I never use the tipping point analysis in my work but I heard this method from my PPD colleague several years ago. This method did not attract my attention until I read a paper written by FDA statistical reviewers who listed the application of “tipping point analysis” as one of the statistical innovations in medical device. Tipping point analysis was initiated a decade ago by division of biostatistics of FDA CDRH. Since then, this simple but appealing method has attracted much attention in medical device community and has been discussed many times at FDA medical device advisory committee panel meetings. So I want to share my understanding of this topic with team today.

Suppose there is a superiority trial comparing investigational medical device product with the active-control in treating heart disease. The primary endpoint is binary outcome, 0 for failure and 1 for success. Both groups have 50 subjects, and 10 missing outcomes. Among the 40 non-missing subjects, there are 36 successes in investigational medical device group and 20 successes in control group.

There are many methods to handle missing values, such as, delete all missing values and use all the complete cases, use the worst case imputation, impute the missing values based on MCAR assumption and etc. Suppose the imputed method based on MCAR assumption is used as primary analysis for this study and the worst case imputation is used as sensitivity analysis. As we can see in the below table, the statistical p value for MCAR imputation is smaller than 0.05 but the p value for worst case imputation is larger than 0.05. The primary analysis gets the significant result but the sensitivity analysis get the non-significant result. How to interpret the conflict result between primary analysis and sensitivity analysis.

A feasible method is to investigate the impact of missing value handling on the study results. As we know, there are 10 missing values in treatment group, so there are 11 possible results for these missing values, from 0 success number to 10 success number for these ten missing subjects. The control group is same with treatment group. So the total possibility is 11\*11, equals to 121. In this graph, all the possibilities for missing values handling in the treatment and control group are displayed with p values.

* The worst case imputation (all the missing values in treatment group are failure and all the missing values in the control group are success) is located in the extreme upper left;
* The best case scenario is located in the extreme lower right;
* The intersection of two dotted lines is MCAR imputation;
* For each ordered possible scenario of missingness, it is possible to categorize the point as to whether the missing scenario changes the study conclusion or not, e.g., from statistical significance to non-significance of hypothesis testing at a pre-specified significance level. The non-significant areas in upper left are tipping points;
* The staircase region that separates the points that correspond to significant study results from those that represent non- significant results, i.e., the region that marks where the study conclusion changes, is called the tipping-point boundary;

We can see clearly that the non-significant cases are only small proportions and significant cases are the dominated proportions. The impact of missing values on the study result is tiny thus we can conclude that the study result is significant with high confidence using this tipping point analysis plot. Also, such a graph with all possible missing data pattern and tipping points can be conveyed to clinical reviewers regarding the robust degree of the study and the impact of missing data.

What is the tipping point? Tipping point is defined as the difference of means or difference of the number of events between the treatment groups in the missing cohort at which the study conclusion is changed. The tipping-point boundary is the staircase **[ˈsterkeɪs]** region that marks where the study conclusion changes. A tipping point analysis replaces the missing value with some values so that the resulting p value of treatment comparison is equal to (or larger but close to) a prespecified significance level (e.g., 0.05). For a dichotomous [**dɪˈkɒtəməs]** clinical outcome, e.g., success or failure at a patient level, the tipping point analysis presents a matrix of all possible patterns for the missing data in the two groups, the investigational treatment group on the y-axis **[ˈæksɪs]** and the control group on the x-axis. For each ordered possible scenario of missingness, it is possible to categorize the point as to whether the missing scenario changes the study conclusion or not, e.g., from statistical significance to non- significance of hypothesis testing at a pre-specified significance level.

Here is the algorithm and SAS code for tipping point analysis with binary outcome. Let n\_tm and n\_cm be the number of missing values in the treatment and control group, respectively. Starting at 0 and ending at n\_cm, the success rate in the control group is changed by adding 1 success at a time; meanwhile, the success rate in the control group is changed by adding 1 success at a time starting at 0 and ending at n\_tm; The p value for proportion difference between treatment group and control group can be produced. If the p-value is greater than 0.05, it is a tipping point. A simple nested SAS do-loops can implement the tipping point algorithm.

As there is no universal best method or one-size-fits-all for handling missing data in a clinical trial setting and when the missing data mechanism is unknown or unverifiable, sensitivity analyses with different methods usually provide a useful tool in the statistical evaluation of robustness of study result. Tipping point analysis is a more powerful tool than other sensitivity since it can display all the possible missing patterns in a graph with p values demonstrating the tipping points and impact of missing values on the study results. Since it is initiated by CDRH statistical reviewers in 2009, it attracted more attention in medical device community and then also applicated in drug clinical trial as it is a simple but appealing method. It has many advantages, for example:

* Tipping-point analysis does not need to postulate any missing data mechanism.
* Tipping-point analysis does not involve model uncertainty and assumptions.
* The result of the tipping-point analysis may be more understandable to non-statisticians than that of the model-based approach.
* It can be used to evaluate the robustness of statistical significance when the extent of missing data is reasonable.

No any advanced statistical analysis should be a substitute for a good clinical plan that minimizes patient dropouts. It is important to continue following patients even after they withdraw from a clinical study. The reasons for loss to follow-up need to be documented and investigated. From the perspective of sponsor, the primary goal of clinical trials is to obtain a statistically significant result. But from the perspective of statisticians, the more important goal is to make sure the significant result reliable and robust. With this aspect, tipping point sensitivity analyses for missing data handling play an important role in the determination of robustness and reliability of study results.

For more information on tipping point analysis for continuous outcome, please read the first paper. For the enhanced display of tipping point algorithm, please read the third paper. The R package TippingPoint is a great tool to produce the tipping point graph.

When the tipping points are just only small proportion in tipping point graph, we can conclude the study result is significant with high confidence. When the proportion of tipping points area is close to 50% percentage or larger than 50% percentages, it indicates that the study conclusion is more sensitive. We may investigate the distribution of baseline characteristics between completers and dropouts, as well as other factors that may have impact on the missing data to make judgment whether or not the tipping-point outcome is likely to happen. The clinical plausibility of the combinations on the tipping point boundary will be discussed in the Clinical Study Report to evaluate robustness of study conclusions to missing data.