Paper RE04

# A Beginner's Guide to Overcome the Attrition Bias in Clinical Research: The Method of Inverse Probability of Attrition Weighting (IPAW)

Parichay Srivastava, Cytel, Pune, India

## ABSTRACT

Longitudinal studies play a crucial role in understanding changes over time, but they are often subject to attrition bias. Attrition bias occurs when participants drop out from a study or are lost to follow-up and the characteristics of those who drop out differ from those who remain, leading to biased estimates. To address this issue, Inverse Probability of Attrition Weighting (IPAW) has emerged as a powerful technique. IPAW aims to adjust for attrition bias by reweighting the observations in the sample. It involves assigning weights to participants based on the probability of remaining in the study at each time point using predictive modelling technique accounting participant characteristics and other relevant factors over time. A competing alternative method, namely propensity score matching involves matching participants who have similar propensity scores.

In conclusion, IPAW offers a valuable tool for researchers looking to reduce attrition bias in longitudinal studies and provide efficient results.

**KEY WORDS:** Attrition bias, Inverse probability of attrition weighting, propensity score

.

## INTRODUCTION

Clinical research plays a crucial role in advancing medical knowledge, evaluating treatment efficacy, and improving patient outcomes. However, the validity and reliability of study results heavily rely on minimizing biases that may arise during the research process. One such bias that poses a significant challenge in clinical research is attrition bias, which occurs when the loss of participants during a study is not random and leads to systematic differences between the remaining participants and those who dropped out. Failure to adequately address attrition bias can compromise the internal and external validity of the findings, potentially leading to erroneous conclusions and misguided clinical decisions.

To establish the context and significance of this study, it is crucial to examine the existing body of literature on attrition bias in clinical research. Previous studies have highlighted the adverse consequences of attrition bias, emphasizing the need for robust methods to address this issue. Although it is not possible to define an acceptable attrition rate applicable to all studies in general due to the extent of possible variation across studies, attrition rate of less than 5% is commonly acceptable and not be a concern. However, there are few studies which state that attrition between 5% and 20% may still be a concern and attrition rate above 20% is more likely to produce biased estimates and can impact the outcome of the trial.

To mitigate the impact of attrition bias, researchers have developed various statistical methods, one of which is the method of Inverse Probability of Attrition Weighting (IPAW). IPAW is a rigorous and increasingly popular technique that allows researchers to account for the complex relationship between participant attrition and study outcomes. Using weighted data based on the inverse probability of attrition, IPAW effectively balances the characteristics of the remaining participants with those who dropped out, thereby minimizing the bias introduced by attrition. The purpose of this research paper is to provide a comprehensive beginner's guide to the method of IPAW and its application in clinical research. We aim to equip researchers with a clear understanding of the underlying principles, assumptions, and practical implementation of IPAW, enabling them to address attrition bias effectively. Furthermore, we will provide a non-technical comparison of IPAW with two traditional methods to handle attrition bias namely complete case analysis (CCA) and propensity score matching (PSM).

In conclusion, this research paper aims to empower researchers with the knowledge and skills required to overcome attrition bias in clinical research using the method of IPAW. By addressing this pervasive bias, researchers can enhance the validity and reliability of their study findings, ultimately contributing to evidence-based medical practices and improved patient care.

## ATTRITION BIAS:

Attrition bias, also known as dropout bias or selective attrition, occurs when there is a systematic difference between the participants who remain in the study and those who drop out. It can introduce bias in the estimation of treatment effects and compromise the internal validity of the study.

## CAUSES OF ATTRITION:

Participants may drop out or be lost to follow-up for various reasons, such as:

1. **Adverse Events:** Participants experiencing adverse effects from an intervention or treatment may be more likely to withdraw from the study.
2. **Loss of Interest:** Some participants may lose interest in the study over time, leading to attrition.
3. **Logistical Reasons:** Practical difficulties, such as transportation issues or scheduling conflicts, may result in participants dropping out.
4. **Health-related Factors:** Participants with deteriorating health conditions might be more prone to attrition.
5. **Non-Compliance:** Participants who do not adhere to study protocols may be more likely to drop out.

## TYPE OF ATTRITION:

a) **Random Attrition:** Participants drop out due to various reasons such as adverse events, lack of interest, or inability to continue participation. Non-random attrition can lead to imbalances in the characteristics of the remaining participants, affecting the representativeness of the sample.

b) **Informative/Systematic Attrition:** This type of attrition occurs when the reason for dropout is related to the study outcomes. For example, in a clinical trial evaluating a new medication, participants experiencing severe side effects may be more likely to drop out. Informative attrition can seriously bias treatment effect estimates.

## EXAMPLES OF ATTRITION BIAS IN CLINICAL RESEARCH:

Let us consider a scenario to illustrate the impact of attrition bias in a clinical trial:

**Example: Clinical Trial Evaluating a New Medication for Depression**

In this example, a pharmaceutical company conducts a randomized controlled trial (RCT) to assess the efficacy of a new antidepressant medication compared to a placebo in treating depression. The study enrolls 200 participants with diagnosed depression, randomly assigning them to either the treatment group (new medication) or the control group (placebo) in 1:1 ratio.

**Attrition Scenario:** Over the course of the trial, 30 participants dropped out of the study. Most of the participants who dropped out were in the treatment group and cited severe side effects, such as nausea and dizziness, as the reason for discontinuing the medication.

**Impact of Attrition Bias:** The participants who dropped out due to side effects may have experienced different treatment outcomes than those who completed the study. This informative attrition introduces bias, as the participants remaining in the treatment group may have a milder form of depression and tolerate the medication better. Consequently, the observed treatment effect may be overestimated, leading to the new medication appearing more effective than it truly is.

There are few statistical methods to handle such kinds of situations arising due to uncontrolled factors. One common approach to deal with this problem is complete case analysis (CCA), where only units that are complete on the variables in an analysis are included. However, it may potentially lead to selection bias if the missing data is not missing completely at random (MCAR). Also, CCA discards cases with missing data, leading to a reduction in sample size and potential loss of valuable information. Hence, there is a need for such a method that can mitigate the attrition bias without loss of information from the data. One such method is Inverse Probability of Attrition Weighting (IPAW).

IPAW analysis is needed to address attrition bias in research studies, particularly in situations where participants drop out or are lost to follow-up. There are several reasons why IPAW analysis is crucial:

**a) Mitigating Bias:**

Attrition bias can lead to systematic differences between participants who remain in the study and those who drop out. These differences can introduce bias in treatment effect estimates or associations, potentially distorting the true relationship between variables. IPAW helps adjust for the differences between participants who dropped out and those who remained, reducing the potential for bias.

b) **Improving Sample Representativeness:**
Participants who drop out from a study may differ in their characteristics from those who complete the study. This can lead to a biased and non-representative sample. IPAW reweights the data to make it more representative of the original target population by giving higher weights to individuals who are less likely to drop out, ensuring a more balanced sample.

c) **Strengthening Internal Validity:**
Attrition bias can threaten the internal validity of a study by making it difficult to establish causal relationships between variables accurately. IPAW analysis helps improve internal validity by accounting for potential biases related to attrition, allowing researchers to draw more accurate causal inferences.

d) **Enhancing External Validity:**
When attrition leads to a non-representative sample, the generalizability of study findings to the broader population can be compromised. IPAW analysis helps improve the external validity of the study by creating a more balanced and representative sample, increasing the likelihood that the results can be generalized to the target population.

e) **Handling Missing Data:**
Attrition often results in missing data, which can impact the precision and power of statistical analyses. IPAW provides a mechanism to address missing data by reweighting the available information and incorporating it into the analysis effectively.

## INVERSE PROBABILITY OF ATTRITION WEIGHTING

Inverse Probability of Attrition Weighting (IPAW) analysis is a statistical technique used to address attrition bias in research studies, particularly in longitudinal studies or randomized controlled trials (RCTs) where participants drop out or are lost to follow-up.

IPAW aims to reweight the data to account for the missing information due to attrition, making the analysis more robust and representative of the original target population. It involves assigning weights to each participant based on the inverse (reciprocal) of probability of their attrition status. In other words, participants who have a low probability of dropping out are given higher weights, while those with a high probability of dropping out are given lower weights.

## A STEP-BY-STEP APPROACH OF IPAW PROCEDURE

***Step 1: Identify the Outcome and Covariates***

- Define the outcome variable(s) you want to analyze, such as treatment response, health outcome, or any other relevant endpoint.
- Identify the covariates (potential confounders) that could be associated with attrition and needed to be included in the analysis to control potential sources of bias.

***Step 2: Assess Attrition and Missing Data***

- Determine the extent of attrition in your dataset by identifying participants who dropped out or are lost to follow-up at various time points.
- Examine the patterns of missing data to understand if there is any systematic relationship between missingness and participant characteristics.

***Step 3: Estimate the Propensity Scores***

- Propensity scores represent the probability of attrition for each participant, given their covariates. This can be estimated using logistic regression or other appropriate methods as per the nature of response variable.
- In the logistic regression model, the outcome variable should be attrition status (binary: 0 for dropout, 1 for completion), and the covariates should include relevant baseline characteristics.

***Step 4: Calculate the Inverse of Probability Weights***

- For each participant, calculate the inverse (reciprocal) of probability of attrition weight (IPW) based on their propensity score. The IPW or unstabilized weights is calculated as the reciprocal of the propensity score for each participant.

***Step 5: Check for Extreme Weights***

- Inspect the distribution of the inverse probability weights. Look for extreme weights, as these may indicate potential issues in the propensity score estimation.

- Consider stabilizing the weights if necessary to avoid unnecessary influence of extreme weights on the analysis. The stabilized weights can be calculated as: ratio of proportions of patients with favorable events (in our case, patients with complete data) to the predicted probability of attrition obtained in Step 3.
- While both types of weights, unstabilized and stabilized, are used in IPAW to adjust for attrition bias, the key difference lies in their stability and efficiency. Unstabilized weights are simpler but may be less reliable in extreme cases, for example, when probability of attrition is close to 0 or 1. While stabilized weights are designed to be more robust and precise.

***Step 6: Apply the Weights in the Analysis***

- Apply the inverse probability weights to your analysis model. This weighting procedure gives more importance to participants who are more similar to those who dropped out, thus making the sample more representative of the original population.
  For example, if you are using a regression model to estimate treatment effects, incorporate the IPW as a weight in the regression analysis.

Let us try to understand the practical implication of IPAW using a case study.

## EXAMPLE DATA

This was a multicenter, multinational, randomized, placebo-controlled, double-blind,
parallel-group study involving patients with multiple sclerosis. There were a total of 1088 patients randomized (1:1) into two study arms (Treatment and Control). However, 772 patients were identified with available MRI data at timepoints baseline and Year 2 (Week 108). Further, more exclusion happened due to technical limitations identified during the trial resulting in a total of 318 subjects meeting the quality criteria. The intention of the IPAW analysis was to make use of the full MRI data that could have given us an improved treatment result.
Here the analysis population includes those subjects who had non-missing MRI assessments at baseline and non-missing baseline brain volume change at Year 2 (Week 108). Since brain volume loss was expected to slow down over time in the treatment arm, reduction will indicate efficacious results.
Now, to proceed with further calculations, first we need to calculate the probability of attrition. For this we have used logistic regression with list of covariates which could be associated with the attrition in the study and calculate the predicted probability. To implement the logistic regression for calculating probability of attrition we classified the data into binary variable. Hence, subjects who have full MRI data were classified into a binary indicator as 1 and remaining subjects with missing MRI data as 0.

## BASELINE CHARACTERISTICS:

The baseline characteristics of importance are: Age, gender, time since first diagnosis of MS, baseline EDSS score, proportional GD lesion at baseline, prior DMT using in past 2 years, total lesion volume, T2 lesion volume, normalized brain volume, T25FTW Z score and fatigue impact score.

Multivariate logistic regression with binary variable (0 and 1, as discussed above) based on availability of complete MRI data as response and above listed baseline characteristics as covariates was performed using SAS® procedure namely PROC LOGISITC to obtain the predicted probabilities.

Further, unstabilized weights were derived as reciprocal of predicted probabilities. We have also calculated the proportion of patients who have non-missing MRI assessment. Then the stabilized weights were calculated as: ratio of proportions of patients with MRI data to the predicted probability obtained using logistic regression.
Once we have calculated the weights, we utilized them to conduct the weighted analysis.

Table 1 below describes the summary based on the characteristics of the unweighted analysis (original population with attrition) and weighted analysis (after implementing IPAW) for placebo and treatment arm.

| Table 1: Summary of baseline characteristics | | | | |
|---|---|---|---|---|
| **Characteristics** | **Placebo** | | **Treatment** | |
| | **Unweighted analysis** | **Weighted analysis** | **Unweighted analysis** | **Weighted analysis** |
| Age, mean (SD) | 36.80 (8.87) | 36.68 (8.10) | 38.69 (7.19) | 38.57 (7.50) |
| Gender, Female n (%) | 81 (85.3%) | 78 (78.1%) | 96 (79.3%) | 94 (73.4%) |
| Time since first diagnosis of MS (years), mean (SD) | 4.17 (4.71) | 4.25 (4.51) | 6.82 (6.06) | 6.42 (5.42) |
| Baseline EDSS score, mean (SD) | 2.43 (1.22) | 2.71 (1.26) | 2.50 (1.15) | 2.67 (1.17) |
| Proportional GD lesion at baseline, (=0) n (%) | 52 (55.91) | 50 (54.95) | 84 (69.42) | 83 (70.94) |
| Prior DMT using in past 2 years | 19 (20.0%) | 73 (78.0%) | 39 (32.2%) | 69 (67.4%) |
| Total lesion volume, mean (SD) | 18.23 (18.56) | 20.05 (18.86) | 15.76 (15.35) | 17.13 (15.58) |
| T2 lesion volume, mean (SD) | 15.35 (15.72) | 16.83 (15.92) | 13.14 (12.98) | 14.20 (13.02) |
| Normalized brain volume, mean (SD) | 1496.71 (80.10) | 1497.41 (78.88) | 1494.45 (72.93) | 1497.13 (72.78) |
| T25FTW Z score, mean (SD) | 0.00 (0.88) | -0.21 (1.00) | -0.11 (1.28) | -0.69 (2.51) |
| Fatigue impact score, mean (SD) | 36.56 (30.70) | 41.60 (30.60) | 50.35 (38.85) | 57.90 (39.22) |

From Table 1, it can be clearly examined that there is a fair balance in the baseline characteristics of participants based on original data having attrition and weighted analysis after implementation of IPAW. Therefore, weighting does not create any problem or imbalance in the baseline characteristics of the participants.

Table 2 below provides the results after conducting a linear regression analysis separately using the original data (with attrition) and stabilized weights after implementing IPAW on annualized change from baseline to year 2 in brain volume:

| Table 2: LSMeans results with and without weights | | | | |
|---|---|---|---|---|
| **Analysis Type** | **Treatment** | **LSMeans** | **SE** | **95% CI** |
| Original data with attrition | Placebo | -13.96 | 22.16 | -57.63, -90.56 |
| | Treatment | -51.7 | 19.71 | -90.56, -12.84 |
| With IPAW | Placebo | -21.31 | 22.27 | -65.22, 22.59 |
| | Treatment | -85.79 | 18.85 | -122.97, -48.63 |

It can be clearly observed that if we had all the data available, we would have experienced approximately 65% less brain volume loss in the treatment arm. The intention here is not to compare the two estimates but to provide a detailed process for implementing IPAW and highlight its importance if there is attrition in a clinical trial to ensure the internal validity of the trial.

As discussed earlier in this research paper, there are few traditional methods that can be used to handle attrition bias, however, IPAW has some benefits over the others. Here, we have discussed a non-technical comparison of IPAW with two such methods namely complete case analysis (CCA) and propensity score matching (PSM).

| Table 3: Comparison of IPAW with traditional methods to handle attrition bias | | | |
|---|---|---|---|
| | **IPAW** | **CCA** | **PSM** |
| **Approach** | Assigns weights based on estimated attrition probabilities to balance sample characteristics | Includes only participants with complete data, excludes individuals with missing data | Matches participants having missing data with similar retained participants based on observed characteristics |
| **Utilization of all available data** | Yes | No | Partial (Depends on matched participants) |
| **Handling differential attrition** | Yes | No | Partial (Depends on matching variables) |
| **Addressing biases introduced by attrition** | Yes | No | Partial (Depends on matching quality) |
| **Preserving internal validity** | Yes | Risk of selection bias | Partial (Depends on matching quality) |
| **Treatment effect estimation** | Overall, more accurate estimates | Potential bias due to exclusion of individuals with missing data | Potential bias if unobserved confounders are not accounted for |

## CONCLUSION

In conclusion, this research paper has provided a comprehensive beginner's guide to overcoming the attrition bias in clinical research using the method of Inverse Probability of Attrition Weighting (IPAW). Attrition bias is a common challenge in clinical studies, and addressing it is crucial to ensure the validity and reliability of research findings. Throughout this paper, we have discussed the underlying causes of attrition bias, the importance of recognizing its presence, and the potential consequences it can have on study outcomes. We have also presented a detailed explanation of the IPAW method, outlining its principles and practical implementation steps.

By delving into the methodology, we have demonstrated how researchers can effectively utilize IPAW to adjust for attrition bias in their study designs. This method offers a robust and statistically sound approach to reweighting the observed data, effectively accounting for missing data while preserving the integrity of the research. Furthermore, we have highlighted non-technical comparison with two traditional methods complete case analysis and propensity score matching to handle attrition bias acknowledging its successful application. In essence, this beginner's guide serves as a valuable resource for researchers looking to enhance the quality of their clinical research by mitigating attrition bias. By implementing the method of Inverse Probability of Attrition Weighting, researchers can contribute to more accurate, reliable, and actionable scientific knowledge, ultimately advancing our understanding of health and medical phenomena.

This research paper has laid a foundational understanding of the IPAW methodology, offering a comprehensive guide to mitigate attrition bias in clinical research, particularly within the realm of medical and health sciences. However, as with any analytical approach, IPAW is not devoid of limitations and potential areas for improvement. Further research should aim to explore the robustness and efficiency of IPAW in diverse clinical research settings, considering varying degrees of attrition rates and study designs. Additionally, investigating the comparative effectiveness of IPAW against other attrition bias correction techniques would provide valuable insights, aiding researchers in selecting the most appropriate methodology for their specific study context. Overall, continued research in this area holds the promise of refining and optimizing the application of IPAW, ensuring its efficacy and relevance in addressing attrition bias and advancing the quality of clinical research.

## REFERENCES


1. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4351790/
2. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1444839/
3. https://www.withpower.com/guides/attrition-in-clinical-research
4. https://www.stratos-initiative.org/sites/default/files/2022-10/00491241221113873.pdf
5. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4351790/
6. https://www.ingentaconnect.com/content/wk/ede/2018/00000029/00000001/art00018

## ACKNOWLEDGEMENT

The authors would like to thank all the colleagues at Cytel who gave their valuable input as and when required.

## CONTACT INFORMATION

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

**Author Name:** Parichay Srivastava

**Company:** Cytel Statistical Software & Services Pvt. Ltd.

**Address:** Lohia-Jain IT Park– A Wing, Survey No.150, Paud Road, Kothrud, Pune, India - 411038

**Email:** Parichay.Srivastava@cytel.com