A Retrospective Cohort Study of Case-Based Scenario Simulation Teaching in Cardiopulmonary Resuscitation Training for Medical and Nursing Interns

Statistical Analysis Plan (SAP)

NCT Number:	NCT2026-0693
Document Date:	2025-08-25
Version:	1.0

1. PURPOSE AND SCOPE

This Statistical Analysis Plan (SAP) describes the statistical methods and procedures to be applied in the retrospective cohort study comparing case-based scenario simulation teaching (CBSST) versus traditional lecture-based teaching (LBT) in cardiopulmonary resuscitation (CPR) training for medical and nursing interns. This SAP is prepared before database lock and data analysis to ensure objectivity and reproducibility.

2. STUDY DESIGN AND OBJECTIVES

2.1 Design
Single-center, two-arm retrospective cohort study.

2.2 Primary Objectives
• To compare theoretical examination scores between LBT and CBSST groups.
• To compare Mini-CEX total scores between LBT and CBSST groups.

2.3 Secondary Objectives
• To compare Mini-CEX dimension sub-scores (knowledge mastery, operational skills, clinical judgment, team collaboration).
• To compare teaching satisfaction scores across four dimensions.
• To compare self-perceived learning burden between groups.

3. ENDPOINT DEFINITIONS

3.1 Primary Endpoints
• Theoretical Examination Score: Continuous variable, range 0–100, measured immediately post-training via standardized closed-book examination.
• Mini-CEX Total Score: Continuous variable, range 4–36, sum of four 9-point dimension scores, measured immediately post-training by blinded assessors.

3.2 Secondary Endpoints
• Mini-CEX Dimension Scores: Each continuous, range 1–9.
• Teaching Satisfaction Dimensions (course content, resources, interaction, methods): Each ordinal (treated as continuous), range 1–5.
• Learning Burden: Ordinal (treated as continuous), range 1–5.

3.3 Covariates
• Gender (categorical: male, female);
• Age (continuous, years);
• Professional Category (categorical: medical student, nursing student);

• Rotation Department (categorical);
• Training Date (continuous, used to account for temporal trends).

4. ANALYSIS POPULATIONS

4.1 Full Analysis Set (FAS)
All enrolled participants who meet inclusion criteria and have at least one post-training assessment record (theoretical or Mini-CEX). This is the primary analysis population.

4.2 Per-Protocol Set (PPS)
A subset of FAS excluding participants with:
• Attendance < 90% but still meeting minimum threshold;
• Major protocol deviations (e.g., crossover between teaching modalities, incomplete training);
• Missing data > 20% on primary endpoints.
Sensitivity analyses will compare FAS and PPS results.

4.3 Safety Population
Not applicable; this is an educational retrospective study with no clinical intervention or adverse event monitoring.

5. HANDLING OF MISSING DATA

5.1 Missing Data Assessment
The pattern and mechanism of missing data will be evaluated (Missing Completely At Random [MCAR], Missing At Random [MAR], or Missing Not At Random [MNAR]).

5.2 Primary Approach
Complete-case analysis (listwise deletion) for the primary endpoints, given the small sample size and the expectation of minimal missing data (< 5%).

5.3 Sensitivity Analyses
• Multiple Imputation by Chained Equations (MICE) under MAR assumption for primary endpoints if missing data exceeds 5%;
• Best-case / worst-case scenario imputation for secondary endpoints;
• Pattern-mixture models if MNAR is suspected.

5.4 Missing Covariates
For missing baseline covariates (gender, age, professional category), multiple imputation will be used if > 5% missing; otherwise, a separate "unknown" category will be created for categorical variables or median imputation for continuous variables.

6. STATISTICAL METHODS

6.1 Descriptive Statistics

• Continuous variables: Mean, standard deviation (SD), median, interquartile range (IQR), minimum, and maximum. Normality assessed via Shapiro-Wilk test and visual inspection (Q-Q plots, histograms).
• Categorical variables: Frequencies and percentages.
• Baseline characteristics tabulated by study group; no formal statistical testing for baseline comparability in retrospective designs (descriptive only).

6.2 Primary Endpoint Analysis

6.2.1 Theoretical Examination Score

Independent samples t-test if normality and homogeneity of variance assumptions hold (Levene's test, $P > 0.05$). If violated, Welch's t-test or Mann-Whitney U test will be used. Effect size: Cohen's d with 95% confidence interval (CI); Hedges' g correction applied.

6.2.2 Mini-CEX Total Score

Same approach as above. Additional analysis of covariance (ANCOVA) adjusting for age, gender, and professional category if baseline imbalances are observed (standardized difference > 0.1).

6.3 Secondary Endpoint Analysis

6.3.1 Mini-CEX Dimension Scores

Independent t-test or Mann-Whitney U test per dimension. Bonferroni correction applied: adjusted $\alpha = 0.0125$ (0.05/4) for four comparisons.

6.3.2 Teaching Satisfaction Dimensions

Mann-Whitney U test preferred given Likert-scale ordinal nature; if distributions are approximately normal, independent t-test may be used. Bonferroni correction: adjusted $\alpha = 0.01$ (0.05/5) for five satisfaction/burden dimensions.

6.3.3 Learning Burden

Independent t-test or Mann-Whitney U test; included in the Bonferroni-corrected satisfaction analysis family.

6.4 Subgroup Analyses

Planned subgroups (if sample size permits, $n \geq 10$ per subgroup):
• Professional category (medical students vs. nursing students);
• Gender (male vs. female);
• Age tertiles (≤ 21, 22, ≥ 23 years).
Subgroup analyses will use the same statistical tests as primary analyses within each

stratum. Interaction terms (teaching modality × subgroup factor) will be tested in regression models to assess effect modification.

6.5 Multivariable Analyses
Multiple linear regression for primary and secondary continuous outcomes, with teaching modality as the independent variable and age, gender, professional category, and rotation department as covariates. Model diagnostics include residual analysis, multicollinearity assessment (VIF < 5), and influence statistics (Cook's distance).

6.6 Sensitivity Analyses
• Per-protocol analysis vs. full analysis set;
• Exclusion of outliers (values > 3 SD from group mean) with documentation;
• Non-parametric alternatives for all primary comparisons;
• ANCOVA adjusting for pre-training baseline if any baseline knowledge assessment data are retrospectively available.

7. MULTIPLICITY CONTROL

To control family-wise Type I error inflation:
• Mini-CEX four dimensions: Bonferroni-adjusted $\alpha = 0.0125$;
• Satisfaction and burden five dimensions: Bonferroni-adjusted $\alpha = 0.01$;
• Primary endpoints (theoretical score, Mini-CEX total): No adjustment (two independent primary hypotheses);
• Secondary endpoints are exploratory; nominal P-values reported with adjusted significance thresholds noted.

If hierarchical testing is preferred, a fallback procedure will be applied: theoretical score → Mini-CEX total → Mini-CEX dimensions → satisfaction dimensions.

8. EFFECT SIZE AND CONFIDENCE INTERVALS

Cohen's d calculation:
$d = (M_1 - M_2) / S$
where $S = \sqrt{[((n_1 - 1)SD_1^2 + (n_2 - 1)SD_2^2) / (n_1 + n_2 - 2)]}$

Hedges' g correction:
$g = d \times [1 - (3 / (4(n_1 + n_2) - 9))]$

95% CI for mean difference: independent t-test CI output.
95% CI for Cohen's d: calculated using the standard error approximation: $SE = \sqrt{[(n_1 + n_2) / (n_1 \times n_2) + d^2 / (2(n_1 + n_2))]}$.

Interpretation:
• d = 0.2: small effect
• d = 0.5: medium effect
• d = 0.8: large effect

9. INTERIM ANALYSES

Not applicable. This is a retrospective study with a fixed, complete dataset; no interim analyses are planned.

10. SOFTWARE AND REPRODUCIBILITY

• Statistical Analysis: SPSS version 23.0 (IBM Corp., Armonk, NY, USA);
• Effect Size and Power: G*Power 3.1 (Heinrich-Heine-Universität Düsseldorf, Germany);
• Data Management: Microsoft Excel 2019 (verification and cross-checking);
• All syntax files, output logs, and annotated datasets will be archived for a minimum of 5 years post-publication to ensure reproducibility.

11. TABLES AND LISTINGS

Planned outputs:
Table 1: Baseline Demographic and Clinical Characteristics by Teaching Modality
Table 2: Comparison of Theoretical Examination Scores Between Groups
Table 3: Comparison of Mini-CEX Total and Dimension Scores Between Groups
Table 4: Comparison of Teaching Satisfaction and Learning Burden Between Groups
Table 5: Multivariable Linear Regression Results for Primary Endpoints
Table 6: Subgroup Analysis Results (if performed)

Listing 1: Complete Case Data Listing (anonymized ID)
Listing 2: Missing Data Pattern Summary

12. DEVIATIONS FROM SAP

Any deviation from this SAP will be documented in the final study report with justification. Major deviations (e.g., change in primary analysis population, addition of unplanned analyses) will be reported to the ethics committee and noted in manuscript limitations.