Statistical Considerations of Adhesion Data Analysis in NDA

Yu-Ting Weng[1], Caroline Strasinger[2], Meiyu Shen[1@],

[1]Division of Biometrics VI, Office of Translational Sciences, Center for Drug Evaluation and Research, Food and Drug Administration
[2]Division of New Drug Products II, Office of New Drug Products, Office of Pharmaceutical Quality, Center for Drug Evaluation and Research, Food and Drug Administration
[@]Contact author: Meiyu Shen
Division of Biometrics VI, Office of Biostatistics,
CDER, FDA
10903 New Hampshire, Silver Spring, MD 20903
Tel: 301-796-0995
Email: Meiyu.shen@fda.hhs.gov

*This article reflects the views of the author and should not be construed to represent FDA's views or policies.

Abstract
The Draft Guidance for Industry: Assessing Adhesion with Transdermal Delivery Systems and Topical Patches for ANDAs was issued by FDA in 2016. In the guidance, the primary endpoint is the mean adhesion score based on a 5-point numerical scaled score that corresponds to a specified range of adhered surface area of a transdermal or topical system. For the primary endpoint analysis, the non-inferiority test is used to demonstrate adequate adhesion performance of the test product compared to the reference product. However, currently no guidance is available for evaluating the adhesion performance of new drugs in New Drug Applications (NDAs) which generally do not have a reference product. Thus, in this paper for the primary endpoint, we compare the use of an ordinal-scaled score and a continuous scaled percent in an adhesion evaluation. For the primary endpoint analysis, a test-product-only approach based on the superiority test is explored to evaluate the adhesion performance of a transdermal or topical system without a reference product (e.g. new drugs of an NDA) and how to determine a margin is provided for the discussion.

Key Words: Adhesion, New Drug Application, Non-Inferiority Test, Superiority Test.

1. Introduction

Transdermal Delivery Systems (TDS) are discrete dosage forms that deliver Active Pharmaceutical Ingredients (APIs) across the skin to the systemic circulation (USP 36 <1151>, 2009). Two advantages for TDS (Banerjee et al. 2014) include the following: first, they are non-invasive and can easily terminate the therapy, and second, they improve bioavailability of some molecules by avoiding first pass metabolism. Of course, TDS have some limitations, including limited number of drug molecules suitable for passive permeation through the skin and potential for skin irritation.

Currently, TDS can be evaluated by following the approaches recommended in Draft Guidance for Industry: Assessing Adhesion with Transdermal Delivery Systems and Topical Patches for Abbreviated New Drug Applications (ANDAs) (2016). In general, ANDAs refer to a Reference-Listed Drug (RLD), and with respect to the adhesion assessment of a TDS, the proposed ANDA product (test) is compared to the adhesion performance of the RLD (reference). The primary endpoint defined in the draft guidance is the mean adhesion score derived for a TDS or a topical patch from individual adhesion scores at each assessment time point, averaged across all the equally spaced time points (except the baseline or time0). The evaluation approach for the primary endpoint uses the following hypothesis

$$H_0: \mu_T - \mu_R > \delta' \text{ vs. } H_a: \mu_T - \mu_R \leq \delta'$$

where μ_T and μ_R are mean adhesion scores for test product and reference product, respectively and the NI margin (δ') is 0.15. The evaluating criterion for the primary endpoint is that the test product is statistically non-inferior to the reference product if 95% one-sided Upper Confidence Limit (UCL) for the unweighted mean difference between the two products is less than or equal to 0.15.

In contrast, New Drug Applications (NDAs) often have no reference product, and thus NDAs, with respect to adhesion, assess the adhesion performance of a test product only. Because of this difference, the non-inferiority evaluation approach in the ANDAs' draft guidance is not applicable. Additionally, no corresponding guidance is available for NDAs. Thus, in this paper, we explore a test-product-only approach to evaluate the adhesion performance of TDS.

We briefly highlight the adhesion score system currently used for adhesion assessment and then discuss the statistical consideration of the score system as it relates to a non-inferiority approach. Next, we compare the use of an ordinal-scaled score and a continuous-scaled percent in an adhesion evaluation as part of a test-product-only hypothesis and provide an evaluation of simulation results utilizing the alternative test-product-only approach.

2. Adhesion Score System and Statistical Considerations

First, we describe the adhesion score system currently used for adhesion assessment and then discuss relevant statistical consideration of the score system.

2.1 Adhesion Score System

ANDAs' draft guidance recommends that for each assessment, applicants use a 5-point ordinal scale in which each score corresponds to a specified range of adhered surface area of the TDS as shown in Table 1:

Table 1: Recommended Adhesion score system in Draft Guidance for Industry: Assessing Adhesion with Transdermal Delivery Systems and Topical Patches in ANDAs

Score	Adhesion
0	≥ 90% adhered (essentially no lift off the skin)
1	≥ 75% to < 90% adhered (some edges only lifting off the skin)
2	≥ 50% to < 75% adhered (less than half of the TDS lifting off the skin)
3	> 0% to < 50% adhered (not detached, but more than half of the TDS lifting off the skin without falling off)
4	0% adhered (TDS detached; completely off the skin)

2.2 Statistical Considerations of Adhesion Score System

The ordinal-scaled score is a categorical variable from 0 to 4, despite that percent is a continuous variable from 0% to 100%. Further, the ordinal-scaled score utilizes an unequally-spaced scale with regard to percent of area attached to skin surface.
The direct consequence of using score instead of percent for comparing a test product to its RLD is that the difference between these two products can be exaggerated or overlooked. The difference can be exaggerated because if the percentages of the test product and its RLD are close to the score boundary, two percentages can be classified into different score groups, such as 91% and 89% are classified as scores of 0 and 1, respectively. The difference can be overlooked because each score corresponds to a specified range of adhered surface area of the TDS, such as 75% and 89% are considered equivalent scores of 1. In more detail, Table 2 shows the number of times for a specific percentage was recorded in 180 total observations at Reference and Test periods from a simulated dataset (15 subjects with six time points of measurement at Reference and Test periods).

Table 2: Number of times for a specific percentage in 180 total observations at Reference and Test periods in a simulated dataset

Treatment	Percent, (0-100%)	Score, (0-4)	Number of times in 180 total obs[1]
Reference	89	1	12
	76	1	16
Test	89	1	21
	91	0	29

1: 15 subjects and each subject has 6 observations (obs) at Reference and Test periods

3. Approach Exploration

First, we describe the test-product-only hypothesis and then explore an approach to identify the pre-specified margin in the hypothesis.

3.1 Model and Hypothesis

Because NDAs usually do not have an RLD, a test-product-only hypothesis is explored below for the continuous variable: percent. The hypothesis is as follows.

$$H_0: \mu_T \leq \mu_0 \text{ vs. } H_a: \mu_T > \mu_0$$

where μ_T is the mean adhesion percent for test product and μ_0 denotes the pre-specified percent-scale margin.
The test product is statistically superior to the pre-specified percent-scale margin μ_0 if 95% one-sided lower confidence limit for the mean of percent of the test product (μ_T) is greater than the pre-specified percent-scale margin μ_0.

3.2 Support Vector Machine

A Support Vector Machine (SVM) identifies pre-specified percent-scale margin μ_0. SVM proposed by Cortes and Vapnik (1995) finds the test-product-only margin to classify two groups ("pass" and "fail"). Figure 1 illustrates the concept of SVM. Support vectors, which are the data points close to the border of the group, are used to construct the optimal hyperplanes. The optimal hyperplanes are defined as linear decision function with optimal margin between the support vectors of the two groups and the optimal hyperplane here is the test-product-only margin.

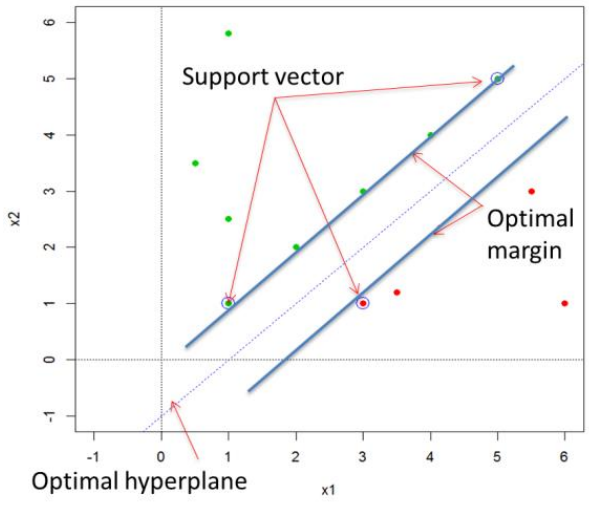

Figure 1: Concept of SVM

4. Simulation

First, we describe the simulation setup and then show the corresponding results.

4.1 Simulation Setup

Historical data from 46 ANDAs or NDAs was collected and consolidated into a simulation database. All of the studies utilized an ordinal-scale approach, and only three of 46 studies utilized a percent-scale dataset. Thus, the interim goal is to explore a test-product-only hypothesis for the ordinal variable: score. The hypothesis is as follows.

$$H_0: \mu'_T \geq \mu'_0 \text{ vs. } H_a: \mu'_T < \mu'_0$$

where μ'_T is the mean adhesion score for test product and μ'_0 denotes the pre-specified score-scale margin. The test product is statistically inferior to the pre-specified score-scale margin μ'_0 if 95% one-sided UCL for the mean of score of the test product (μ'_T) is smaller than the pre-specified score-scale margin μ'_0. Then, we extrapolate from the score-scale margin μ'_0 to a percent-scale margin μ_0 by the assumption of linearity.
After data cleaning and pre-processing, for each study, 95% UCL of mean adhesion score was calculated. In addition, 34 studies were classified in the "pass" group, and 12 studies were classified in the "fail" group. Three studies, which have percent-scale data, were placed in the test dataset, and the remaining 43 studies were placed in the training dataset. SVM with ten-fold cross-validation was conducted.

4.2 Simulation Results

The result shows that the cross-validation error rate is 0%, indicating SVM can be a reasonable approach. The pre-specified score-scale margin was calculated by using all 43 remaining studies, resulting in a margin of 0.721. This score-scale margin was applied to the test dataset, and subsequently all three studies labelled "pass" can be classified in the "pass" group. The score-scale margin was extrapolated to a percent-scale margin by the assumption of linearity, resulting in a margin of 80%. This percent-scale margin was applied to the test dataset, and achieved a consistent result showed that all three studies labelled "pass" can still be classified in the "pass" group.

5. Discussion

Percent can be a measurement of adequate adhesion performance; however, the data collection is still at an investigational stage, and additional data must be collected because of the following three reasons. First, from the results, we know that a big gap exists between the "pass" group and the "fail" group. The margin can be easily shifted if more data are included. Second, only three historical studies utilized a percent-scale data. Thus, extrapolation of the score-scale margin can be obtained by the assumption of linearity; a different percent-scale margin can be obtained if a different assumption was used. Third, SVM appears to provide a reasonable margin because the cross validation error rate is 0%. However, we know that a big gap exists between the "pass" group and the "fail" group. We might use a different approach to achieve 0% cross validation error rate. Although both margins provide consistent results in the analysis, the three historical studies with a percent-scale data are not close to both margins. Thus, the feasibility of the

approach is untested for products that perform close to the pre-specified margin of adhesion.

Acknowledgements

The authors thank our colleagues Drs. Yi Tsong and Wanjie Sun for their informative inputs.

References

General Chapter <1151> Pharmaceutical Dosage Forms (2009)

Banerjee S., Chattopadhyay P., Ghosh A., Datta P., Veer V. (2014). Aspect of adhesives in transdermal drug delivery systems. International Journal of Adhesion & Adhesives 50: 70-84.

Draft Guidance for Industry: Assessing Adhesion with Transdermal Delivery Systems and Topical Patches in ANDAs (2016)

Cortes C., Vapnik V. (1995). Support-vector networks. Machine Learning. 20 (3): 273–29