Paper PP10

Tipping Point Analysis in SAS® and R: A Comparative Approach to Handling Missing Data in Clinical Research

Jeet Agrawal, Parexel International (India) Private Ltd., Hyderabad, India
Sushil Kale, Parexel International (India) Private Ltd., Hyderabad, India

ABSTRACT

This paper explores the implementation of Tipping Point Analysis (TPA) in SAS® and R, comparing their approaches to assessing the robustness of clinical trial results in the presence of missing data. TPA is a vital sensitivity analysis method, helping to determine how much missing data can influence study conclusions. We discuss the theoretical foundations of TPA and demonstrate its practical application using SAS® and R, utilizing simulated data representing a typical randomized controlled trial with a continuous primary outcome and various patterns of missing data. The paper covers the syntax, functionality, and output interpretation for TPA in each platform, highlighting their strengths and limitations. We examine how R's flexibility compares with SAS®'s structured procedures and established presence in the industry.

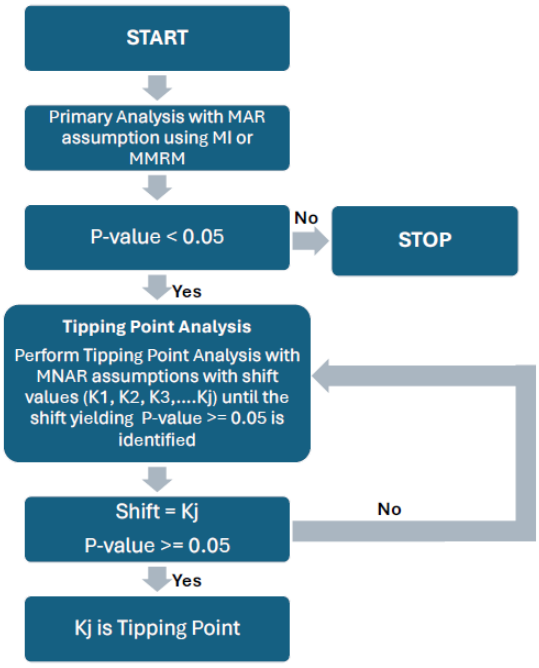
INTRODUCTION

Tipping point analysis is a statistical technique to assess robustness of clinical trial results. Its purpose is to evaluate the impact of missing data on study conclusions. TPA is used when analyses of endpoints are performed based on Missing at Random (MAR) assumptions and is requested by regulatory agencies as a sensitivity analysis under Missing Not at Random (MNAR) assumptions. It explores plausibility of missing data assumptions and can be applied using SAS® and R software. The focus is to identify tipping points where the study conclusion reverses due to different assumptions about missing data.

The objective of this study is to demonstrate Tipping Point Analysis implementation in both SAS® and R on simulated data and compare results.

METHODOLOGY

Decision Criteria for TPA



The decision to proceed with tipping point analysis follows a structured approach: First, determine if there are significant results in primary/secondary analyses. Second, check if there is over 5% missing data. Third, verify if MAR

assumption was used for the primary analysis. If yes to all, proceed with tipping point analysis by applying incremental shift values (K) to predicted values for subjects with missing data.

Simulated Data Description

The study objective was to assess the efficacy of an antidepressant compared to placebo in reducing depression severity over time. The data description included: a randomized controlled trial design with a sample size of 200 participants, divided into Antidepressant (n=100) and Placebo (n=100) groups. The outcome measure was the Hamilton Depression Rating Scale (HAM-D) assessed at time points: Baseline, Week 8, Week 16, and Week 24, with the primary endpoint being HAM-D score at Week 24. Key features included baseline HAM-D scores ranging from 14-30 (moderate to severe depression) and age range of 18-65 years. Key assumptions were Missing at Random (MAR) missing data mechanism with 26% missing data. SAS® was used for simulation.

The 26% missing data rate in this study substantially exceeds the 5% threshold that typically triggers the need for tipping point analysis. This level of missingness makes sensitivity analysis particularly crucial, as it could potentially influence study conclusions if the missing data mechanism deviates from the MAR assumption. The relatively high proportion of missing data provides a realistic scenario for demonstrating the value of TPA in assessing the robustness of treatment effect estimates.

treatment	baseline_score	age	sex	week8_score	week16_score	week24_score
Antidepressant	22	19	F	19	17	14
Antidepressant	24	65	M	21	14	12
Antidepressant	24	29	M	22	21	18
Antidepressant	19	38	M	.	11	8
Antidepressant	22	18	M	15	12	11
Placebo	21	63	F	16	18	19
Placebo	22	27	F	21	18	18
Placebo	20	38	F	16	.	17
Placebo	24	59	F	21	.	15
Placebo	18	37	M	16	.	.

Table 1: Simulated sample data

Note: '.' indicates missing values

PRIMARY ANALYSIS USING SAS®

Missing Data Patterns

Two types of missing data patterns were identified: Monotone Missing Data Pattern, where once a missing value is encountered, all subsequent variables in the ordered list are also missing; and Arbitrary Missing Data Pattern, where values are missing in a scattered way.

Imputation Process

Missing data patterns were examined using PROC MI in SAS®. Imputation was done in two parts due to mixed (arbitrary & monotone) missing data patterns. Part 1 converted arbitrary missing patterns to monotone using MCMC method. Part 2 imputed remaining monotone missing data using regression method.

```
/* Create monotone missing pattern using MCMC method */
proc mi data=monotone_data out=imputed_data nimpute=25 seed=12345;
monotone reg(baseline_score week8_score week16_score week24_score = age);
var age baseline_score week8_score week16_score week24_score;
by treatment;
run;
```

Analysis and Pooling

An ANCOVA model was fit for each imputed dataset to assess treatment effect.

```
/* Perform mixed model analysis for each imputation */
ods output diffs=diffs lsmeans=lsmeans;
proc mixed data=imputed_data;
by _imputation_; class treatment;
```

```
model week24_score = baseline_score treatment / solution;
lsmeans treatment / diff=all cl alpha=0.05;
run;
```

Results were combined across imputations to obtain overall treatment effect estimate and p-value using PROC MIANALYZE

```
/* Use PROC MIANALYZE to pool the results */
proc mianalyze data=diffs;
modeleffects estimate;
stderr stderr;
ods output ParameterEstimates=pooled_results;
run;
```

Parm	Estimate	StdErr	LCLMean	UCLMean	DF	Min	Max	Theta0	tValue	Probt
estimate	-7.886326	0.626770	-9.12199	-6.65066	207.16	-8.650365	-7.283783	0	-12.58	<.0001

The results were statistically significant (p<0.0001). The primary analysis suggested that antidepressant was effective in reducing depression. Due to the presence of missing data and significant result, a sensitivity analysis was recommended to ensure robustness of the results.

SENSITIVITY ANALYSIS WITH TIPPING POINT APPROACH

SAS® Implementation

The sensitivity analysis involved three steps:

- Step 1 involved performing imputation under MAR and then under MNAR assumptions with varying shift values.

```
proc mi data=simulated_data out=imputed_mar nimpute=25 seed=54321;
    var baseline_score week8_score week16_score week24_score;
    mcmc impute=monotone;
    by treatment;
run;

/* Imputation under MNAR */
%macro tipping_point_analysis;
%let shifts = 17 18 19 20 21 22 23 24 25 26 27;
%do i = 1 %to 11;
%let shift = %scan(&shifts, &i);
proc mi data=imputed_mar out=imputed_shifted_&i nimpute=1 seed=67890;
class treatment;
var treatment baseline_score week24_score;
monotone reg;
mnar adjust(week24_score /
adjustobs=(treatment='Antidepressant') shift=&shift);
by _imputation_;
run;
%end;
%mend;
```

- Step 2 involved conducting analysis using PROC MIXED on each imputed dataset.

```
proc mixed data=imputed_shifted_&i;
by _imputation_;
class treatment;
model week24_score = baseline_score treatment;
lsmeans treatment / diff=all;
ods output diffs=diffs_&i;
run;
```

- Step 3 involved pooling results using PROC MIANALYZE.

```
proc mianalyze data=diffs_&i;
modeleffects estimate;
```

```
stderr stderr;
ods output ParameterEstimates=results_&i;
run;
```

SAS® Results

Parm	Estimate	StdErr	LCLMean	UCLMean	DF	Min	Max	Theta0	tValue	Probt	shift
estimate	-3.641856	0.809509	-5.22909	-2.05463	3100.7	-4.233427	-3.224479	0	-4.50	<.0001	17
estimate	-3.441293	0.843551	-5.09517	-1.78742	3656.1	-4.032864	-3.023915	0	-4.08	<.0001	18
estimate	-3.240729	0.878123	-4.96230	-1.51915	4293.3	-3.832300	-2.823351	0	-3.69	0.0002	19
estimate	-3.040165	0.913164	-4.83037	-1.24996	5020.7	-3.631736	-2.622787	0	-3.33	0.0009	20
estimate	-2.839601	0.948623	-4.69925	-0.97995	5847.2	-3.431172	-2.422224	0	-2.99	0.0028	21
estimate	-2.639038	0.984455	-4.56888	-0.70920	6782	-3.230609	-2.221660	0	-2.68	0.0074	22
estimate	-2.438474	1.020620	-4.43916	-0.43779	7834.8	-3.030045	-2.021096	0	-2.39	0.0169	23
estimate	-2.237910	1.057084	-4.31004	-0.16578	9015.9	-2.829481	-1.820532	0	-2.12	0.0343	24
estimate	-2.037346	1.093817	-4.18144	0.10675	10336	-2.628917	-1.619969	0	-1.86	0.0625	25
estimate	-1.836782	1.130794	-4.05332	0.37976	11806	-2.428353	-1.419405	0	-1.62	0.1043	26
estimate	-1.636219	1.167990	-3.92564	0.65321	13438	-2.227790	-1.218841	0	-1.40	0.1613	27

Table 2: SAS® Results

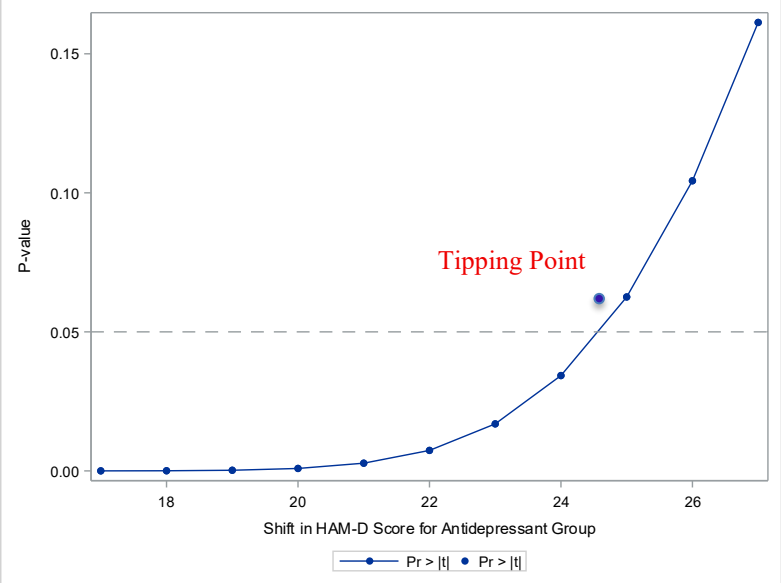


Figure 1: Tipping point plot using SAS®

The antidepressant remained significant up to shift value 24, with the tipping point at shift value 25 (p-value became > 0.05). At shift 25, the treatment effect became non-significant. When the imputed Hamilton Depression Rating Scale (HAM-D) scores for the Antidepressant increased by 25 units, the treatment effect became non-significant. This means that if the true values of the missing data were 25 units higher than the imputed values, the conclusion that antidepressant was effective would no longer hold.

PRIMARY ANALYSIS USING R

R code using "MICE" (Multiple Imputation by Chained Equations) package was implemented. Missing data patterns were examined, multiple imputations were created under MAR assumption using pmm (predictive mean matching), and each imputed dataset was analyzed under MAR with pooled p-values calculated.

Create multiple imputations

imp <- mice(data, m = 25, method = "pmm", print = FALSE)

Analyze each imputed dataset under MAR

```
mar_models <- vector("list", 25)
mar_p_values <- numeric(25)
for (j in 1:25) {
mar_data <- complete(imp, j)
mar_models[[j]] <- lm(week24_score ~ treatment + baseline_score,
                      data = mar_data)
mar_p_values[j] <- summary(mar_models[[j]])$coefficients[2, 4]
}
mar_p_value <- mean(mar_p_values)
```

Result: Pooled p-value for treatment effect: 0

R Tipping Point Implementation

The tipping point analysis function created multiple imputations and applied shifts to imputed values for MNAR implementation. For each shift value, complete datasets were created from imputation and shifts were applied to treatment group's imputed values.

```
tipping_point_analysis <- function(data, shifts = seq(17, 27, by = 1)) {
   imp <- mice(data, m = 25, seed = 54321, print = FALSE)

   for (i in seq_along(shifts)) {
      shift_val <- shifts[i]
      imp_datasets <- vector("list", 25)

      for (j in 1:25) {
         imp_datasets[[j]] <- complete(imp, j)
         missing_indices <- which(is.na(data[[outcome_var]]) &
                                  data[[treatment_var]] == "Antidepressant")
         if (length(missing_indices) > 0) {
            imp_datasets[[j]][missing_indices, outcome_var] <-
               imp_datasets[[j]][missing_indices, outcome_var] + shift_val
         }}}}}
```

Analysis and P-value Extraction

```
# Analyze each imputed dataset and extract p-values
p_values <- numeric(25)
for (j in 1:25) {
  model <- lm(week24_score ~ treatment + baseline_score, data = imp_datasets[[j]])
  p_values[j] <- summary(model)$coefficients[2, 4]  # p-value for treatment
}
```

R Results

	shift	p_value	
1	17	0.0001145240	Significant
2	18	0.0004823658	Significant
3	19	0.0016521290	Significant
4	20	0.0047338687	Significant
5	21	0.0116328586	Significant
6	22	0.0250531705	Significant
7	23	0.0481848891	Significant
8	24	0.0841190378	Not Significant
9	25	0.1351777936	Not Significant
10	26	0.2023889502	Not Significant
11	27	0.2852606287	Not Significant

Table 3: R Results

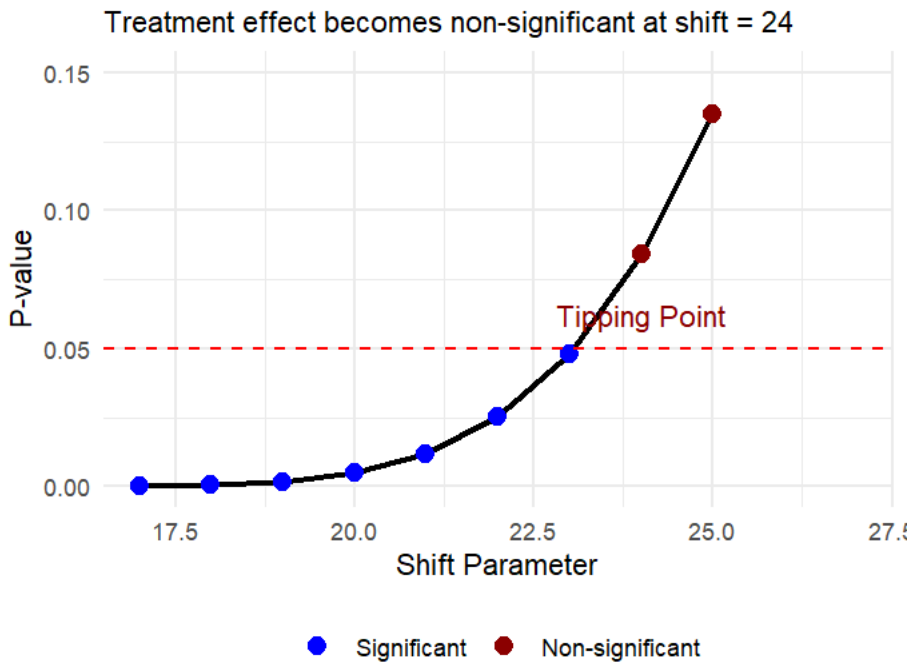


Figure 2: Tipping point plot using R

The antidepressant remained significant up to shift value 23, with the tipping point at shift value 24 (p-value became > 0.05). At shift 24, the treatment effect became non-significant.

COMPARISON OF SAS® AND R

Aspect	SAS®	R
Results	Tipping point at shift = 25	Tipping point at shift = 24
Code Efficiency	More structured, verbose	More concise, flexible
Implementation	PROC MI with structured procedures	mice package with customizable functions
Visualization	ODS Graphics, more setup required	ggplot2, highly customizable with less code
Performance	Optimized for large datasets	Better for iterative analyses, interactive exploration
Methods	Primarily parametric approaches	Default to PMM, fewer distributional assumptions
Regulatory	Established history in submissions	Increasing acceptance, requires standardization
Learning Curve	Steeper for new users, familiar for clinical programmers	Gentler entry, more intuitive for statisticians
Integration	Better with CDISC standards	Superior for end-to-end analysis workflow
Recommendation	Preferred for standard analyses in regulated environments	Ideal for exploratory analyses, complex visualizations

Understanding the One-Unit Difference

The one-unit difference in tipping points between SAS® (shift=25) and R (shift=24) can be attributed to several methodological factors. First, SAS® primarily uses regression-based imputation methods, while R's MICE package defaults to predictive mean matching (PMM). PMM is a semi-parametric method that imputes values by matching predicted values to observed values, while regression imputation uses parametric models. Second, despite using the same random seeds, the underlying random number generation algorithms differ between platforms, introducing minor variation in the imputed values. Third, numerical precision and rounding differences in statistical computations can accumulate across 25 imputations. However, both platforms yield consistent conclusions about the robustness of the treatment effect. The convergence to nearly identical tipping points (24-25) validates the reliability of both implementations.

CONCLUSION

The tipping point range (24-25) provides a more informative sensitivity boundary than a single value, and importantly, represents a shift magnitude that exceeds clinically plausible scenarios given the baseline HAM-D range of 14-30 units. This indicates that the treatment effect is robust to realistic departures from the MAR assumption. Both SAS® and R yield consistent conclusions despite minor implementation differences. The magnitude of shift required (24-25 units) suggests findings are robust to plausible MNAR mechanisms. When interpreting tipping point results, researchers should evaluate whether the shift magnitude required to reverse conclusions is clinically plausible within the context of the outcome measure's scale and the study population's characteristics. In studies with high missing data rates, conducting TPA in both SAS® and R can provide additional confidence in the robustness assessment, as demonstrated by the convergent results in this analysis.

REFERENCES

- Heymans, M.W. and Twisk, J.W.R. (2022) 'Handling missing data in clinical research', Journal of Clinical Epidemiology, 151, pp. 185-188.
- Patil, V. (2024) 'Unveiling the Tipping Point: An In-depth Analysis', PHUSE US Connect 2024.
- Smith, C. and Kosten, S. (2017) 'Multiple Imputation: A Statistical Programming Story, PharmaSUG 2017.
- Yuan, Y. (2014) 'Sensitivity Analysis in Multiple Imputation for Missing Data', SAS® Global Forum 2014.

CONTACT INFORMATION

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

Author name: Jeet Agrawal
Company: Parexel International (India) Private Ltd.
Address: Hyderabad
Work Phone: +91 95886 09641
Email: jeet.agrawal@parexel.com
Website: https://www.parexel.com/

Co-Author name: Sushil Kale
Company: Parexel International (India) Private Ltd.
Address: Hyderabad
Work Phone: +91 98811 35985
Email: sushil.kale@parexel.com
Website: https://www.parexel.com/