Paper AS02

Framework for Incorporating Estimands within the Statistical Analysis Plan

Inderjeet Arora, GenInvo, Toronto, Canada

ABSTRACT

In clinical trials, estimands establish a structured framework to define the precise treatment effect of interest, ensuring that statistical analyses are aligned with clinical objectives. Within a Statistical Analysis Plan (SAP), the biostatistician's role is to specify each estimand clearly, consistent with ICH E9(R1) guidance. This involves detailing its key components: the target population (e.g., all randomized subjects), the variable or endpoint of interest (e.g., primary efficacy measure), strategies for handling intercurrent events (e.g., treatment policy, hypothetical, or composite), and the chosen summary measure (e.g., mean difference, hazard ratio). The SAP should also outline approaches to missing data and include sensitivity analyses to evaluate robustness. Collaboration with clinical teams is essential to ensure that estimands address clinically meaningful questions. By explicitly linking estimands to analysis methods, models, and assumptions, the SAP improves transparency, regulatory compliance, and interpretability, ultimately supporting reliable decision-making in drug development.

INTRODUCTION

The publication of the ICH E9(R1) addendum marked a fundamental shift in how clinical trial objectives are formulated and how treatment effects are interpreted. Historically, clinical trials often relied on loosely defined objectives, with ambiguity regarding how post-randomization events, missing data, and protocol deviations were handled in the analysis. This disconnect sometimes resulted in analyses that were difficult to interpret and challenging for regulators to assess.

The estimand framework was introduced to resolve these ambiguities by requiring explicit definition of the treatment effect of interest, independent of the statistical methods used to estimate it. By separating the what from the how, estimands promote clarity, transparency, and alignment between clinical objectives and statistical analyses.

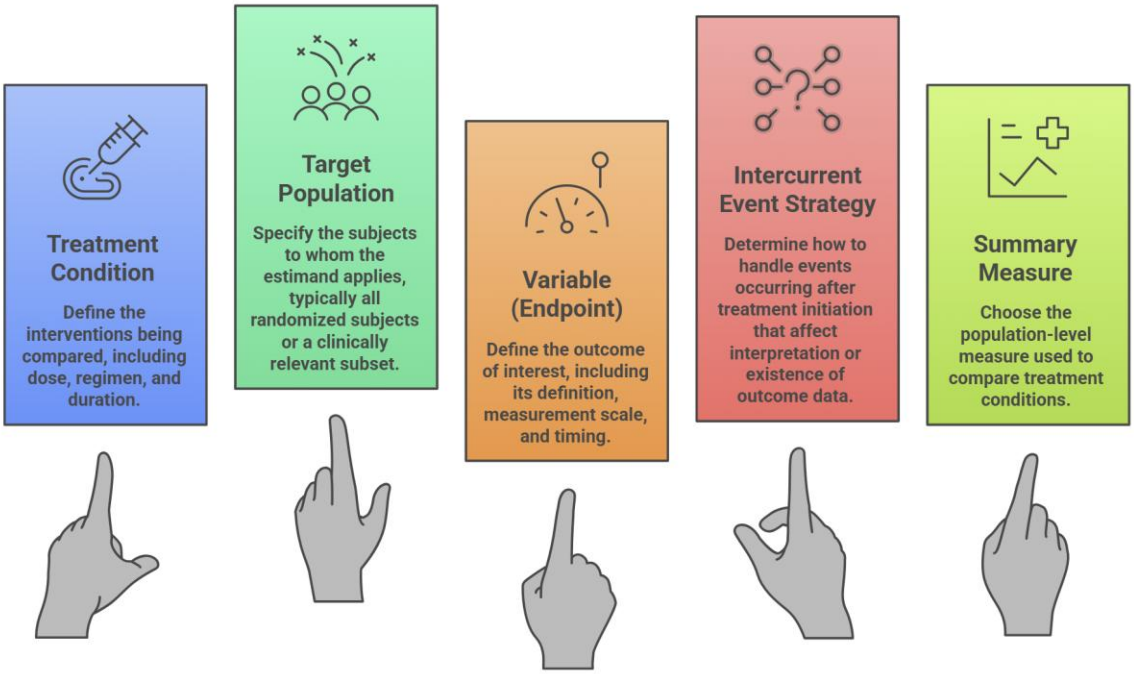
The Statistical Analysis Plan (SAP) is the key document in which estimands are operationalized. This paper proposes a practical framework for incorporating estimands within the SAP, with emphasis on regulatory expectations, cross-functional collaboration, and implementation considerations, including statistical programming and dataset design.

OVERVIEW OF ICH E9(R1) ESTIMAND FRAMEWORK

ICH E9(R1) defines an estimand through five key attributes: treatment condition, population, variable (endpoint), intercurrent event handling strategy, and summary measure. Together, these attributes precisely describe the treatment effect of interest. The estimand framework separates the definition of the estimand from the statistical method used to estimate it, allowing greater flexibility and clarity.

Within the SAP, it is essential to demonstrate how each estimand aligns with the study objectives and endpoints defined in the protocol. Clear traceability between protocol objectives, estimands, and analyses supports regulatory review and minimizes post hoc interpretation.

FIGURE 1: ESTIMAND FRAMEWORK ATTRIBUTES



ROLE OF ESTIMANDS WITHIN THE STATISTICAL ANALYSIS PLAN

The SAP translates the high-level estimand definitions into actionable statistical procedures. This includes specifying analysis populations, defining analysis variables, selecting statistical models, and describing assumptions related to missing data and intercurrent events.

Incorporating estimands into the SAP typically involves the inclusion of dedicated estimand sections or tables that summarize the attributes of each estimand. These sections should be written in a manner that is accessible to both statistical and non-statistical reviewers, while still providing sufficient technical detail to support implementation.

Importantly, the SAP should avoid treating estimands as purely descriptive statements. Instead, each estimand should be explicitly linked to the corresponding primary and secondary analyses, sensitivity analyses, and estimators. This linkage is critical for ensuring traceability from trial objectives to reported results.

SPECIFICATION OF ESTIMAND COMPONENTS IN THE SAP

The SAP should clearly specify each component of the estimand. The target population typically includes all randomized subjects, consistent with the intention-to-treat principle. The variable defines the endpoint of interest, such as a change from baseline or time-to-event outcome.

Intercurrent event strategies must be explicitly stated, including treatment policy, hypothetical, composite, while-on-treatment, or principal stratum approaches. The summary measure, such as a mean difference or hazard ratio, should be justified based on clinical relevance and regulatory expectations.

Following are the two examples of defining the estimand section in SAP:

1) TEXTUAL APPROACH

Treatment:

- Investigational product (IP) or comparator
- Example:
  - Experimental:
    - Therapy for anaemia
  - Control:
    - Placebo

Population:

- Participants for whom IP is being studied

- Example:
    - Adults (18 years or older) with chemotherapy induced anaemia

Variable:

- The specific clinical variable or outcome being measured in the trial
- Example:
    - Change in haemoglobin concentration

Intercurrent Events (ICEs):

- ICEs are the events that can impact the interpretation of the treatment effect
- Example:
    - Treatment discontinuation due to reason other than lack of efficacy or worsening of the adverse events. Data from all visits following the ICE will be marked as missing [Hypothetical Strategy]

Population-level summary:

- In what way the treatment effect is summarized
- Example:
    - Difference in the haemoglobin concentration between Therapy for anaemia and placebo.

2) TABULAR APPROACH: ABOVE CAN BE UTILISED AS TABULAR FORM

TABLE 1: ESTIMAND

<table>
<tr><td colspan="2">Treatment: Investigational product (IP) or comparator.

Example:
Experimental:
• Therapy for anaemia
Control:
• Placebo</td></tr>
<tr><td colspan="2">Population: Participants for whom IP is being studied.

Example:
• Adults (18 years or older) with chemotherapy induced anaemia</td></tr>
<tr><td colspan="2">Variable: The specific clinical variable or outcome being measured in the trial.

Example:
• Change in haemoglobin concentration</td></tr>
<tr><td>Intercurrent Events (ICEs): ICEs are the events that can impact the interpretation of the treatment effect.

Example:
1. Treatment discontinuation due to reason other than lack of efficacy or worsening of the adverse events.</td><td>Handling ICEs: Strategies to handle ICE

Example:
1. Data from all visits following the ICE will be marked as missing [Hypothetical Strategy]</td></tr>
<tr><td colspan="2">Population-level Summary: In what way the treatment effect is summarized

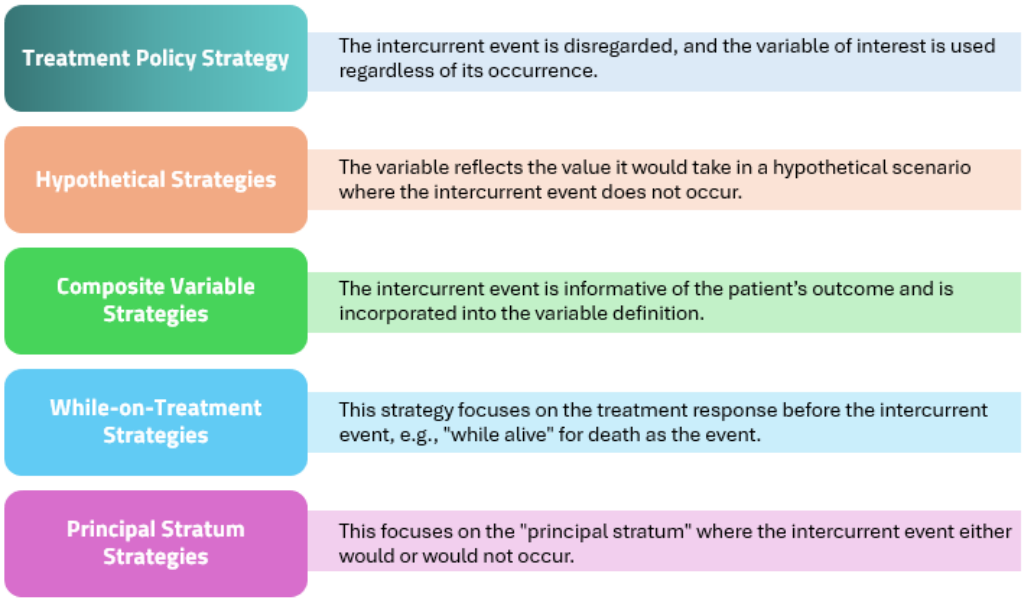
Example:
• Difference in the haemoglobin concentration between therapy for anaemia and placebo.</td></tr>
</table>

HANDLING INTERCURRENT EVENTS

Intercurrent events occur after treatment initiation and can affect interpretation of the endpoint. Examples include treatment discontinuation, use of rescue medication, or death. The estimand framework requires explicit strategies to handle such events, which must be clearly documented in the SAP.

Different strategies may be appropriate for different estimands. For example, a treatment policy strategy may reflect real-world effectiveness, while a hypothetical strategy may address efficacy under ideal conditions. The SAP should justify the chosen approach and describe its implementation in the analysis.

Figure 2: Strategies for Addressing ICEs



Following example illustrates how ICE can be handled using different strategies:

- Treatment Policy Strategy: Utilize data as observed
- Hypothetical Strategies: Set to missing from ICE point onwards
- Composite Variable Strategies: Set data to not achieving the endpoint or not having the expected observation value (actual or change)
- While-on-Treatment Strategies: Utilize data before ICE
- Principal Stratum Strategies: Utilizing data from a group of participants who will potentially experience the ICE (or group of participants who will not experience ICE)

MISSING DATA CONSIDERATIONS

Missing data handling is a property of the estimator, not the estimand itself; however, the estimand strategy dictates how missingness arises and should be treated in the analysis. This is clear in the conceptual distinction that missing data may occur as a consequence of intercurrent events (e.g., withdrawal after treatment discontinuation), but missing data are not themselves intercurrent events.

For example:

- Under a treatment policy strategy, post-intercurrent event outcomes are *relevant* and thus considered observed data; absent values in these post-event intervals are truly missing and must be addressed in the estimator.

- In contrast, under a while-on-treatment strategy, outcomes after treatment discontinuation may not be relevant to the estimand and therefore are not considered missing relative to that strategy.

Because estimand strategies alter the definition of when missingness occurs, the SAP must articulate the estimand-specific missing data definition and corresponding analytic approach.
The SAP should explicitly state the assumed mechanism of missingness aligned with the estimand.
Classical statistical theory categorizes missing data mechanisms as:

- Missing Completely At Random (MCAR): Missingness is unrelated to observed or unobserved data.
- Missing At Random (MAR): Missingness can be explained by observed data variables.
- Missing Not At Random (MNAR): Missingness depends on unobserved data.

Selecting an appropriate mechanism is crucial because it underpins the choice of analysis method (e.g., likelihood-based models or multiple imputation). In many clinical contexts, particularly when missingness is related to treatment response or adverse events, the MAR assumption may be too strong, necessitating sensitivity analyses under MNAR assumptions.
Once the estimand and missing data mechanism are defined, the SAP must prespecify a primary analysis method that yields unbiased and efficient estimates under the assumed missing data mechanism.
Commonly used methods include:

- **Likelihood-based models** such as mixed models for repeated measures (MMRM), which implicitly handle MAR under certain conditions;
- **Multiple imputation (MI)**, which replaces missing values multiple times to reflect uncertainty about the missing data;
- **Model-based estimators** that incorporate assumptions about post-intercurrent event trajectories consistent with the estimand.

Importantly, for treatment policy estimands, the SAP should justify the choice of method and its compatibility with the estimand's definition of missing data. For example, point estimators that implicitly impute outcome data may not be suitable when missing data mechanisms are non-ignorable or when the estimand strategy does not support certain modelling assumptions.
ICH E9(R1) emphasizes that sensitivity analyses are not optional but essential to assess the robustness of the primary analysis to alternative plausible assumptions about missing data. Sensitivity analyses should:

- Use alternative missing data assumptions (e.g., MNAR, control-based imputation, tipping point analysis);
- Explore different estimators, such as reference-based MI that assume outcomes after dropout follow a different distribution;
- Evaluate impact on key estimand conclusions, not just statistical significance.

For example, reference-based MI approaches can be used to assess robustness by assuming that, after dropout, patients behave like those in a comparator group (e.g., following control arm trajectories). These alternative assumptions produce sensitivity results that illuminate how conclusions might vary under different plausible scenarios.
Sensitivity analyses should be pre-specified in the SAP with clear linkage to the estimand, chosen missing data mechanisms, and interpretation criteria. This includes documenting the rationale for each analysis, the assumptions being tested, and the decision rules for interpreting consistency or divergence from the primary analysis.
From an implementation perspective, the SAP should include:

- **Detailed imputation model specifications**, including variables, correlations, and predictors;
- **Software, macros, or scripts** used for primary and sensitivity analyses;
- **Tables and listings** that tabulate patterns of missing data, reasons for missingness, and the distribution of prognostic variables in missing vs observed data.

These details support reproducibility, audit readiness, and regulatory inspection. Explicit documentation of missing data assumptions and sensitivity analyses fosters transparency in regulatory submissions and supports confidence in estimated treatment effects.

EXAMPLE: DOCUMENTATION OF MISSING DATA PATTERNS AND REASONS

SAP SPECIFICATION EXAMPLE
Missing data patterns will be summarized by treatment group using descriptive statistics and graphical displays. The number and percentage of subjects with missing primary endpoint data at each scheduled visit will be tabulated. Reasons for missingness, including treatment discontinuation, adverse events, withdrawal of consent, and loss to follow-up, will be summarized.

EXAMPLE: PRIMARY ANALYSIS IMPLEMENTATION (MAR ASSUMPTION)

SAP SPECIFICATION EXAMPLE
The primary analysis will be performed using a mixed model for repeated measures (MMRM) under a Missing At Random (MAR) assumption. The model will include treatment, visit, treatment-by-visit interaction, baseline value, and stratification factors as fixed effects.

PROGRAMMING AND ADAM CONSIDERATIONS

- The analysis dataset includes:
    - One record per subject per visit
    - Observed post-baseline values only (no implicit imputation)
- Missing values are handled implicitly through likelihood-based estimation.
- ADICE is not directly used in the primary analysis but provides traceability to justify MAR plausibility by documenting post-randomization events.

EXAMPLE: REFERENCE-BASED MULTIPLE IMPUTATION SENSITIVITY ANALYSIS

SAP SPECIFICATION EXAMPLE
A sensitivity analysis will be conducted using reference-based multiple imputation assuming a hypothetical estimand in which subjects discontinuing treatment would have followed the response trajectory of the control group.

IMPLEMENTATION DETAILS

- Missing values following treatment discontinuation (identified via ADICE) are imputed using a **Jump-to-Reference (J2R)** approach.
- Imputation model includes:
    - Baseline endpoint value
    - Treatment group
    - Visit
    - Key prognostic covariates
- 50 imputed datasets are generated and combined using Rubin’s rules.

PROGRAMMING CONSIDERATIONS

- ADICE is used to:
    - Identify the imputation start point (date of intercurrent event)
    - Apply imputation only after the relevant intercurrent event

- Parameter-driven macros allow reuse across endpoints.

EXAMPLE: TIPPING POINT ANALYSIS FOR MNAR ASSESSMENT

SAP SPECIFICATION EXAMPLE
A tipping point analysis will be performed to assess the robustness of conclusions under Missing Not At Random (MNAR) assumptions by systematically varying post-dropout imputed values.

IMPLEMENTATION DETAILS
- Incremental shifts (deltas) are applied to imputed values for missing data in the treatment group.
- The analysis identifies the point at which statistical conclusions change.

PROGRAMMING AND REPORTING
- Deltas are parameterized and looped across a predefined range.
- Results are summarized in:
    - A table of treatment effects by delta value
    - A graphical display highlighting the tipping point

This analysis provides regulators with a clear understanding of how extreme assumptions must be to overturn conclusions.

SENSITIVITY AND SUPPLEMENTARY ANALYSES
Sensitivity analyses are a fundamental component of the estimand framework and play a critical role in assessing the robustness of conclusions drawn from the primary estimand. In accordance with ICH E9(R1), sensitivity analyses are designed to evaluate the extent to which the results of the primary analysis depend on the assumptions made regarding intercurrent events, missing data, and model specification. As such, sensitivity analyses should be pre-specified in the SAP and directly linked to the primary estimand.
The SAP should clearly describe the objectives of each sensitivity analysis, the alternative assumptions being tested, and how the results will be interpreted relative to the primary analysis. Sensitivity analyses should not redefine the clinical question but rather assess whether reasonable deviations from the primary assumptions materially affect the estimated treatment effect.

EXAMPLE: SENSITIVITY ANALYSES FOR INTERCURRENT EVENTS AND MISSING DATA
For a primary estimand defined using a treatment policy strategy, outcomes following treatment discontinuation are considered relevant. The primary analysis may assume a Missing At Random (MAR) mechanism and be implemented using a likelihood-based method such as a mixed model for repeated measures (MMRM). To assess robustness, the SAP may pre-specify sensitivity analyses that relax the MAR assumption, including:
- A reference-based multiple imputation approach (e.g., jump-to-reference), in which post-discontinuation outcomes are imputed under the assumption that subjects follow the response trajectory of the control group.
- A delta-adjusted multiple imputation analysis, applying systematic shifts to imputed values to explore Missing Not At Random (MNAR) scenarios.
- A tipping point analysis to identify the magnitude of departure from MAR assumptions required to alter the study conclusion.

These sensitivity analyses evaluate whether the treatment effect remains consistent in direction and clinically meaningful magnitude under alternative, yet plausible, assumptions.

SUPPLEMENTARY ESTIMANDS
In addition to sensitivity analyses, supplementary estimands may be defined to address additional clinical or scientific questions beyond the primary objective. Supplementary estimands represent distinct treatment effects rather than alternative assumptions for the same estimand and therefore require separate definitions of intercurrent event handling, endpoints, and summary measures.

Supplementary estimands are particularly useful for providing complementary perspectives on treatment benefit, such as understanding efficacy while on treatment or under hypothetical scenarios that may be of clinical interest.

EXAMPLE: SUPPLEMENTARY ESTIMAND DEFINITION

While the primary estimand may use a treatment policy strategy to reflect real-world effectiveness, a supplementary estimand may be defined using a while-on-treatment strategy to assess the pharmacological effect of the investigational product prior to treatment discontinuation.

For example:

- **Primary estimand**: Mean difference in change from baseline at Week 24 regardless of treatment discontinuation (treatment policy strategy).

- **Supplementary estimand**: Mean difference in change from baseline at Week 24 considering data only up to treatment discontinuation (while-on-treatment strategy).

Each estimand addresses a different clinical question and is analyzed using methods appropriate to its definition.

CROSS-FUNCTIONAL COLLABORATION

Successful implementation of the estimand framework depends on early and sustained cross-functional collaboration among statisticians, clinicians, regulatory experts, data managers, and statistical programmers. Because estimands define the clinical question of interest rather than merely the analysis technique, their development must be driven by a shared understanding of both clinical relevance and operational realities.

Early cross-functional discussions, ideally initiated during protocol development, are critical to ensure that estimands reflect clinically meaningful treatment effects while remaining feasible to implement and interpret. Clinicians provide insight into which intercurrent events are clinically relevant and how treatment effects should be interpreted in the presence of events such as treatment discontinuation or rescue medication. Statisticians translate these clinical objectives into well-defined estimands and corresponding analysis strategies consistent with ICH E9(R1). Regulatory experts contribute perspective on evolving regulatory expectations and ensure alignment with precedents and guidance. Data managers and programmers assess whether the necessary data can be collected, structured, and analyzed in a manner that supports the proposed estimands.

The Statistical Analysis Plan serves as a central communication and alignment document across functions. By explicitly documenting the agreed-upon estimands, intercurrent event strategies, missing data assumptions, and corresponding analysis methods, the SAP provides a single source of truth for study conduct and analysis. This transparency reduces the risk of misinterpretation, late-stage revisions, or implementation inconsistencies and supports efficient regulatory review.

Ongoing collaboration throughout study conduct and analysis remains essential, particularly when unanticipated intercurrent events or data issues arise. A well-documented SAP, grounded in cross-functional agreement, ensures that any adaptations remain consistent with the original estimand framework and preserve the integrity and interpretability of the trial results.

CONCLUSION

The ICH E9(R1) estimand framework represents a significant evolution in clinical trial analysis by requiring explicit definition of the treatment effect of interest and separating this definition from the statistical methods used to estimate it. By clarifying how intercurrent events, missing data, and post-randomization outcomes are handled, estimands address long-standing sources of ambiguity that have historically complicated interpretation and regulatory review. The Statistical Analysis Plan (SAP) is the critical document in which estimands are operationalized, providing traceability from protocol objectives to analysis methods and reported results.

Effective incorporation of estimands within the SAP requires clear specification of estimand components, prespecified strategies for intercurrent events and missing data, and robust sensitivity analyses to evaluate the impact of alternative assumptions. Practical implementation relies on close cross-functional collaboration and fit-for-purpose data standards, including ADaM and emerging datasets such as ADICE, to ensure transparent and reproducible analyses. When implemented consistently, the estimand-driven SAP enhances interpretability, supports regulatory decision-making, and strengthens confidence in clinical trial conclusions.

REFERENCES

ICH E9(R1). Addendum on Estimands and Sensitivity Analysis in Clinical Trials. International Council for Harmonisation, 2019.

ICH E9. Statistical Principles for Clinical Trials. International Council for Harmonisation, 1998.

Mallinckrodt CH, et al. Estimands, Estimators, and Sensitivity Analysis in Clinical Trials. Statistics in Biopharmaceutical Research, 2012.

Rutherford MJ, et al. Estimands and the Analysis of Clinical Trials. Pharmaceutical Statistics, 2020.

FDA. Adjusting for Intercurrent Events in Clinical Trials. U.S. Food and Drug Administration, 2021.

Suzie Cro, et al. A four-step strategy for handling missing outcome data in randomised trials affected by a pandemic, 2020

A. Atkinson, et al. Application of the estimand framework for an emulated trial using reference based multiple imputation to investigate informative censoring, 2024

Helle Lynggaard, et al. Principles and recommendations for incorporating estimands into clinical study protocol templates, 2022

Brennan C Kahan, et al. The estimands framework: a primer on the ICH E9(R1) addendum, 2024

Yun Wang, et al. Statistical methods for handling missing data to align with treatment policy strategy, 2023

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Inderjeet Arora

GenInvo

1408 East Empire Street 61701

Bloomington (Illinois), United States

+1 309-807-5006

inderjeet.a222@geninvo.com