Trimmed mean to handle missing/meaningless outcomes - a recommendation from FDA

Fan Tang, David Kardatzke, and Hans Ulrich Burger

Agenda

- Context setting
- Preliminary work/simulation
- Implications

Context setting

- Phase III study in an Ophthalmology indication
- Primary endpoint is visual accuacy usually sampled as a longitudinal endpoint
- Missing data can occur in the range of 10%
- Standard method of analysing was MMRM
  - Treatment policy approach
- For this study a modified approach still based on MMRM was proposed
  - patients potentially considered as worst outcomes are censored at the time of the event (discontinue due to AE or LCE, or receive rescue treatment, etc.)
- Feedback from FDA was not in agreement with proposal but proposed a trimmed mean approach.

Concept of Trimmed Mean Approach

- Applicable to longitudinal outcomes measures at visits - not time to event type endpoints.
- Concept is that missing is not missing at random and is an outcome.
- By trimming equal proportions from each treatment arm, the treatment effect of the remaining better outcomes is estimated
- A good example is missing due to death where death is the ultimate outcome to be prevented but happens at too low a rate to design the study for an overall survival benefit.
- Another example is a chronic treatment in which a subset of patients respond to treatment. The estimand is comparing the better outcome patients between treatment arms

Concept of Trimmed Mean Approach

Approaches for handling this type of missing:

- Imputing the worst value observed in the study or worst possible value
  - Problem, often introduces skewed result, no good bases for choice of value
- Multiple imputation tipping point analysis
  - Problem, once tipping point is known conclusion is not clear - no standard testing approach like p<=0.0500 is applicable
- Trimmed Mean
  - Straightforward unbiased estimate of the trimmed mean treatment effect if assumption holds
  - Trimmed mean is however different from standard mean

Summary of Trimmed Mean Approach

- Define event scenarios that will meet trimming criteria
- Subjects meeting trimming criteria and dropouts (must trim patients) are always ranked the worst. Order the rest of data (completers) using either observed responses or model-based methods (adjusted responses).
- Trim equal fractions of worse subjects for each treatment arm based on the ranked data. The percentage to be trimmed can be either defined as the maximum must trim patient rate in any treatment group (flexible or adpative trimming method) or a fixed percentage (exceeding must trim patient rate in any treatment group).
- Compute a summary statistics (estimate of interest) based on trimmed data.
- Repeat the above steps on permuted data to construct a reference distribution for testing and interval estimation for the summary statistics.

Graphical explanation and potential label presentation

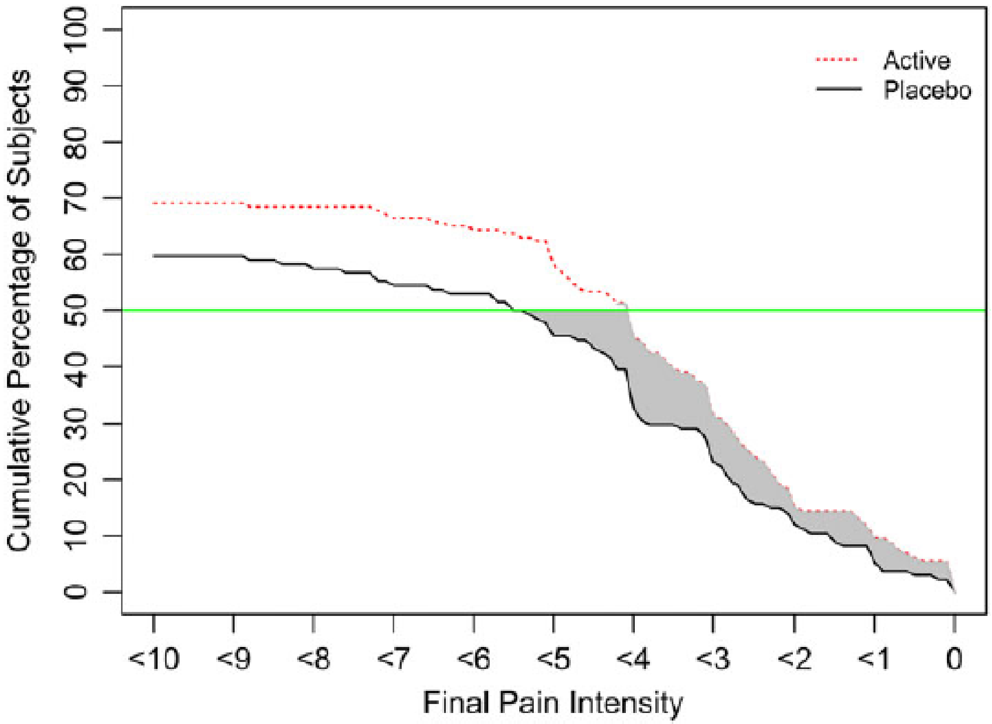

Figure 1. Empirical CDF. The shaded area is proportional to the difference in 50% trimmed means.

Reference: Permutt T, Li F. Trimmed means for symptom trials with dropouts. Pharm Stat 2017; 16(1):21.

Impact on Sample Size

- Assess the relative efficiency for the variance of the trimmed mean under different distributions

Relative Efficiency (Variance of Trimmed Mean / Variance of ordinary sample mean) N of Samples =1000; Length of Sample = 200							
Distribution	Characteristics	Trimmed Mean (One-sided Fixed Trimming)					
		10%	20%	25%	30%	40%	50%
Standard Normal	Symm.	104.8%	109.8%	113.0%	116.9%	125.5%	137.1%
Normal (4,10)	Symm.	102.9%	108.6%	112.0%	115.6%	124.6%	136.3%
t-3 (mu=4)	Heavy tail	78.3%	86.4%	92.2%	99.7%	120.3%	153.0%
t-5 (mu=4)	Heavy tail	92.8%	97.7%	102.1%	107.6%	122.2%	144.8%
t-10 (mu=4)	Heavy tail	99.1%	104.3%	107.9%	111.8%	122.5%	137.8%
t-50 (mu=4)	Similar to Normal	103.6%	110.0%	114.4%	119.6%	130.5%	146.0%
Skewed Normal (4, 10)	Left Skewed, shape=-1	101.6%	104.6%	107.3%	110.5%	117.9%	127.8%
Skewed Normal (4, 10)	Left Skewed, shape=-2	96.1%	96.4%	96.7%	97.2%	99.6%	103.9%
Skewed Normal (4, 10)	Left Skewed, shape=-4	89.5%	84.9%	82.5%	80.2%	76.3%	72.9%

*Depending on the underlying assumption of distribution, trimmed mean approach may/may not reduce the efficiency

Notes (I)

- Although it has been shown that the trimmed mean approach provides an unbiased estimate of the trimmed mean treatment effect it is not an estimate of the population mean.
- Clinicians will need to become familiar with this approach and used to thinking about the clinical significance of differences in the trimmed population.
- The estimand for the trimmed mean approach will be different from common ITT/mITT analysis.
- More fundamentally the trimmed mean is only unbiased for the trimmed population mean under the assumption that the trimming is of truly bad outcome patients in both groups.
    - Example: Assuming you selectively trim only one treatment arm based on a criteria unrelated to outcome and then trim from the other arm the worst patient. This will have a negative effect on the treatment arm trimmed on the criteria unrelated to outcome.

Notes (II)

- If the patient population is largely stable with a small subset that has rapid decline there is a risk of trimming the population where the potential treatment benefit is present.
- Full picture of ITT population will require some description of the population trimmed
- Trimmed mean approach does not really fit the estimand of the mean of the overall population, but the estimand of the mean of the trimmed population. The trimming proportion becomes part of the estimand.
    - But: Why would we select a trimmed mean (in an adaptive setting not even defined)? Why are we not interested in the population mean?

Notes (III)

- Concept in handling dropouts/missing data as clinically relevant information is not new
    - For binary endpoints we move patients with missing data into non-responding category. Usually makes sense and does not affect population based estimand
    - Ranked ANOVA is very similar: Missing data/dropouts get worse ranks
        - However: Ranked ANOVA still estimates the overall population
        - Point estimate of Ranked ANOVA however not easily interpretable
    - Trimmed approach is similar to binary situation, however trimmed mean differentiates among the better outcomes not the worse outcomes. Also like there we are interested in patients doing better and move equal proportions of patients with missing data/drop outs into worse category. However, definition of «better» is not clear and up front defined
    - In survival we would probably move to a composite endpoint. That however seems more transparent