

# Comments on Statistical Issues in May 2013

Commentary

Yong Gyu Park

Department of Biostatistics, The Catholic University of Korea College of Medicine, Seoul, Korea

In this section, we explain the analysis data sets obtained from clinical trials; in the article titled, “Effect of *Lactobacillus gasseri* BNR17 on overweight and obese adults: a randomized, double-blind clinical trial”, published in March 2013 by Jung et al.<sup>1)</sup>

## ANALYSIS DATA SETS IN CLINICAL TRIALS

The following is from the International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use guideline E9,<sup>2)</sup> “Statistical Principles for Clinical Trials”, concerning the analysis data sets obtained from clinical trials. Decisions concerning the analysis set should be guided by the following principles: 1) to minimize bias and 2) to avoid inflation of type I error.

### 1. Full Analysis Set

The intention-to-treat principle implies that the primary analysis should include all randomized subjects. Compliance with this principle would necessitate complete follow-up of all randomized subjects for study outcomes. In this document the term ‘full analysis set’ is used to describe the analysis set which is as complete as possible and as close as possible to the intention-to-treat ideal of including all randomized subjects. Preservation of the initial randomization in analysis is important in preventing bias and in providing a secure foundation for statistical tests.

There are a limited number of circumstances that might lead to exclude randomized subjects from the full analysis set, including the failure to satisfy major entry criteria (eligibility violations), the failure to take at least one dose of trial medication and the lack of any data post randomization. Such exclusions should always be

justified.

Subjects who fail to satisfy an entry criterion may be excluded from the analysis without the possibility of introducing bias only under the following circumstances: 1) the entry criterion was measured prior to randomization; 2) the detection of the relevant eligibility violations can be made completely objectively; 3) all subjects receive equal scrutiny for eligibility violations (This may be difficult to ensure in an open-label study, or even in a double-blind study if the data are un-blinded prior to this scrutiny, emphasizing the importance of the blind review.); 4) all detected violations of the particular entry criterion are excluded.

### 2. Per Protocol Set

The ‘per protocol’ set of subjects, sometimes described as the ‘valid cases’, the ‘efficacy’ sample or the ‘evaluable subjects’ sample, defines a subset of the subjects in the full analysis set who are more compliant with the protocol and is characterized by criteria such as the following: 1) the completion of a certain pre-specified minimal exposure to the treatment regimen; 2) the availability of measurements of the primary variable(s); 3) the absence of any major protocol violations including the violation of entry criteria.

### 3. Safety Set

For the overall safety assessment, the set of subjects to be summarized is usually defined as those subjects who received at least one dose of the investigational drug.

From the definitions of the three analysis data sets above, we can presuppose that, in most cases, the full analysis set has the largest number of observations, followed by the safety set, and the

per-protocol set has the smallest amount of data.

When the full analysis set is used, we usually replace the missing measurements which occurred in the subjects withdrawn after receiving one or more doses, with some conservative values as follows: 1) regard as a failure when the outcome is measured as success or failure; 2) use the value from the carrying forward of the last observation for continuous variables.

## CONFLICT OF INTEREST

---

No potential conflict of interest relevant to this article was reported.

## REFERENCES

---

1. Jung SP, Lee KM, Kang JH, Yun SI, Park HO, Moon Y, et al. Effect of *Lactobacillus gasseri* BNR17 on overweight and obese adults: a randomized, double-blind clinical trial. *Korean J Fam Med* 2013;34:80-9.
2. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH). ICH guideline E9: statistical principles for clinical trials [Internet]. Geneva: International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use; 1998 [cited 2013 May 11]. Available from: <http://www.ich.org/products/guidelines/efficacy/article/efficacy-guidelines.html>.