Comparison of Dissolution Similarity Assessment Methods for Products with Large Variations: \( f_2 \) Statistics and Model-Independent Multivariate Confidence Region Procedure for Dissolution Profiles of Multiple Oral Products

Hiroyuki Yoshida, Hiroko Shibata, Ken-ichi Izutsu, and Yukihiro Goda

\( ^a \) Division of Drugs, National Institute of Health Sciences; 1–18–1 Kamiyoga, Setagaya-ku, Tokyo 158–8501, Japan and \( ^b \) Division of Biological Chemistry and Biologicals, National Institute of Health Sciences; 1–18–1 Kamiyoga, Setagaya-ku, Tokyo 158–8501, Japan.

Received November 21, 2016; accepted February 21, 2017

The current Japanese Ministry of Health Labour and Welfare (MHLW)'s Guideline for Bioequivalence Studies of Generic Products uses averaged dissolution rates for the assessment of dissolution similarity between test and reference formulations. This study clarifies how the application of model-independent multivariate confidence region procedure (Method B), described in the European Medical Agency and U.S. Food and Drug Administration guidelines, affects similarity outcomes obtained empirically from dissolution profiles with large variations in individual dissolution rates. Sixty-one datasets of dissolution profiles for immediate release, oral generic, and corresponding innovator products that showed large variation in individual dissolution rates in generic products were assessed on their similarity by using the \( f_2 \) statistics defined in the MHLW guidelines (MHLW \( f_2 \) method) and two different Method B procedures, including a bootstrap method applied with \( f_2 \) statistics (BS method) and a multivariate analysis method using the Mahalanobis distance (MV method). The MHLW \( f_2 \) and BS methods provided similar dissolution similarities between reference and generic products. Although a small difference in the similarity assessment may be due to the decrease in the lower confidence interval for expected \( f_2 \) values derived from the large variation in individual dissolution rates, the MV method provided results different from those obtained through MHLW \( f_2 \) and BS methods. Analysis of actual dissolution data for products with large individual variations would provide valuable information towards an enhanced understanding of these methods and their possible incorporation in the MHLW guidelines.

Key words  dissolution test; generic product; dissolution similarity; \( f_2 \) statistic; bootstrap analysis; multivariate ANOVA

Ensuring similarity in the performance of oral solid dosage forms, such as between pre- and post-changed approved products or innovator and generic products, requires appropriate methods in comparing their dissolution profiles. The European Medicines Agency (EMA), the U.S. Food and Drug Administration (FDA), and the Ministry of Health Labour and Welfare (MHLW, Japan) guidelines on bioequivalence list the model-independent method using similarity factors (Method A), difference factor \((f_1)\), or a similarity factor \((f_2)\), as main methods for comparing dissolution profiles. The FDA and EMA guidelines also allow the use of the model-independent multivariate confidence region procedure (Method B) for the comparison of profiles that show large variations in individual dissolution rates, without detailed description of the procedure. However, the current MHLW guideline does not describe the use of Method B.

Several procedures have been reported to assess the similarity of dissolution profiles having large variations. A “bootstrap” method (BS method), which compares dissolution profiles by calculating the statistical distribution of \( f_2 \) values, simulates their distribution of expected values and calculates the 90% confidence interval.\(^5\) The method compares the lower limit of this interval of expected \( f_2 \) value \((f_2^*)\) with the \( f_2 \) criterion to assess dissolution similarity. Dissolution profiles with large variations can also be compared through another multivariate analysis procedure (MV method), which uses the Mahalanobis distance \((D_M)\) for this purpose.\(^6\) The \( D_M \) values are based on the pooled variances at each time point and correlations among these,\(^7\) while \( f_2 \) statistics are based on the Euclid distance, established on the differences of averaged dissolution rates. Limited studies describe the effects of different methods of analysis on the similarity outcome using a small number of dissolution profiles and/or simulated ones.\(^8\)–\(^10\) However, it is unclear whether the data used for evaluations represent the dissolution profiles of many formulations on the market.

Consequently, this study compares the methods used to assess the similarity of dissolution profiles with large variations in individual dissolution rates by analyzing dissolution profiles of innovator and generic products marketed in Japan. Sixty-one datasets of dissolution profiles with large variations in individual dissolution rates of generic products (Fig. 1) out of 360 datasets were assessed on their similarity through the \( f_2 \) statistics defined in the MHLW guidelines (MHLW \( f_2 \) method) and two Method B procedures, that is, the BS and MV methods. The analysis of actual dissolution data of products provided valuable information for an enhanced understanding of the utilized methods and their possible incorporation in the MHLW guidelines.

* To whom correspondence should be addressed.  
  e-mail: h.yoshida@nih.go.jp

© 2017 The Pharmaceutical Society of Japan
MATERIALS AND METHODS

Data  Datasets of dissolution profiles of generic and corresponding reference innovator products (six tablets each), obtained for quality evaluation activities of products on the market by the Formulation Evaluation Working Group in Expert Committee on Quality of Generic Pharmaceutical Products in Japan were used for this study.11,12) The working group is composed of 11 public health science institutes. Among 360 datasets of dissolution profiles (90 generic formulations, 4 medium, 10 active pharmaceutical ingredients) reported in the 13th Committee, 61 datasets of dissolution profiles assessed using the $f_2$ statistics as per the MHLW guidelines3) and showed large variations in the individual dissolution rates of generic products (coefficients of variation were above 20% at 5 or 10 min, or above 10% at 15 min or later) were used for the comparison (Fig. 1).

$f_2$ Statistics According to MHLW Guidelines (MHLW $f_2$ Method)  The comparison of dissolution profiles using the MHLW $f_2$ method was based on MHLW guidelines,3) which specify testing times as follows: 2 h in a medium of pH 1.2 and 6 h in other test media, or the time at which the average dissolution of the reference product reaches 85%. The dissolution rates were obtained at three or four time points, chosen according to the average dissolution of reference products. The acceptance criteria for dissolution similarity were set based on the average dissolution of reference products, namely, $f_2 \geq 42$ (the average dissolution of reference product reaches 85% within the specified testing time), $f_2 \geq 46$ (the average dissolution of reference product is between 50 and 85% at the specified testing time), and $f_2 \geq 53$ (the average dissolution of reference product is below 50% at the specified testing time).

Method B: BS and MV Methods  The BS method was applied for the assessment of dissolution similarity using “PhEq_bootstrap version 1.2,” an open source software programmed to assess dissolution similarity by FDA’s and EMA’s $f_2$ computation with the bootstrap technique.9,13) All data points until average dissolutions of both reference and generic products that reached 85% were used according to FDA recommendations. The datasets of dissolution profiles were created by random sampling of individual dissolution rates at each time point in the original data. The number of artificially created dissolution profiles was set to 5000, which was sufficient to stabilize the results. The confidence interval was set to 90%. Dissolution similarity was concluded when the lower confidence interval for the expected $f_2$ value ($f_2^*$) calculated by the software was not below the similarity acceptance criteria. The same acceptance criteria as for the MHLW $f_2$ method were applied for similarity assessment.

The MV method was applied as previously described6) using Microsoft Excel. Global similarity limits were set based on the average dissolutions of reference products to ±15, ±12, and ±9% (the average dissolution of reference product reaches 85%, between 50 and 85%, and below 50% at the specified time points, or less than 20% at specified testing time).
testing time, respectively) according to MHLW guidelines. The similarity limits were calculated using the global similarity limit, subsequently compared to the upper 90% limit of the confidence interval of true $D_M$ ($D_{Mu}^*$). Dissolution similarity was concluded when the $D_{Mu}^*$ value was not above the similarity limit. The time points used for the analysis of Method B are not specified in FDA guidelines. Therefore, all data points or those until the reference product reached 85% dissolution were used, according to MHLW guidelines.

RESULTS AND DISCUSSION

Among the 219 dissolution datasets of reference/generic products required to be assessed using MHLW $f_2$ methods, the coefficient of variation of 61 datasets (28%) exceeded the criteria for the use of Method A in the FDA and EMA guidelines (Fig. 1). The relatively high ratio of formulations with large variations in individual dissolution data indicates the necessity for setting appropriate methods (e.g., Method B) to evaluate dissolution similarities. Figure 2 shows the number of reference and generic product dissolution profile datasets that showed similar or non-similar dissolution profiles when assessed using the MHLW $f_2$ method and the two types of Method B. Among the 61 datasets of dissolution profiles of reference/generic products, analysis through the MHLW $f_2$, BS, and MV methods categorized 40, 38, and 24 combinations as “similar,” respectively. Moreover, the MHLW $f_2$ and BS methods showed almost similar conclusions (Fig. 2a). The sensitivity of the BS method to the MHLW $f_2$ method was higher ($36/40 = 0.9$) than that of the MV method ($18/40 = 0.45$), which agreed with the results of the previous studies. The large variations in individual dissolution rates expanded the confidence intervals of the expected $f_2$ value, resulting in a decrease in the $f_2^*$ value. Although most $f_2^*$ values obtained by the BS method were lower than the $f_2$ values obtained by the MHLW $f_2$ method (Fig. 3), the difference in the time points used in the analysis could explain the higher $f_2^*$ values obtained through the BS method compared to the $f_2$ values of the MHLW $f_2$ method. Therefore, adopting the BS method in the MHLW guidelines would require specifying appropriate measurement time points for calculating $f_2$ values.

Fig. 3. Relationships between $f_2$ Value Obtained by MHLW $f_2$ Method and Lower Confidence Interval of Expected $f_2$ ($f_2^*$) Value by BS Method in 61 Datasets of Dissolution Profiles

A solid line represents equal values.

Fig. 4. Example of Dissolution Datasets of a Generic and Corresponding Reference Product That Showed Different Conclusion of Dissolution Similarity between the MHLW $f_2$ Method, BS Method, and MV Method

(a) Dissolution data categorized as similar by using the MHLW $f_2$ or BS methods and as non-similar by the MV method. (b) Dissolution data categorized as similar by using the MV method and as non-similar by the MHLW $f_2$ or BS methods. Each result represents the mean±S.D. of six formulations.

**Table 1. Similarity Assessment of Dissolution Profiles Shown in Fig. 4a by 3 Different Methods**

<table>
<thead>
<tr>
<th>Method</th>
<th>$f_2$</th>
<th>Similarity criteria</th>
<th>Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>MHLW $f_2$ method</td>
<td>64.13</td>
<td>$\geq 42$</td>
<td>Similar</td>
</tr>
<tr>
<td>BS method</td>
<td>51.14</td>
<td>$\geq 42$</td>
<td>Similar</td>
</tr>
<tr>
<td>MV method</td>
<td>$D_{Mu}=9.61$, $D_{Mu}^*$=Limit</td>
<td>$D_{Mu}^*$=Limit</td>
<td>Non-similar</td>
</tr>
</tbody>
</table>

**Table 2. Similarity Assessment of Dissolution Profiles Shown in Fig. 4b by 3 Different Methods**

<table>
<thead>
<tr>
<th>Method</th>
<th>$f_2$</th>
<th>Similarity criteria</th>
<th>Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>MHLW $f_2$ method</td>
<td>40.26</td>
<td>$\geq 42$</td>
<td>Non-similar</td>
</tr>
<tr>
<td>BS method</td>
<td>28.96</td>
<td>$\geq 42$</td>
<td>Non-similar</td>
</tr>
<tr>
<td>MV method</td>
<td>$D_{Mu}=9.49$, $D_{Mu}^*$=Limit</td>
<td>$D_{Mu}^*$=Limit</td>
<td>Similar</td>
</tr>
</tbody>
</table>
Analysis through the MV method categorized a smaller number of datasets as similar dissolution profiles compared to those analyzed by the MHLW $f_2$ and BS methods. Around half of the datasets categorized as similar dissolution profiles by using the MHLW $f_2$ or BS methods were concluded to be non-similar in the MV method (Figs. 2b, c). This was different from the comparison between the BS and MHLW $f_2$ methods (Fig. 2a). Some datasets were categorized as non-similar, although the differences of mean dissolution rates were within 15% at all time points (Fig. 4a, Table 1). In case of two dissolution curves crossing each other, analysis through the MV method was reported to be less sensitive to that through the $f_2$ method, requiring special attention on using the MV method in this case. Conversely, six dissolution datasets, categorized as non-similar on using the MHLW $f_2$ method, were concluded to be similar through the MV method. Among these, two datasets showed nearly 100% dissolution of test products at an early time point and reduced the variance-covariance matrix, resulting in a tendency to meet the similarity criteria of the MV method. An example of a dissolution profile showing different similarities depending on the analysis method is shown in Fig. 4b and Table 2.

The number of tablets used in the analysis was associated with the second degree of freedom in the F-distribution. A relatively small number of tablets (six tablets) and the limited time points (up to nine) used for the analysis in this study would explain the different conclusions on dissolution similarity through the MV method compared to MHLW $f_2$.

CONCLUSION

This study, using actual dissolution profiles with large individual variations, indicated a large difference in the dissolution similarity assessed by the current MHLW $f_2$ method and two Method B procedures (BS and MV methods). The use of datasets on actual dissolution profiles is as important as using simulated datasets for a better understanding of the advantages and drawbacks of the analysis methods. However, these results may vary depending on the criteria applied in each method. Adopting these assessment methods in the MHLW guidelines would require deliberate consideration, including specifying a detailed procedure, and should be based on an appropriate rationale.

Acknowledgments We are grateful to Dr. Yomota of the Pharmaceutical and Medical Device Regulatory Science Society of Japan (PMRJ), Mr. Takahashi of Astellas Pharma Inc., and members of the study group in the Office of Generic Drugs, Pharmaceuticals and Medical Devices Agency (PMDA) for the valuable discussions regarding analysis by Method B. This work is partly supported by the Research on Regulatory Science of Pharmaceuticals and Medical Devices from the Japan Agency for Medical Research and Development (AMED), and by a Health and Labour Sciences Research Grant from the Ministry of Health, Labour and Welfare of Japan.

Conflict of Interest The authors declare no conflict of interest.

REFERENCES