

Advancing Clinical Trial Analysis with Trimmed Means: Estimands, Intercurrent Events, and Regulatory Alignment

Prasanth Palanisamy
SENSAN Biosciences

- Introduction
- Intercurrent Events and the ICH E9(R1) Estimand Framework
- Application of Trimmed Means in Clinical Trials for Intercurrent Events
- Statistical Methodology of Trimmed Means
- Case Study
- Regulatory Landscape and FDA Approval
- SAS Implementation Example
- Conclusions

Definition and Conceptual Basis of Trimmed Means:

- A trimmed mean removes a prearranged percentage of the lowest and highest values before calculating the average.
- It is used to reduce the influence of outliers and provide a more stable estimate of central tendency.
- The approach emphasizes estimating the central tendency of the main body of the distribution, not the extremes.

Advantages in Handling Outliers and Skewed Data:

- Trimmed means are robust estimators, less sensitive to non-normality and outliers compared to traditional means.
- In skewed or asymmetric distributions, trimmed means better capture the true central value.
- They lead to lower standard errors and potentially increased statistical power in some settings.
- Clinically, they allow for a clearer treatment effect interpretation, especially when extreme outcomes (like dropout or toxicity) may mask the effect on typical patients.
- There is a trade-off between robustness and efficiency, more trimming improves robustness but may reduce precision.

Comparison with Mean and Median:

- The mean includes all data points and is **highly sensitive to extreme values**.
- The median is more robust, using only the central value(s), but may ignore too much information.
- The trimmed mean balances the two, it retains more data than the median while still minimizing outlier influence.
- The choice between these methods should be based on the clinical question and estimand definition, as emphasized by ICH E9(R1).

Intercurrent Events and the ICH E9(R1) Estimand Framework

Understanding Intercurrent Events (ICEs) in Clinical Trials

- Intercurrent events (ICEs) are post-treatment initiation events that affect the interpretation or availability of outcome data, such as use of **rescue medication, treatment discontinuation, or death**.
- ICEs are classified as **treatment-modifying or truncating events**, and accurate classification is critical to choosing appropriate estimand strategies and avoiding biased conclusions.

Overview of the ICH E9(R1) Estimand Framework and its Attributes

- ICH E9(R1) expands the PICO (Population, Intervention, Comparator and Outcome) framework by adding two key attributes: summary measure and strategy for handling ICEs, aligning statistical analysis with the clinical question.
- The framework outlines **five strategies, treatment policy, hypothetical, composite, while on treatment, and principal stratum**, each targeting a distinct clinical interpretation of treatment effect in the presence of ICEs.

Intercurrent Events and the ICH E9(R1) Estimand Framework (Cont...)

Intercurrent Events and the
ICH E9(R1) Estimand
Framework



ICE Becomes Part of Endpoint: In the composite strategy, ICEs are incorporated into the outcome by assigning specific values, modifying the endpoint attribute of the estimand.

For example, in an asthma study, participants who discontinue treatment might be assigned a moderately abnormal FEV1 value (e.g., 60%) or a severely abnormal value (e.g., 50%), with each choice defining a distinct estimand



Clinical Justification Required: The value assigned to ICEs (e.g., for death or dropout) must be clinically meaningful; poorly chosen values can lead to misinterpretation or misleading conclusions about the treatment effect.

Application of Trimmed Means in Clinical Trials for Intercurrent Events

Trimmed Means as a Robust Approach for Missing Data (e.g., Dropout)

- **Dropout as a “Bad Outcome”:** The trimmed means approach treats patient dropout (typically MNAR) as a proxy for poor clinical outcomes, rather than merely missing data, offering a more realistic estimation of treatment efficacy.
- **Trimming Post-Imputation:** Missing values are imputed with the worst observed outcome, and then a portion of the distribution, including these imputed values, trimmed, underlining patients who stayed in the trial and had better outcomes.

Integrating Trimmed Means with the Composite Strategy for ICEs

- **Implicit Composite Endpoint:** The trimmed means method integrates intercurrent events into the endpoint by assigning a worst-case value, efficiently implementing a composite strategy without taking to define multiple explicit outcomes.
- **Rank-Based Robustness:** The actual imputed value need not be precise, only “**bad enough**” to be trimmed making the method robust and less sensitive to the exact numeric imputation.

Scenarios Where Trimmed Means are Particularly Applicable (e.g., MNAR data)

- **Best for MNAR Situations:** Trimmed means are particularly well-suited for Missing Not At Random (MNAR) scenarios, where missingness (e.g., dropout due to inefficacy) reflects a clinically meaningful poor outcome.
- **Not for MAR Data:** They are not appropriate for Missing At Random (MAR) cases, where methods like multiple imputation are more suitable; careful justification is required for use based on the nature of missingness and clinical rationale.

Statistical Methodology of Trimmed Means

Detailed Calculation of Trimmed Means and Trimming Percentages

Order Data: All observations in the dataset are first arranged in ascending order

Determine Trimming Count: For a sample of n observations and a specified trimming percentage α (expressed as a decimal, e.g., 0.2 for 20%), the number of observations to be removed from each end of the ordered data is calculated as $g = \text{floor}(\alpha * n)$.

Remove Extremes: The calculated number of observations, g , is then discarded from both the lowest and highest ends of the ordered dataset.

Calculate Mean: Finally, the arithmetic mean of the remaining $n - 2g$ observations is computed.

Step	Data
Original (n=20)	39, 92, 75, 61, 45, 87, 59, 51, 87, 12, 8, 93, 74, 16, 32, 39, 87, 12, 47, 50
Sorted	8, 12, 12, 16, 32, 39, 39, 45, 47, 50, 51, 59, 61, 74, 75, 87, 87, 87, 92, 93
Trim Count (g)	4 (20% of 20)
Trimmed Values Removed	Lowest: 8, 12, 12, 16; Highest: 87, 87, 92, 93
Remaining Values	32, 39, 39, 45, 47, 50, 51, 59, 61, 74, 75, 87
Mean and Median	53.30 and 50.50
Trimmed Mean	54.92

One-Sided Trimmed Means (Permutt and Li Approach):

The Permutt and Li approach is a specialized application of trimmed means, particularly suited for continuous outcomes where missing values are explicitly assumed to represent poor health outcomes.

This method is designed to address missing data that are Missing Not At Random (MNAR), where the reason for missingness (e.g., patient dropout due to lack of efficacy) provides information about the patient's outcome.

Define New Outcome (U):

- For observed outcomes, $U_i = Y_i$
- For each patient, we define a new outcome **U_i** .
- If the patient's outcome is observed:

Set $U_i = Y_i$ (just use the real value)

For missing outcomes, assign $U_i = \min(Y) - \epsilon$ (slightly worse than worst observed).

- If the patient's outcome **is missing (e.g., dropout)**:
- Assign **$U_i = \min(Y) - \epsilon$**
- This means assign them a value **just slightly worse** than the **worst observed value** (so they rank at the very bottom).
- Example: if the worst observed value is 30, assign 29.999 ($\epsilon = 0.001$)

Statistical Methodology of Trimmed Means (Cont.)

Trim Distribution:

- One-sided trimming from the poor-outcome end.
- Trim at least the maximum dropout percentage between treatment arms.
- Example: if one group has 15% dropout, you trim at least 15% from **each group**.
- Assigned values for missing data ensure these are trimmed out but aren't directly averaged.

Calculate Trimmed Means:

- Compute arithmetic mean of the remaining $(1 - \alpha)\%$ observations in each arm after trimming.
- Results in μ^{T1} and μ^{T0}

Estimate Treatment Effect:

- Treatment effect $\mu^{T\Delta} = \mu^{T1} - \mu^{T0}$
- Use permutation tests for inference and confidence intervals.

Statistical Methodology of Trimmed Means (Cont.)

Statistical Assumptions:

Missing Values as Poor Outcomes: The method is fundamentally designed for scenarios where missing values are justifiably assumed to represent poor health outcomes.

Ability to Rank Outcomes: The approach requires that all outcomes, including those that are "missing" but assigned a "worst" value, can be meaningfully ranked.

Location Family Assumption

Assumes treatment and control outcome distributions differ only by a constant shift (Δ).
Enables valid comparison of trimmed means as proxies for overall average treatment effect.

Strong MNAR Assumption

Assumes all missing outcomes fall below the trimming threshold.
Ensures trimming captures all "bad" outcomes, maintaining consistency between observed and composite outcome distributions.
Some robustness to mild violations of this assumption is supported by simulations.

Combining Trimmed Means with Multiple Imputation:

In situations where the reasons for study discontinuation are known and can be categorized, the trimmed means approach can be combined with Multiple Imputation (MI) to enhance precision. This combined methodology involves:

- **1. Categorizing Missing Data:** Missing data are partitioned into observed data (Yr1), Missing At Random (MAR) or Missing Completely At Random (MCAR) data (Yr2) and Missing Not At Random (MNAR) data (Yr3).
- **2. Multiple Imputation for MAR/MCAR Data:** Multiple imputation is performed specifically, for the Yr2 component, assuming a valid imputation model based on the MAR assumption. This process generates m imputed datasets for Yr2.
- **3. Trimming for MNAR Data:** The Yr3 (MNAR) observations are then handled by the trimmed means approach, meaning they are effectively trimmed out of the analysis after being assigned a "worst outcome" value.
- **4. Calculating Trimmed Means on Imputed Datasets:** For each of the m partially imputed datasets, the trimmed means statistic ($\mu^{(l)}T\Delta$) is calculated.
- **5. Summarizing Results:** Rubin's rules are subsequently applied to combine the results from the m partially imputed datasets, yielding a combined estimate ($\mu T\Delta$) and its associated variance.

A Phase III, Multicenter, Randomized, Double-Masked, Active Comparator-Controlled Study to Evaluate the Efficacy and Safety of Faricimab (RO6867461) in Patients With Diabetic Macular Edema (YOSEMITE AND RHINE) → **NCT Number: NCT03622580**

- The primary efficacy endpoint is the change from baseline in BCVA averaged over Weeks 48, 52, and 56.

➤4.4.1.4 Supplementary Analyses:

Trimmed Mean Analysis: The trimmed mean analysis will be used to assess the difference in BVCA(Best Corrected Visual Acuity) between each faricimab treatment group and the active comparator using a truncated distribution, truncating patients with the worst outcome, with the assumption that patients have the worst outcome after non-COVID-19 related intercurrent events.

Study Population (ITT / Treatment-Naïve DME Patients)

BCVA Assessments
(Weeks 48, 52, 56)

Intercurrent Events Handling

- Non-COVID-19 events → Assume worst outcome
- COVID-19 events → Censor post-event data

Patient Classification

- Completers (≥ 1 BCVA assessment)
- Missing at random → Excluded
- “Must be trimmed” patients identified

Ranking & Trimming

- Fit ANCOVA (covariates only)
- Rank by adjusted BCVA values
- “Must be trimmed” ranked lowest
- Trim equal fractions (nominally 10%)

Trimmed Dataset

(Best ~90% per treatment arm)

The best $(1-0.1)*100\%(=90\%)$ in each group will be used for the analysis

ANCOVA on Trimmed Data

- Treatment effect estimated
- Adjusted trimmed mean difference

Permutation Inference

- ~30,000 random permutations
- Re-rank → re-trim → re-fit ANCOVA
- CI and p-values derived

Regulatory Landscape and FDA Approval



ICH E9(R1) and Regulatory Acceptance of Estimands

The ICH E9(R1) addendum introduces a structured estimand framework that enhances clarity, consistency, and regulatory alignment in defining and analyzing treatment effects, promoting transparent handling of intercurrent events and robust sensitivity analyses in clinical trials.



Examples of Clinical Trials Using Trimmed Means as Primary Efficacy Endpoint Leading to FDA Approval

While the trimmed means approach is recognized for its robustness and aligns with ICH E9(R1) principles, there is limited public evidence of its use as a primary efficacy endpoint in FDA-approved trials; it is more commonly applied in sensitivity analyses to assess the impact of missing data.

Regulatory Landscape and FDA Approval (Cont.)



Trimmed Means and the Eye Disease Study:



A **Phase III trial** for an **eye disease** (likely ophthalmology) considered using the **trimmed mean** method to analyze **visual acuity** (how well a person can see).



The **FDA** reportedly **recommended or supported** the idea of using a **trimmed mean approach** for that outcome.



This was just **conceptual or proposed** — it was **not confirmed** that the **trimmed mean was used as the primary efficacy endpoint** in the final, approved submission.



There's **no public documentation** that the FDA **approved a drug** based solely on a **trimmed mean as the primary analysis method**.

- Trimmed means can be computed in SAS using procedures like PROC UNIVARIATE. For instance, PROC UNIVARIATE supports a `trimmed=` option.

```
/* Set ODS to select and output the trimmed means table */
ods select trimmedmeans;
ods output trimmedmeans=my_trimmed_means_data;

/* Invoke PROC UNIVARIATE to calculate the 20% trimmed mean
*/
proc univariate trimmed = 0.2 data=sashelp.shoes;
var sales; /* Specify the variable for which to calculate the
trimmed mean */
run;
/* Optionally, view the created dataset */
proc print data=my_trimmed_means_data;
run;
```

➤ **Handling of Outliers and Missing Data:**

Trimmed means provide a stable estimate of central tendency by removing extreme values, making them effective in trials with outliers, skewed data, or non-random missingness.

➤ **ICH E9(R1) and Composite Strategy:**

The method fits well within the ICH E9(R1) estimand framework, especially when using the composite strategy for handling intercurrent events like treatment discontinuation.

➤ **Clinical Interpretation of Dropout as Outcome:**

By treating dropout or other ICEs as "bad outcomes" and trimming them, the method shifts the focus from imputing data to redefining the estimand for a more clinically meaningful subpopulation.

➤ **Need for Careful Assumption Justification:**

Key assumptions, such as the strong MNAR and location family assumptions in the Permutt and Li approach, must be clearly justified to ensure the resulting estimand is valid and interpretable.

➤ **Limited Use in Regulatory Approvals:**

While methodologically strong, trimmed means are rarely cited as the primary basis for FDA approvals and are more often used in sensitivity analyses or for secondary estimands in clinical trials.

Thank you!



The Global Healthcare Data Science Community

Contact Channels

 UK +44 1843 609600
 office@phuse.global
 phuse.global

Social Media

 [@phusetwitta](https://twitter.com/phusetwitta)
 [/phusebook](https://facebook.com/phusebook)
 [/phusetube](https://youtube.com/phusetube)
 [/company/phuse](https://linkedin.com/company/phuse)