal, low/moderate, or high according to the prespecified ISTH SSC classification) may similarly be explored if sample size and outcome distribution are adequate. These interaction analyses are supportive/exploratory.
- A POST per-protocol/as-treated supportive analysis will be restricted to participants managed in concordance with the predefined APTTO-supported recommendation. Participants whose clinicians elect to await additional testing despite an APTTO recommendation permitting earlier clearance will remain in the POST group for the primary intention-to-implement analysis; adherence and documented reasons for non-adherence will be reported.
- Complete-case analyses for essential dates, IPTW versus unweighted estimates, and matching-based estimates will be used as sensitivity analyses where appropriate.

18. Paper 3 Module C - Incremental value of CWA morphology

18.1 Objective and population

This analysis will address a question distinct from Paper 2 validation: whether CWA morphology provides functional information regarding the clinical relevance and severity of abnormalities underlying prolonged APTT. The principal analysis population will comprise participants in the Paper 2 external-validation cohort with evaluable waveform morphology and definitive etiological characterization. Participants in the prospective Paper 3 preoperative cohort meeting the same analytical definitions may be analyzed separately as a supportive replication cohort.

18.2 Morphology definitions

The principal exposure is dichotomous non-atypical versus atypical waveform morphology, defined as patterns A-B versus C-D, respectively. The detailed categories - A, strictly normal; B, minimal deviation; C, shoulder-like; and D, biphasic/double-peaked - will be analyzed separately and will not be conflated with the dichotomous classification.

18.3 Outcomes

- Presence and type of factor deficiency.
- Residual factor activity as a continuous variable.
- Presence of a prespecified clinically relevant bleeding-associated hemostatic abnormality.
- Distribution of etiological diagnoses according to morphology.
- Relationship between APTT ratio, morphology and factor activity.

For the principal prespecified morphology analysis, a clinically relevant bleeding-associated hemostatic abnormality will be defined a priori as FVIII or FIX activity <40 IU/dL; FXI activity <25 IU/dL; VWF antigen and/or platelet-dependent VWF activity <0.30 IU/mL; confirmed type 2 or type 3 von Willebrand disease; or another established hemostatic disorder requiring specific perioperative management. These are prespecified operational study definitions informed by published disease-classification and bleeding-phenotype evidence and are not intended as absolute biological thresholds of perioperative bleeding risk. The FVIII/FIX threshold is aligned with the established mild hemophilia range in the WFH guidelines (Srivastava et al., 2020), in which factor levels of 5-40 IU/dL may be associated with severe bleeding during major trauma or surgery. The FXI <25 IU/dL threshold is an operational cut-off informed by EN-RBD data (Peyvandi et al., 2012), which showed poor correlation between FXI activity and overall bleeding severity but identified approximately 26 IU/dL as the activity associated with asymptomatic status and <25 IU/dL with Grade III bleeding; therefore, bleeding history and procedural context remain essential for FXI interpretation. For VWF, the 2021 ASH/ISTH/NHF/WFH diagnostic guideline (James et al., 2021) supports <0.30 IU/mL regardless of bleeding to confirm type 1 VWD and <0.50 IU/mL in patients with abnormal bleeding. Accordingly, VWF 0.30-0.50 IU/mL and FXI 25-39 IU/dL will be treated as intermediate/context-dependent findings rather than automatically classified as bleeding-associated abnormalities. Factor XII deficiency and isolated lupus anticoagulant will not be classified as bleeding-associated abnormalities; the absence of a bleeding phenotype in FXII deficiency is supported by established hemostasis literature (Stavrou and Schmaier, 2010). Factor and VWF measures will also be analyzed continuously so that the principal analysis does not depend solely on categorical thresholds.

18.4 Incremental-value models

The principal analysis will characterize the relationship between waveform morphology and residual factor/VWF activity, retaining factor activity as continuous whenever possible. A base model using APTT ratio may be compared with an extended model adding waveform morphology using change in discrimination/AUC, likelihood-ratio testing, and calibration where appropriate. Clinically motivated thresholds such as 30% may be used for prespecified supportive analyses but will not replace continuous analysis. Decision-curve analysis is not prespecified for this morphology module.

The emerging local observation that non-atypical morphology may be associated with preserved residual factor activity will be treated as a hypothesis to be tested and will not be used to define the outcome or choose thresholds.

19. Micro-costing analysis

19.1 Scope

The economic analysis is secondary and pragmatic. It will quantify direct diagnostic resource use rather than attempt a full health-economic evaluation. Its purpose is to complement the primary perioperative-flow analysis.

19.2 Components

- Local reagent costs for repeat APTT, mixing studies, FVIII/FIX/FXI/FXII, VWF testing, lupus anticoagulant testing and other locally used investigations.
- Technical hands-on time for sample preparation, mixing studies, factor panels, lupus anticoagulant testing and technical post-analytical work.
- Physician time for simple/complex validation, integrated interpretation, further-test selection and clinically relevant communication.

19.3 Time measurement

Continuous patient-level time-and-motion observation will not be required. Centers will perform repeated representative measurements of diagnostic modules, with the number of replicates determined pragmatically according to workflow variability and feasibility. Mean hands-on time will be multiplied by the relevant personnel cost per unit time.

19.4 Cost calculation and analysis

Direct diagnostic cost = reagent cost + technical personnel cost + physician personnel cost. Costs will be summarized by center and diagnostic pathway. Conventional versus APTTO-supported strategies will be compared using observed resource use in implemented cohorts and clearly labeled modeled/projection analyses when the APTTO pathway has not been directly implemented. Because cost distributions are commonly skewed, non-parametric bootstrap confidence intervals or suitable generalized linear models will be considered.

20. Missing data - Paper 3

Missingness will be quantified by cohort and variable. Dates defining primary time outcomes will not be imputed when they cannot be reliably established. Multiple imputation may be considered for missing baseline covariates if the amount and missingness pattern justify it. The imputation model will include exposure, outcome-related information and relevant baseline covariates but will not be used to manufacture unobserved pathway dates.

The extent of missingness in historical PRE data will be explicitly compared with POST data, and sensitivity analyses will assess whether conclusions are robust to differential completeness.

21. Analysis hierarchy and interpretation

Paper 2 primary external-validation interpretation will focus on performance of the frozen APTTO models in external data. Paper 3 primary implementation interpretation will focus on the adjusted absolute difference in T1-T4 time between FJD PRE and POST periods under the prespecified adjustment strategy. The multicenter pathway, T1-T5, resource, safety, morphology, DiD, subgroup and micro-costing analyses are supportive or exploratory unless otherwise explicitly designated in this SAP.

The central Paper 3 clinical interpretation is that APTTO implementation may be associated with improved perioperative workflow by allowing hematological clearance to be decoupled from completion of extensive diagnostic work-up in selected low-risk patients. Causal language will be used cautiously in this observational PRE/POST design; estimates will be described as implementation-associated effects unless assumptions supporting stronger causal interpretation are met.

22. Tables and figures planned before analysis

22.1 Paper 2

- Participant flow diagram by center and analysis population.
- Baseline/etiological characteristics.

- ROC curves and AUCs for APTTO1/APTTO2.
- Calibration plots.
- Classification-performance table at frozen cut-offs.
- Decision curves.
- False-negative clinical characterization.
- Center/subgroup forest plot where estimable.

22.2 Paper 3

- Multicenter pathway-time distribution and center variability plot.
- FJD PRE/POST participant flow and baseline balance table.
- Love plot before/after IPTW.
- Absolute time-distribution plots for T1-T4 and T1-T5, with time-to-event plots added when censoring is material.
- Absolute pathway-time decomposition (T1-T3, T3-T4, T4-T5).
- Resource-use and direct-cost comparison.
- DiD figure if the supportive analysis is feasible and assumptions are reasonably credible.
- Morphology versus factor-activity distribution and incremental model performance.
- Safety outcomes table.

23. Key methodological references

Gamble C, Krishan A, Stocken D, et al. Guidelines for the Content of Statistical Analysis Plans in Clinical Trials. JAMA. 2017;318(23):2337-2343. doi:10.1001/jama.2017.18556.

Velasco-Rodríguez D, Revilla N, Mahíllo-Fernández I, et al. APTTO: a clot waveform analysis-based algorithm for diagnostic triage of prolonged activated partial thromboplastin time. Thromb Res. 2026;262:109731. doi:10.1016/j.thromres.2026.109731.

Riley RD, Debray TPA, Collins GS, Archer L, Ensor J, van Smeden M, Snell KIE. Minimum sample size for external validation of a clinical prediction model with a binary outcome. Stat Med. 2021;40(19):4230-4251. doi:10.1002/sim.9025.

Devreese KMJ, Bertolaccini ML, Branch DW, et al. An update on laboratory detection and interpretation of antiphospholipid antibodies for diagnosis of antiphospholipid syndrome: guidance from the ISTH-SSC Subcommittee on Lupus Anticoagulant/Antiphospholipid Antibodies. J Thromb Haemost. 2025;23(2):731-744. doi:10.1016/j.jtha.2024.10.022.

Shaw JR, Abdulrehman J, Frosi Stella S, et al. Surgical and procedural bleed risk stratification for anticoagulated patients undergoing planned surgery: guidance from the ISTH SSC Subcommittee on Perioperative and Critical Care Thrombosis and Hemostasis. J Thromb Haemost. 2026;24(7):2640-2655. doi:10.1016/j.jtha.2025.12.031.

Schulman S, Angerås U, Bergqvist D, Eriksson B, Lassen MR, Fisher W; Subcommittee on Control of Anticoagulation of the Scientific and Standardization Committee of the International Society on Thrombosis and Haemostasis. Definition of major bleeding in clinical investigations of antihemostatic medicinal products in surgical patients. J Thromb Haemost. 2010;8(1):202-204. doi:10.1111/j.1538-7836.2009.03678.x.

Srivastava A, Santagostino E, Dougall A, et al. WFH Guidelines for the Management of Hemophilia, 3rd edition. Haemophilia. 2020;26(Suppl 6):1-158. doi:10.1111/hae.14046.

Makris M, Oldenburg J, Mauser-Bunschoten EP, et al. The definition, diagnosis and management of mild hemophilia A: communication from the SSC of the ISTH. J Thromb Haemost. 2018;16(12):2530-2533. doi:10.1111/jth.14315.

James PD, Connell NT, Ameer B, et al. ASH ISTH NHF WFH 2021 guidelines on the diagnosis of von Willebrand disease. Blood Adv. 2021;5(1):280-300. doi:10.1182/bloodadvances.2020003265.

Peyvandi F, Palla R, Menegatti M, et al.; European Network of Rare Bleeding Disorders Group. Coagulation factor activity and clinical bleeding severity in rare bleeding disorders: results from the European Network of Rare Bleeding Disorders. J Thromb Haemost. 2012;10(4):615-621. doi:10.1111/j.1538-7836.2012.04653.x.

Stavrou E, Schmaier AH. Factor XII: what does it contribute to our understanding of the physiology and pathophysiology of hemostasis & thrombosis. Thromb Res. 2010;125(3):210-215. doi:10.1016/j.thromres.2009.11.028.

24. Sign-off

Role	Name / signature	Date
Principal Investigator		
Study Statistician		
Additional scientific reviewer		