A Case Study of Clinical Endpoint Bioequivalence Study with Missing Data and Non-compliance Data

Sun W*

Office of Biostatistics, Division VIII, Center for Drug Evaluation and Research, Food and Drug Administration (CDER/FDA), Silver Spring, USA

Abstract

Missing data and non-compliance data questions are especially important in evaluating locally acting generic drugs because primary equivalence analyses in clinical endpoint bioequivalence (BE) studies are based on the per-protocol (PP) population (generally, completers and compliers). However, no case studies have been reported to address the status of missing data and non-compliance in clinical endpoint BE studies. A case study using a clinical endpoint study for topical drugs for treatment of acne vulgaris was used to illustrate the extent and pattern of drop out and non-compliance reported in the Sun et al. meta-analysis results, and whether drop out and non-compliance are random or not random, independent or correlated with each other, balanced or not balanced between the generic (TEST) and the innovator drug, i.e., the reference listed drug (RLD).

Keywords: Generic drugs; Bioequivalence; Missing data; Non-compliance; Meta-analysis

*Corresponding author: Dr. Wanjie Sun

✉ Wanjie.Sun@fda.hhs.gov

Office of Biostatistics, Division VIII, Center for Drug Evaluation and Research, Food and Drug Administration (CDER/FDA), Silver Spring, USA.

Tel: 1-888-463-6332

Introduction

Missing data and non-compliance are very common in clinical trials [1]. They are especially important in clinical endpoint bioequivalence (BE) studies for evaluating locally acting generic drugs because the primary equivalence analyses are usually based on the per-protocol (PP) population (generally, completers and compliers). Sun et al. [2] presented the results from a meta-analysis using six approved Abbreviated New Drug Application (ANDA) based on clinical endpoint BE studies for topical drugs as an example and reported the missing data and non-compliance data status in these ANDA studies. In this paper, we give one case study to illustrate the prevalence and breakdown categories of drop out and non-compliance, and whether drop out and non-compliance are random or not random, independent or correlated with each other, balanced or not balanced between the generic and the innovator drug.

Methods

Clinical endpoint BE studies aim to establish bioequivalence between TEST and RLD [3]. A placebo or vehicle arm (VEH) is usually included in these studies in order to test the sensitivity of the study. The study design is usually a blinded, randomized 3-arm parallel clinical trial. In our case study, there are four visits: baseline (Week 0), Week 4, Week 8, and Week 12 (the test of cure visit). The co-primary endpoints are percent changes from baseline to Week 12 in inflammatory lesion count and in non-inflammatory lesion count. According to the FDA guidance [4], in order to establish bioequivalence between TEST and RLD, a study has to pass two superiority tests (TEST vs. VEH: $H_{0t} : \mu_T \leq \mu_V, H_{1T} : \mu_T > \mu_V$; and RLD vs. VEH: $H_{0R} : \mu_R \leq \mu_V, H_{1R} : \mu_R > \mu_V$; for validation of the sensitivity of the study, and an equivalence test [5] (TEST vs. RLD: $H_0 : \frac{\mu_T}{\mu_R} \leq 0.8 \, or \frac{\mu_T}{\mu_R} \geq 1.25, H_1 : \frac{\mu_T}{\mu_R} < 1.25$ for establishment of equivalence between TEST and RLD, where μ_T and μ_R are the mean percent changes from baseline to Week 12 in inflammatory or non-inflammatory lesion count in TEST and RLD. Only when all three null hypotheses ($H_{0t}, H_{0R}, and H_0$) are rejected, the bioequivalence between TEST and RLD can be established. The primary superiority tests are based on the modified intent-to-treatment (ITT) population, which in general is defined as the randomized subjects who took at least one dose of treatment. The primary equivalence test is based on the per-protocol (PP) population, which is usually composed of completers and compliers. For treatment of acne vulgaris, those who dropped out due to noted lack of efficacy are also included in the PP population. In this paper, we focus on the equivalence evaluation.

Results

Figure 1 shows that among the 285 TEST and 283 RLD randomized subjects, 79.7% of the TEST and 77.7% of the RLD subjects were completers and compliers; 7.7% (TEST) vs. 6.4% (RLD) subjects were both dropouts and non-compliers; 4.9% of the TEST vs. 5% of the RLD subjects dropped out of the study but complied with the study protocol; and 7.7% of the TEST vs. 11% of the RLD subjects completed the study but did not comply with the study protocol. The latter three categories were excluded from the PP population except those dropouts due to noted lack of efficacy (LOE) (0.4% in TEST and 0% in RLD). Therefore, in total, 80.1% of the TEST subjects and 77.7% of the RLD subjects were included in the PP population and used in the primary equivalence analysis between TEST and RLD. In this clinical endpoint BE study, the equivalence of TEST and RLD was established in the co-primary endpoints – both 90% CIs of the ratio of means were contained in the BE margin [0.8, 1.25].

Breakdown of the dropout and non-compliance categories **(Figure 2)** shows that among the non-PP population (TEST: 19.9%; RLD: 22.3%), 12.6% (TEST) vs. 11.3% (RLD) dropped out of the study or were lost to follow up; 9.1% (TEST) vs. 8.8% (RLD) missed three consecutive doses and were non-compliant; 6.3% (TEST) vs. 7.4% (RLD) were out of window for the EOC visit; 6.3% (TEST) vs. 6% (RLD) were either hyper-dosed or under-dosed (i.e., treatment non-compliant); 0.7% (TEST) vs. 1.8% (RLD) took restricted medication; and 0.4% (TEST) vs. 0% (RLD) had other non-compliance. These categories are not mutually exclusive because one subject could drop out and miss three consecutive doses. For example, 2.5% of the TEST and RLD subjects had two dropout and non-compliance categories, and 6.3% of the TEST vs. 5.3% of the RLD subjects had three dropout and non-compliance categories.

Figure 3 reveals that among the 12.6% of the TEST dropouts and 11.3% of the RLD dropouts, most did not specify dropout reasons (TEST: 8.4%; RLD: 7.1%), 2.5% (TEST) vs. 3.2% (RLD) were non-study related (e.g. move out of the town), 1.4% (TEST) vs. 1.1% (RLD) were terminated by the sponsor due to non-compliance or by the subject, and only 0.4% of the TEST subjects dropped out due to noted adverse events (AE) or lack of efficacy (LOE).

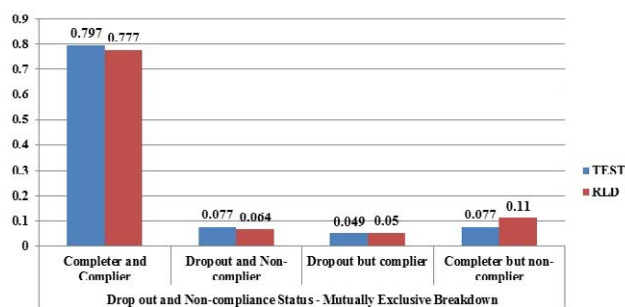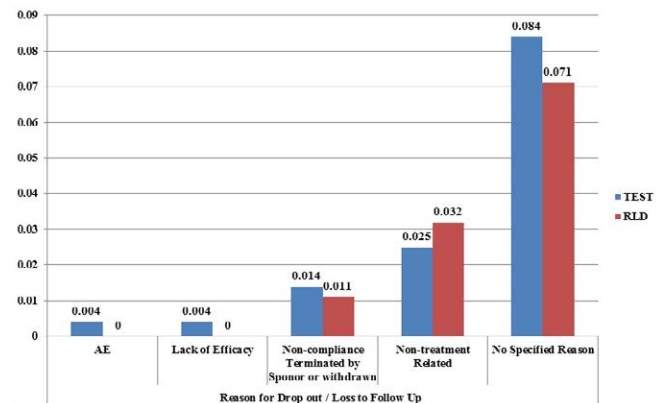In order to evaluate whether dropout (and non-compliance) was

Figure 2 Breakdown of the dropout and non-compliance categories proportion among randomized subjects (%).
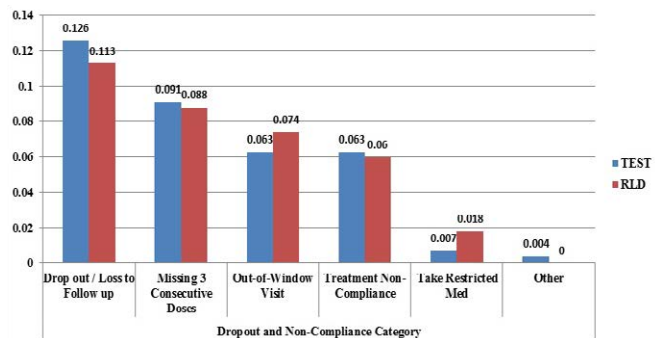

Figure 3 Breakdown of causes for dropout or loss to follow-up and proportion among randomized subjects (%).

at random or not, a logistic regression was conducted to evaluate the association between dropout (or non-compliance, as the outcome) and treatment effect (as the risk factor), i.e., percent change from baseline lesion count (average of inflammatory and non-inflammatory) at the last visit (TOC for completers and last measurement for those who drop out). It turns out that one standard deviation (SD) (27) higher percent change from baseline in lesion count at the last visit was associated with 38.7% (95% CI: 19.5-52.2%) of odds reduction in dropout (odds ratio (OR): 0.61, 95% CI: 0.48, 0.81, P=0.0003), and 35.3% (95% CI: 15-50.9%) of odds reduction in non-compliance (odds ratio (OR): 0.65, 95% CI: 0.49, 0.85, P=0.0025). This means that a worse treatment effect (i.e., lack of efficacy) was associated with a higher risk to drop out and a higher risk to be non-compliance. Therefore, drop out and non-compliance was not at random.

Further analysis shows that dropout and non-compliance were correlated (Pearson Correlation = 0.64, P<0.0001). This means that those who dropped out were more likely to be non-compliant, and those who did not comply with the study protocol were more prone to drop out of the study.

Discussion

Missing data and non-compliance data are very important to clinical endpoint BE studies because they directly affect the

Figure 1 Dropout and non-compliance status in a clinical endpoint is study and proportion among randomized subjects (%).

statistical conclusion of equivalence, which is based on the PP population (completers and compliers in general). However, no case study has been reported to address the status of missing data and non-compliance in clinical endpoint BE studies. The case study in this paper revealed some interesting findings that will help regulators better understand the extent and pattern of drop out and non-compliance in clinical endpoint BE studies. Although this meta-analysis was conducted for clinical endpoint BE studies for treatment of acne vulgaris, similar conclusions can be generalized to clinical endpoint BE studies for other locally acting drugs.

The study result shows that the prevalence of PP population and breakdown of non-compliance and dropout were generally balanced between TEST and RLD. However, as discussed in Sun et al. [1], those who dropped out with no specified reasons and those terminated by the sponsor due to non-compliance or withdrawn can be a mixture of missing completely at random (MCAR), missing at random (MAR), missing not at random (MNAT) [6,7], or nonexistent data, as defined by LaVange and Permutt [8]: "Drug trials, though, often have to cope with observations that are not missing in this narrow sense but never would have been collected because they do not exist". For example, a subject would drop out of the study due to intolerance of the study drug.

Based on the results from the Logistic regression, drop out was not missing completely at random (MCAR) in this study. It was either MAR or MNAR or a combination of MAR and MNAR. However, based on the observed data, we cannot distinguish between MAR and MNAR because the two missing-data mechanisms are not testable unless in the context of certain distributional assumptions [9]. Likewise, non-compliance was not completely at random either. This is concerning because the primary equivalence analysis is usually based on some missing data mechanism, for example, ANOVA is based on MCAR; linear mixed model is based on MAR. The uncertainty of the missing data mechanism requires sensitivity analysis to verify the robustness of the conclusion of the primary equivalence analysis. On the other hand, the fact that drop out and non-compliance were associated with treatment effect means that drop out and non-compliance contained the outcome information. Therefore, naively excluding those who drop out or non-compliers from the PP population may create bias when evaluating equivalence. We may need to re-evaluate the study population and estimands needed for equivalence, which will be a subject of our future research.

Ciccone et al. reported a feasibility study called Project Leonardo, which aims to evaluate the impact of a disease and care management model and of the introduction of "care manager" nurses, trained in this specialized role, into the primary health care system. Results show that Leonardo was feasible and highly effective in increasing patient health knowledge, self-management skills, and readiness to make changes in health behaviors. The "care manager" model may be a good option to prevent missing data and non-compliance in clinical studies.

Findings from this case study confirmed the results of Sun et al. meta-analysis [10]. The case study highlights the impact of missing data and non-compliance data on statistical conclusions of equivalence. It can help in choosing appropriate analysis population, endpoints and estimands, as well as appropriate primary statistical methods and sensitivity methods for evaluating equivalence in clinical endpoint BE studies in presence of missing data and non-compliance data.

Conclusions

Based on this case study of a clinical endpoint bioequivalence study for treatment of acne vulgaris, we find that:

1) Among all of the randomized subjects, 80.1% of the TEST subjects and 77.7% of the RLD subjects were included in the PP population and used in the primary equivalence analysis between TEST and RLD;19.9% of the subjects were non-compliant and/or dropped out (due to reasons other than lack of efficacy) hence are excluded from the primary equivalence analysis.

2) Among the non-PP population, more than half (12.6% out of 19.9% for TEST vs. 11.3% out of 22.3% RLD) were dropouts or lost to follow up; Of the latter, most not specify their drop out reasons (TEST: 8.4%; RLD: 7.1%).

3) Dropout and non-compliance were not at random. They were significantly associated with the treatment effect, i.e., those who dropped out or did not comply were more likely to have poor treatment effect (P<0.05). Furthermore, dropout and non-compliance were correlated with each other (P<0.0001) – those who dropped out were more likely to be non-compliant and vice versa.

Acknowledgement

References

1 Heyting A, Tolboom JTBM, Essers JGA (1992) Statistical handling of drop-outs in longitudinal clinical trials. Stat Med 11: 2043-2061.

2 Sun W, Zhou L, Grosser S, Kim C (2016) A meta-analysis of missing data and non-compliance data in clinical endpoint bioequivalence studies. Stat Biopharm Res 8: 334-344.

3 Davit BM, Nwakama PE, Buehler GJ, Conner DP, Haidar SH, et al. (2009) Comparing generic and innovator drugs: A review of 12 years of bioequivalence data from the United States Food and Drug Administration. Ann Pharmacother 43: 1583-1597.

4 Schuirmann D (1987) A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability. J Pharmacokinet Biopharm 15: 657-680.

5 Guidance for Industry (2005) Acne vulgaris: Developing drugs for development.

6 Rubin DB (1976) Inference and missing data. Biometrika 63: 581-592.

7   Little RJ, Rubin DB (2002) Statistical analysis with missing data. (2nd edn) John Wiley & sons, NY, USA.

8   LaVange LM, Permutt T (2015) A regulatory perspective on missing data in the aftermath of the NRC report. Stat Med 35: 2853-2864.

9   Jaeger M (2006) On testing the missing at random assumption. Machine Learning: ECML 2006. Springer, Berlin, Heidelberg.

10 Ciccone MM, Aquilino A, Cortese F, Scicchitano P, Sassara M, et al. (2010) Feasibility and effectiveness of a disease and care management model in the primary health care system for patients with heart failure and diabetes (Project Leonardo). Vasc Health Risk Manag 6: 297.