

Randomized clinical trial data analysis

Shengping Yang PhD, Gilbert Berdine MD

I am analyzing data from my recently completed two-arm randomized clinical trial evaluating the effect of a 12-week dietary intervention on blood glucose, compared with placebo control in an adult population. I will adopt an intention-to-treat (ITT) principle in the analysis. I am wondering if there is anything else that I need to be aware of.

Randomized clinical trials (RCTs) are the gold standard for evaluating the effects of interventions on clinical outcomes. Data analysis is a critical component of a successful RCT. Several considerations are essential when planning and conducting the analysis, including adherence to the intention-to-treat (ITT) principle, definition of the estimand and analysis population, handling of missing data, and selection of an appropriate statistical model.

1. REGULATORY GUIDELINES

The ICH E9 guideline, Statistical Principles for Clinical Trials (1998), provides the backbone for statistical guidance in clinical trials.¹ It introduced the concept of the Full Analysis Set (FAS) as the practical realization of the ITT principle, defined the Per-Protocol Set (PPS), and emphasized the pre-specification of the primary analysis, analysis sets, and statistical model. Its addendum, ICH E9(R1), Estimands and Sensitivity Analysis in Clinical Trials, reframed the analysis process around a more fundamental question: before choosing an analysis, the trialist should define the estimand, i.e., the precise description of the treatment effect the trial is meant to answer, specified by five attributes: (i) population, (ii) treatment conditions, (iii) variable (endpoint), (iv) handling of intercurrent events, and (v) population-level summary.² In your study, the estimand might be described as the difference in mean

fasting glucose at 12 weeks between adults randomized to the diet and to control, regardless of intervention compliance.

In addition, Consolidated Standards of Reporting Trials (CONSORT) provides complementary guidance on the reporting of RCTs, including the participant flow diagram, which makes the disposition of randomized participants and the accounting for exclusions, withdrawals, and losses to follow-up transparent to readers.^{3,4}

2. ANALYSIS PRINCIPLES

2.1 THE INTENTION-TO-TREAT (ITT) PRINCIPLE

The ITT principle states that participants should be analyzed according to the treatment to which they were randomized, regardless of the treatment actually received, adherence, treatment discontinuation, or subsequent treatment changes. The rationale is structural: randomization creates treatment groups that are comparable, on average, with respect to both measured and unmeasured prognostic factors. Excluding or reassigning participants after randomization can compromise this protection and introduce selection bias because non-adherence, treatment discontinuation, and protocol deviations are often related to prognosis and outcomes.

The key feature of ITT is therefore preservation of randomized treatment assignment rather than exclusion based on post-randomization events. Ideally, all randomized participants would contribute to the primary analysis.

In practice, however, a participant who is randomized but never treated and has no post-randomization outcome assessment contributes no observed outcome data. This creates an important distinction between the ITT principle and the Full Analysis Set (FAS), which is the group of trial participants used for analysis that should conform as closely as possible

Corresponding author: Shengping Yang
Contact Information: Shengping.Yang@pbrc.edu
DOI: 10.12746/swjm.v14i61.1704

to the ITT principle. The FAS is generally intended to include all randomized participants who provide some post-randomization information, subject to prespecified criteria, and is considered a practical approximation to the ITT principle. In superiority trials, the FAS is generally preferred as the primary analysis population, with appropriate prespecified methods for handling missing outcome data.

The ITT principle can therefore be summarized by two related rules: (i) participants should be analyzed according to the treatment to which they were randomized; and (ii) every effort should be made to retain randomized participants in the analysis rather than excluding them due to post-randomization events. The first rule is fundamental and is preserved by retaining randomized treatment assignment; the second is an ideal that cannot always be achieved when outcome data are missing.

The common data analysis terms, such as FAS, multiple imputation (MI), linear mixed-effects models (LMM), and other missing-data methods, are not different forms of ITT. Rather, they are different tools for implementing an analysis that seeks to preserve the benefits of randomization while addressing incomplete outcome information under explicitly stated assumptions. Thus, the critical question for data analysis is no longer simply “Is this an ITT analysis?” but rather “What treatment effect is being estimated, in which population, how are intercurrent events handled, and what assumptions are required to make inference about unobserved outcomes?”

2.2 THE PER-PROTOCOL (PP) ANALYSIS

The PP analysis restricts the population to participants who sufficiently adhered to the protocol as designed. The purpose of a PP analysis is explanatory: What effect does the treatment achieve when it is administered and followed according to the protocol?

The principal weakness of PP analysis is that it sacrifices some of the protection provided by randomization. Adherence to treatment and compliance with the protocol are not random. Participants who discontinue treatment, deviate from the protocol, or fail to complete follow-up may differ systematically from

those who remain compliant. Consequently, comparing only protocol-compliant participants can introduce selection bias and may exaggerate or otherwise distort the treatment effect.

For superiority trials, the ITT/FAS-based analysis is generally the primary basis for inference, with PP analysis serving as a supportive analysis. Concordant results from ITT/FAS and PP analyses strengthen the evidence for a treatment effect, whereas discordant results require careful investigation, particularly with respect to adherence, treatment discontinuation, protocol deviations, and missing data.

2.3 THE COMPLETER ANALYSIS

A completer analysis includes all participants with an observed end-of-study (EOS) assessment regardless of adherence, including those who discontinued the intervention but remained under follow-up. It should be distinguished from the PP analysis, which further restricts the analysis to protocol-adherent participants; the two coincide only when all followed-up participants were adherent.

A completer analysis nominally retains the primary treatment-policy estimand but relies on stronger assumptions about the relationship between missingness and the unobserved outcome than a likelihood- or MI-based missing at random (MAR) analysis. Under ICH E9(R1), PP analysis generally addresses a different, adherence-related estimand, whereas a completer analysis is better viewed as a supportive or sensitivity analysis of the primary analysis.

A completer analysis can introduce substantial selection bias because completion is generally not a random event. Participants who discontinue early may differ systematically from completers with respect to treatment response, adverse events, disease severity, adherence, or other prognostic factors. Consequently, a completer analysis should generally not replace the primary ITT/FAS analysis.

3. STATISTICAL MODELS

We will focus primarily on two-arm randomized superiority trials with normally distributed continuous

outcomes measured at baseline and at the EOS. Several statistical approaches can be considered for comparing treatment groups. Trials with more complex designs can generally apply similar principles with design-specific considerations.

3.1 UNADJUSTED ENDPOINT COMPARISON

A two-sample *t*-test can be used to compare the mean EOS outcome, Y_1 , between treatment groups. Because treatment assignment is randomized, this provides an unbiased estimate of the treatment effect under the randomization-based framework. However, an unadjusted comparison does not use information from the baseline measurement, which may be strongly prognostic of the EOS outcome and therefore may result in less precise estimates than baseline-adjusted methods.

3.2 CHANGE-SCORE ANALYSIS

A change score is calculated as

$$\Delta Y = Y_1 - Y_0,$$

and the mean change is then compared between treatment groups, typically using a two-sample *t*-test or an equivalent regression model.

The approach is intuitively appealing because each participant serves as his/her own reference. However, its efficiency depends on the correlation between baseline and EOS measurements. If Y_0 and Y_1 have equal variance σ^2 and correlation ρ , then

$$\text{Var}(\Delta Y) = 2\sigma^2(1 - \rho).$$

Thus, under these assumptions, the change-score analysis is more precise than an unadjusted endpoint analysis when $\rho > 0.5$, whereas it is less precise when $\rho < 0.5$.

3.3 ANALYSIS OF COVARIANCE (ANCOVA)

ANCOVA compares the EOS outcome between treatment groups while adjusting for the baseline outcome. A typical model is

$$Y_1 = \beta_0 + \beta_1 \text{Treatment} + \beta_2 Y_0 + \varepsilon.$$

The treatment coefficient, β_1 , estimates the adjusted mean difference between treatment groups at EOS. Analysis of covariance is generally preferred to an unadjusted endpoint or change-score analysis when the baseline measurement is prognostic of the EOS outcome because it uses the optimal weight for baseline measurement.⁵

For a two-arm randomized trial with a continuous outcome measured at baseline and EOS, ANCOVA is often an efficient primary analysis when the primary estimand is the between-group difference in the EOS outcome under a treatment-policy strategy, where the occurrence of the intercurrent event is considered irrelevant to the definition of the treatment effect, with outcome data collected regardless of whether the event occurs.^{2,6}

3.4 LINEAR MIXED-EFFECTS MODEL

A linear mixed-effects model (LMM) treats the baseline and EOS measurements jointly as outcomes. For trials with baseline and EOS data collection only, a marginal model with an appropriate covariance structure can be used.

$$Y_{ij} = \beta_0 + \beta_1 \text{Treatment}_i + \beta_2 \text{Time}_{ij} + \beta_3 \text{Treatment}_i \times \text{Time}_{ij} + \varepsilon_{ij},$$

where $\varepsilon_i = (\varepsilon_{i0}, \varepsilon_{i1})^T \sim N(0, \Sigma_{UN})$.

An important special case is constrained longitudinal data analysis (cLDA), in which the baseline means are constrained to be equal across randomized treatment groups, reflecting the comparability created by randomization. The treatment-by-time interaction at EOS provides the estimated treatment effect.

$$Y_{ij} = \beta_0 + \beta_2 \text{Time}_{ij} + \beta_3 \text{Treatment}_i \times \text{Time}_{ij} + \varepsilon_{ij},$$

where $\varepsilon_i = (\varepsilon_{i0}, \varepsilon_{i1})^T \sim N(0, \Sigma_{UN})$. For a trial with only baseline and EOS measurements, cLDA and ANCOVA can produce the same treatment estimate under commonly used covariance assumptions; differences can arise when different variance/covariance structures are specified.⁷

The practical advantage of the longitudinal framework is that it provides a natural extension to designs

with repeated post-baseline measurements and can use likelihood-based methods to incorporate participants with incomplete outcome trajectories under a missing at random (MAR) assumption.

3.5 GENERALIZED ESTIMATING EQUATIONS

Generalized estimating equations (GEE) provide a population-averaged approach for analyzing correlated longitudinal outcomes. For continuous outcomes, a Gaussian GEE with an appropriate working correlation structure can be used to model baseline and EOS measurements while accounting for within-participant correlation.

$$E(Y_{ij}) = \beta_0 + \beta_1 \text{Treatment}_i + \beta_2 \text{Time}_{ij} + \beta_3 (\text{Treatment}_i \times \text{Time}_{ij}),$$

Generalized estimating equations can provide consistent estimates of population-averaged treatment effects when the mean model is correctly specified and the estimating procedure and missing-data assumptions are appropriate. With only two measurement occasions and a normally distributed continuous outcome, GEE generally offers little advantage over ANCOVA or an appropriately specified linear mixed-effects model.

Standard complete-case GEE generally relies on a missing-completely-at-random assumption, whereas inverse-probability weighting, multiple imputation, or other extensions can be used to address missing data under MAR assumptions.⁸

For a simple two-arm trial with baseline and EOS measurements, ANCOVA or a constrained longitudinal mixed-effects model will generally be more straightforward and efficient choices than GEE.

4. MISSING DATA

In the baseline-EOS design, missing data will primarily involve the EOS outcome.

4.1 MISSINGNESS MECHANISMS

Missing data are conventionally classified as missing completely at random (MCAR), missing at random

(MAR) given observed data, or missing not at random (MNAR). The primary analysis will generally assume MAR, under which the probability that EOS is missing may depend on observed information such as treatment assignment, baseline outcome, and other prespecified prognostic or auxiliary variables, but not on the unobserved EOS outcome after conditioning on those observed variables.

The MAR assumption cannot be verified from the observed data alone. Therefore, the missing-data assumptions should be prespecified and examined through sensitivity analyses.

Importantly, ITT does not by itself specify how missing outcomes should be handled. A method such as multiple imputation or likelihood-based estimation is needed to obtain inference when outcome data are incomplete.

4.2 PRIMARY ANALYSIS

For superiority trials with baseline and EOS data collection, the ITT principle should be followed, with missing EOS outcomes handled under a prespecified MAR assumption.

There is no single universally preferred statistical implementation of this principle. The choice should be prespecified and should be consistent with the estimand, analysis model, and available auxiliary information.

OPTION 1: MULTIPLE-IMPUTATION ANCOVA

Missing EOS values, as well as missing baseline or covariate values when applicable, may be multiply imputed under a model including randomized treatment, baseline outcome, and prespecified auxiliary variables predictive of the EOS outcome or of discontinuation. The imputation model should be compatible with, or at least adequately reflect, the analysis model, including treatment effects and important treatment-by-covariate interactions when appropriate. Results are combined using Rubin's rules.⁹

The choice of imputation method should be based on the outcome distribution, missing-data mechanism, and

analysis model. Recent simulation work suggests that some machine-learning-based imputation approaches can produce poorly calibrated standard errors or confidence intervals when used naively with Rubin's rules in clinical-trial settings.¹⁰ Therefore, the validity of the complete imputation-and-analysis procedure should be evaluated rather than assuming that a particular algorithm is universally appropriate or inappropriate.

OPTION 2: LIKELIHOOD-BASED cLDA

A likelihood-based cLDA model can be fit to all available outcome data by maximum likelihood under MAR, without explicitly imputing missing outcomes. This approach is particularly attractive when the longitudinal model is well specified, and there is no compelling need to incorporate additional auxiliary variables through an imputation model.

When baseline and EOS data are complete, likelihood-based cLDA and ANCOVA can yield equivalent treatment estimates under corresponding model assumptions. With missing EOS data, however, likelihood-based cLDA uses the observed data directly and does not require stochastic imputation.

Selection rule: if prespecified auxiliary variables predictive of the EOS outcome or of discontinuation are available, Multiple-Imputation ANCOVA is preferred, because it can exploit them and thereby make the MAR assumption more plausible. If no such variables exist, both options target the same estimand under the same MAR assumption and yield very similar estimates; Likelihood-Based cLDA is then attractive for its simplicity – valid likelihood-based inference without imputation models or Rubin's rules.

4.3 SENSITIVITY AND SUPPLEMENTARY ANALYSES

Per-Protocol (PP) analysis: Assesses whether the treatment effect is consistent among participants who followed the protocol.

Completer / complete-case ANCOVA: Assesses the robustness of the primary conclusion.

Delta-adjustment or tipping-point analysis: Can be used to examine progressively less favorable

assumptions about the unobserved EOS outcomes. For example, the imputed EOS values for participants with missing outcomes can be shifted by prespecified amounts, and the analysis repeated to determine how large a departure from MAR would be required to change the study conclusion.¹¹

Agreement between the primary FAS/ITT analysis and PP or completer analyses can strengthen confidence in the findings. Discordant results should be investigated in relation to treatment adherence, protocol deviations, reasons for dropout, and the pattern of missing EOS outcomes.

Table 1 is a summary of scenarios for missing data handling across different statistical models.

5. REPORTING AND INTERPRETATION

CONSORT requires reporting the numbers of participants randomized, assigned to each intervention, receiving the intervention, completing follow-up, and included in the analysis, as applicable, together with effect estimates and confidence intervals and analyses that are consistent with the prespecified statistical plan.^{3,4} An example of data reporting is provided in Table 2. Interpretation should state the estimand in words: the population, treatment condition, outcome variable, handling of intercurrent events, and population-level summary being reported.

A superiority claim rests primarily on the prespecified primary analysis. PP and completer analyses should generally be presented as supplementary or sensitivity analyses rather than as replacements for the primary analysis.

6. DISCUSSION

Even the simplest RCT – one baseline measurement, one endpoint, and two arms – requires weighing both clinical and statistical considerations. In the ideal situation of no missing values, ANCOVA and LMM/cLDA can yield equivalent treatment estimates under corresponding model assumptions. A PP analysis, in contrast, restricts the population according to post-randomization adherence and therefore addresses a different, explanatory question.

Table 1. Different Statistical Models

Method	Uses Baseline Info	Missing Data Handling	Key Condition/ Assumption	Typical Role
Unadjusted <i>t</i> -test	None	Complete case (EOS observed)	Randomization; <i>t</i> -test assumptions	Rarely used
Change score	Fixed weight	Complete case (requires both measurements)	Beats unadjusted endpoint only if $\rho > 0.5^a$; never more efficient than ANCOVA when baseline is prognostic	Intuitive; rarely optimal
ANCOVA	Estimated weight	Complete case; usually paired with MI under MAR	Baseline prognostic of EOS; linear model correctly specified	Often preferred primary analysis
cLDA/ LMM	Baseline as outcome	Likelihood-based under MAR (partial data retained)	Correct mean and covariance specification ^b	Equivalent to MI-ANCOVA when no auxiliaries
GEE	Baseline as outcome or covariate	Complete case requires MCAR ^c ; weighted GEE or MI under MAR	Correct mean model for consistency; working correlation affects efficiency only (robust variance)	Population-averaged; little advantage compared to cLDA

^a Under equal baseline and EOS variances; ρ = baseline-EOS correlation.

^b With baseline + EOS data, cLDA(UN) and ANCOVA give identical treatment estimates with complete data. With missing EOS and no auxiliary variables, cLDA(UN) and MI-ANCOVA are asymptotically equivalent under compatible model assumptions.

^c Stronger than the MAR assumption required for likelihood or MI approaches.

Missing values, however, are often unavoidable, and their occurrence introduces additional assumptions because the outcomes that would have been observed under different missing-data scenarios cannot be directly determined. For example, if a patient dies before completing the treatment protocol, was the death caused by severe disease or by treatment toxicity?

Two clarifications are important. First, ITT is not itself a statistical analysis or a statistical model; it is a principle governing how randomized participants are retained in the treatment comparison. Thus, the term “ITT analysis” is often used to describe an analysis

that preserves treatment assignment at randomization and targets a treatment-policy estimand.

Second, ITT does not necessarily guarantee a conservative estimate of treatment effect. Although non-adherence or treatment crossover can attenuate the contrast between randomized groups and thereby reduce power in many superiority trials, the direction and magnitude of bias depend on the nature of the intercurrent events, missing data, and treatment effects. Therefore, it is more precise to say that ITT preserves the benefits of randomization and targets the effect of treatment assignment, with valid

Table 2. Intervention Effect on the Primary Outcome

Outcome	Treatment (n)	Baseline (Mean (SD))	EOS (Mean (SD))	Change from Baseline (Δ , 95% CI)	Difference in EOS Mean* (Beta, 95% CI)	p-value*
Primary outcome	Control (n = xx)	xx (xx)	xx (xx)	xx (xx)		
	Intervention (n = xx)	xx (xx)	xx (xx)	xx (xx)	xx (xx)	xx

* Adjusted for prespecified baseline covariates, including xx.

inference requiring appropriate assumptions about missing outcomes and the specified estimand, rather than simply labeling ITT as “conservative.”

For data analysis, the estimand should be defined first, followed by the analysis population, treatment strategy, handling of intercurrent events, statistical model, missing-data assumptions, and sensitivity analyses. This sequence helps ensure that the reported treatment effect corresponds to the clinical question the trial was designed to answer.

When baseline and covariates are complete, ANCOVA and likelihood-based cLDA can yield essentially identical treatment estimates under corresponding model assumptions. When outcome or covariate data are missing, the choice between likelihood-based analysis and multiple imputation should depend on the analysis model, the missing-data pattern, and the availability of informative auxiliary variables.

Sensitivity analysis is important because MAR cannot be verified from the observed data alone.

Article citation: Yang S, Berdine G. Randomized clinical trial data analysis. *The Southwest Journal of Medicine*. 2026;14(61):118–124

From: Penington Biomedical Research Center, Louisiana State University, Baton Rouge, LA (SY)
Department of Internal Medicine, Texas Tech University Health Sciences Center, Lubbock, TX (GB)

Conflicts of interest: none

This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

REFERENCES

1. European Medicines Agency. ICH E9 statistical principles for clinical trials: Step 5. European Medicines Agency; September 1998. Accessed October 12, 2026. https://www.ema.europa.eu/en/documents/scientific-guideline/ich-e-9-statistical-principles-clinical-trials-step-5_en.pdf
2. European Medicines Agency. ICH E9(R1) addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials: Step 5. European Medicines Agency; February 17, 2020. Accessed October 12, 2026. https://www.ema.europa.eu/en/documents/scientific-guideline/ich-e9-r1-addendum-estimands-and-sensitivity-analysis-clinical-trials-guideline-statistical-principles-clinical-trials-step-5_en.pdf
3. Schulz KF, Altman DG, Moher D; CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. *Ann Intern Med*. 2010; 152(11):726–32. doi: 10.7326/0003-4819-152-11-201006010-00232. Epub 2010 Mar 24. PMID: 20335313.
4. Hopewell S, Chan A, Collins G S, et al. CONSORT 2025 statement: updated guideline for reporting randomised trials *BMJ*. 2025;389:e081123 doi: 10.1136/bmj-2024-081123
5. Vickers, AJ, Altman, DG. Analysing controlled trials with baseline and follow up measurements. *BMJ*. 2001; 323(7321):1123–4.
6. O’Connell NS, Dai L, Jiang Y, et al. Methods for Analysis of Pre-Post Data in Clinical Research: A Comparison of Five Common Methods. *J Biom Biostat*. 2017;8(1):1–8. doi: 10.4172/2155-6180.1000334. PMID: 30555734; PMCID: PMC6290914.
7. Liu GF, Lu K, Mogg R, et al. Should baseline be a covariate or dependent variable in analyses of change from baseline in clinical trials? *Stat Med*. 2009;28(20):2509–30. doi: 10.1002/sim.3639. PMID: 19610129.
8. Fitzmaurice GM, Laird NM, Ware JH. *Applied Longitudinal Analysis*. 2nd ed. Hoboken, NJ: John Wiley & Sons, 2011.
9. Rubin DB. *Multiple Imputation for Nonresponse in Surveys*. New York: John Wiley & Sons, 1987.
10. Tackney MS, Bartlett JW, Williamson E, et al. Comparison of parametric versus machine-learning multiple imputation in clinical trials with missing continuous outcomes. *BMC Med Res Methodol* (2026). doi: 10.1186/s12874-026-02962-6
11. Cro S, Morris TP, Kenward MG, et al. Sensitivity analysis for clinical trials with missing continuous outcome data using controlled multiple imputation: A practical guide. *Statistics in Medicine*. 2020;39:2815–42. doi: 10.1002/sim.8569