SCE 0906 Public Evidence Appendix

Statistical interpretation

Literature review date: September 13, 2026

Controlling principle. Statistics quantify evidence and uncertainty; they do not replace study design or clinical reasoning. Positive, negative, mixed, and inconclusive findings should be judged using the same standards, with effect size and uncertainty interpreted separately from clinical importance.

Point-by-point framework

Controlling Point	Primary role	Interpretive boundary
Sample size	Determines the amount of information available and contributes to precision and detectability.	Small studies may be imprecise; very large studies can detect clinically trivial differences.
Statistical power	Quantifies the planned probability of detecting a specified effect if it exists under design assumptions.	A nonsignificant underpowered trial may remain inconclusive; observed power adds little beyond the estimate and interval.
Confidence intervals	Express the estimated effect together with statistical precision and uncertainty.	Intervals should be compared with clinically relevant benefit, harm, and null thresholds rather than treated as a binary significance test.
Multiple comparisons	Addresses inflation of chance findings when many hypotheses, endpoints, time points, doses, or subgroups are tested.	Exploratory unadjusted findings can generate hypotheses but should not be promoted to confirmatory evidence.
Missing data	Addresses unavailable planned observations and the uncertainty or bias they may introduce.	Complete-case analysis can be biased; methods and sensitivity analyses cannot fully recreate unobserved information.
Subgroup analysis	Tests whether treatment effects differ across defined participant groups.	Credible claims require prespecification, rationale, adequate data, interaction testing, consistency, and preferably replication.
Intention-to-treat analysis	Preserves the randomized comparison by generally analyzing participants according to assigned group.	Does not eliminate missingness or protocol deviations; alternative estimands may answer different prespecified questions.

Evidence-handling rules

A p value should not be interpreted as the probability that a hypothesis is true, and statistical significance should not be equated with clinical importance. Effect estimates, confidence intervals, prespecified clinically meaningful thresholds, harms, durability, and patient priorities remain central.

A negative result should be examined to determine whether the data exclude a clinically important effect. If the confidence interval remains compatible with meaningful benefit and harm, the appropriate conclusion may be uncertainty rather than proof of no effect.

Principal references

1. 1. International Council for Harmonisation. E9 Statistical Principles for Clinical Trials.

2. 2. International Council for Harmonisation. E9(R1) Addendum on Estimands and Sensitivity Analysis in Clinical Trials.

3. 3. International Council for Harmonisation. E6(R3) Good Clinical Practice. 2025.

4. 4. U.S. Food and Drug Administration. Multiple Endpoints in Clinical Trials: Guidance for Industry. 2022.

5. 5. CONSORT Group. CONSORT 2025 Statement and explanatory materials for randomized trials.

6. 6. U.S. Food and Drug Administration. Innovative Designs for Clinical Trials of Cellular and Gene Therapy Products in Small Populations. Draft Guidance for Industry. September 2025.

7. 7. Contemporary clinical-trial methodology literature on confidence intervals, statistical power, subgroup analyses, multiplicity, and missing-data sensitivity analyses.

8. 8. Wasserstein RL, Lazar NA. The ASA Statement on p-Values: Context, Process, and Purpose. Am Stat. 2016;70:129-133.