Biologicals xxx (2017) 1–5

Contents lists available at ScienceDirect

Biologicals

journal homepage: www.elsevier.com/locate/biologicals

Determination of acceptance criteria and sample sizes for accelerated stability comparability studies for biologics

Binbing Yu*, Lingmin Zeng, Harry Yang

MedImmune LLC, Gaithersburg, MD 20878, United States

ARTICLE INFO

Article history:
Received 20 April 2017
Received in revised form
21 June 2017
Accepted 4 July 2017
Available online xxx

Keywords:
Accelerated stability study
Acceptance criteria
Process comparability
Sample size

ABSTRACT

Changes of manufacturing processes are common. It is required by the regulatory agencies that manufacturers establish adequate and appropriate comparability between pre-change and post-change products. The goals of comparability assessments are to demonstrate the comparability and consistency of product quality before and after change and to demonstrate that the changes do not have an adverse effect on safety and efficacy of the drug products. Accelerated or stressed stability studies may shed light on drug quality under stressed environmental conditions and on product differences in the degradation pathways. Comparability of accelerated stability data may provide further evidence on the impact of process change. Equivalence test has been recommended to demonstrate the comparability of stability profiles for accelerated stability studies. Selection of appropriate acceptance criteria for determining comparability is one of the most challenging steps in the comparability studies. Because of the inherent heterogeneity of biologics, the stability profiles may vary considerably from batch to batch. It is more challenging to set the acceptance criteria for comparing the accelerated stability data for biologics. In this article, we present an approach for determining the acceptance criteria and necessary sample sizes for accelerated comparability studies for biologics.

1. Introduction

As a drug product progresses from early phase development, regulatory approval to commercialization, changes in formulation, upstream and downstream processes, scale, delivery devices, manufacture facility are common [1]. These changes are often intended to improve efficiency, to achieve better process control, to meet product supply demands etc [2]. According to Code of Federal Regulation Title 21 (21 CFR 601.12), "An application may submit one or more protocols describing the specific tests and validation studies and acceptable limits" to demonstrate that the process change has no adverse impact on the identity, strength, quality, purity or potency of the product [3,4].

As such, a process change should be adequately assessed by comparing pre- and post-change products and demonstrate that the old and new processes are comparable. The regulatory agencies, including ICH, EMA and US FDA, have issued multiple guidance documents on the principles of demonstrating comparability [5–10]. The goal of a comparability study is to ascertain whether any quality attributes have been impacted by the process change. Because biological products are highly complex and process-dependent, any process change is expected to affect the quality attributes. Comparability does not mean the products are identical, but that their physicochemical and biological properties are sufficiently similar to ensure no adverse impact on quality, safety and efficacy [11]. Regulatory agencies strongly recommend "one or more protocols describing the specific tests and studies and acceptance criteria to be achieved" before carrying out the comparability exercises [3]. Selection of relevant analytical methods and determination of acceptance criteria for comparability may be the most challenging step in a comparability study [12]. Encouraged by the regulatory agencies [7,10], statistical methods have been used to establish comparability acceptance criteria [13] and to evaluate comparability [14] before and after process changes. The determination of acceptance criteria for comparability of analytical methods and key quality attributes has been discussed by Chatfield and Borman [13] and de Fontenay [15].

Although the long-term stability study is the gold standard for evaluating the comparability of stability, accelerated stability studies offer a quick assessment of stability profiles in the presence of a process change. The accelerated stability study is especially advantageous when a product is stable at the long-term storage

* Corresponding author.
E-mail address: yub@medimmune.com (B. Yu).

http://dx.doi.org/10.1016/j.biologicals.2017.07.001

condition. Therefore, evaluating the effect of storage time and environmental factors by accelerated and long-term stability studies is a critical part of the comparability evaluation [11]. Cowley et al. [2] proposed to use the equivalence test to demonstrate the comparability of stability profiles and Burdick and Sidor [16] established the acceptance criteria based on the variance of degradation rates from the pre-change historical data. They considered the common-slope models, where variations in degradation rates (slopes) from different lots are only attributable to analytical method variability. Burdick and Sidor [16] briefly mentioned the random-slope model for balanced stability design, but did not offer details on the general design, where multiple lots may have different followup times and imbalanced stability time points. Because of complexity of most biologics, the lot-to-lot variability in degradation rates cannot be ignored.

In this article, we extended the acceptance criteria calculation to comparing stability data from multiple lots with heterogeneous degradation rates for biologics. We used a simulation-based method to determine the necessary sample sizes for the equivalence test for stability comparison. The proposed method was illustrated by an accelerated stability study for an experimental monoclonal antibody.

2. Determination of acceptance criteria

Usually the acceptance criteria are determined based on representative historical stability data from the old process. Linear mixed-effects model is a popular model for stability data from multiple lots [17–19]. Suppose that the stability data from n historical lots are available. For the ith lot, let x_{ij} be the stability testing time points and let y_{ij} be the observed quality attribute values, $j = 1, .., m_i$. The linear mixed-effects model assumes that

$$y_{ij} = \alpha_i + \beta_i x_{ij} + \varepsilon_{ij}, j = 1, \ldots, m_i, i = 1, \ldots, n. \quad (1)$$

where $\varepsilon_{ij} \sim N(0, \sigma_\varepsilon^2)$ are iid normal random variables for assay variability. To account for the lot-to-lot variability, the intercepts and slopes are assumed to be random effects, where

$$\alpha_i = \alpha + a_i, a_i \sim N\left(0, \sigma_a^2\right),$$

$$\beta_i = \beta + b_i, b_i \sim N\left(0, \sigma_b^2\right),$$

where α, β are the overall intercept and slope and $(a_i, b_i), i = 1, \ldots, n$ are random intercepts and slopes. The parameters to be estimated include $\alpha, \beta, \sigma_a^2, \sigma_b^2$ and σ_ε^2.

For comparability of accelerated stability data, an acceptance criterion Δ represents the largest acceptable difference between the average slopes of the historical and new processes. Burdick and Sidor [16] defined the acceptance criterion as

$$\Delta = ES \times \sqrt{\mathrm{Var}\left(\widehat{\beta}_i\right)}, \quad (2)$$

where ES is the effect size and $\widehat{\beta}_i$ is the slope estimate for the ith randomly sampled lot from the historical process. Burdick and Sidor [16] assumed fixed slope with $\sigma_b^2 = 0$ and discussed in details on how to select the acceptable effect size. They emphasized that a one-size-fits-all criterion for ES may not be appropriate, the ES should be determined based on scientific expectations for a particular application. Some common values are $ES = 2$ or 3, but in some cases the 95% upper bound on σ_b from the historical stability studies can be used.

Under the linear-mixed model formulation, the estimate $\widehat{\beta}_i$ can be interpreted as the random slope of stability data from a future lot and

$$\widehat{\beta}_i = \widehat{\beta} + b_i, \ \ \widehat{\beta} \sim N\left(\beta, \sigma_{\widehat{\beta}}^2\right), \ \ b_i \sim N\left(0, \sigma_b^2\right),$$

where β is the overall degradation slope, $\sigma_{\widehat{\beta}}^2$ is the variance of the overall slope estimate $\widehat{\beta}$ and σ_b^2 is the variance of the random effects b_i. When $\widehat{\beta}$ and b_i are independent, the slope estimate $\widehat{\beta}_i \sim N(\beta, \sigma_{\widehat{\beta}}^2 + \sigma_b^2)$ with $\mathrm{Var}(\widehat{\beta}_i) = \sigma_{\widehat{\beta}}^2 + \sigma_b^2$. For the balanced stability design where all lots have the same stability time points, the variance of $\widehat{\beta}_i$ reduces to equation (8) in Ref. [16].

The parameter estimates of the linear mixed-effects model can be obtained by SAS Proc mixed or R function lmer. An alternative estimation method is the Bayesian Markov chain Monte Carlo (MCMC) method. The MCMC method has been implemented in software packages WinBUGS or OpenBUGS. The Bayesian formulation requires specifications of prior of the parameters. One can use the following priors, $\alpha \sim N(0, \sigma_\alpha^2), \beta \sim N(0, \sigma_\beta^2), \sigma_\varepsilon \sim \mathrm{Uniform}(L_\varepsilon, U_\varepsilon), \sigma_a \sim \mathrm{Uniform}(L_a, U_a), \sigma_b \sim \mathrm{Uniform}(L_b, U_b)$, where $\mathrm{Uniform}(L, U)$ is the uniform distribution with range (L, U). The hyperparameters $\sigma_\alpha^2, \sigma_\beta^2, L_\varepsilon, U_\varepsilon, L_a, U_a, L_b, U_b$ are chosen accordingly depending on whether priors are informative. One advantage of using the Bayesian method is that the scientific knowledge or expert opinion accumulated during the previous pharmaceutical development can be incorporated as informative priors. For example, the analytical method variability σ_ε can be derived as the intermediate precision estimates from the analytical method validation study or robustness study. For a typical bioassay for measuring relative potency, the standard deviation of intermediate precision has a range of (5%, 30%).

3. Equivalence test and sample size determination

When stability study is deemed necessary, a side-by-side study plan comparing material from the new process with that from the old process should be part of the comparability protocol [12]. The goal of the equivalence test is demonstrate with high confidence the average degradation rates from two processes do not differ by more than a pre-defined acceptance margin Δ. The null hypothesis of the equivalence test is such that the difference between the mean degradation rates (slopes) for the accelerated stability data from the two processes exceeds Δ:

$$H_0 : |\beta_H - \beta_N| \geq \Delta \text{ versus } H_1 : |\beta_H - \beta_N| < \Delta, \quad (3)$$

where β_H is the slope for the accelerated stability data from the historical process, and β_N is for the new process. The equivalence of stability profiles from the two processes is achieved if the 90% confidence limits for the difference $\beta_H - \beta_N$ fall inside the equivalency region $(-\Delta, \Delta)$. This test is also called the two one-sided test (TOST) [20].

Let $\boldsymbol{\theta} = (\alpha, \beta, \sigma_a^2, \sigma_b^2, \sigma_\varepsilon^2)$ be the random parameters generated from the posterior distributions. We assume that the mean slope for the historical stability data is $\beta_H = \beta$ and there is a shift of δ for the mean slope for the new stability data, i.e., $\beta_N = \beta + \delta$. Without additional information, we assume that all other parameters are the same for both processes and the number of lots from the old and new processes $n_1 = n_2 = n$. Following ICH Guidance [21], the stability time points are 0, 1, 3, 6, 9 and 12 months.

The power of the equivalence test is related to the number of

Please cite this article in press as: Yu B, et al., Determination of acceptance criteria and sample sizes for accelerated stability comparability studies for biologics, Biologicals (2017), http://dx.doi.org/10.1016/j.biologicals.2017.07.001