Journal of Immunological Methods 442 (2017) 29–34

Contents lists available at ScienceDirect

Journal of Immunological Methods

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

Research paper

Immunogenicity assay cut point determination using nonparametric tolerance limit

Jianchun Zhang [a], Wenjia Li [b], Lorin K. Roskos [a], Harry Yang [a,*]

[a] MedImmune LLC, Gaithersburg, MD, United States
[b] University of Maryland, College Park, MD, United States

ARTICLE INFO

Article history:
Received 22 August 2016
Received in revised form 13 December 2016
Accepted 3 January 2017
Available online 5 January 2017

Keywords:
Bootstrap
Cut point
Immunogenicity
Nonparametric method
Random effect model
Tolerance interval

ABSTRACT

The newly released FDA guidance on immunogenicity assay development and validation recommends use of a lower confidence limit of the percentile of the negative subject population as the cut point in order to guarantee a pre-specified false positive rate with high confidence. The limit is, in essence, a lower tolerance limit. Although in literature several methods are available for determining the tolerance limit, they either fail to take into account the repeated measurement of the data from a typical immunogenicity assay quantification/validation experiment or rely heavily on normality assumption of the data, which is rarely correct. As a result, the methods may result in biased estimates of the cut point, causing the false positive rate to be either lower or higher than expected. To overcome this drawback, we propose two non-parametric methods under repeated measure data structure and without normal distribution assumption. Simulation studies were carried to compare the performance of the two non-parametric approaches with the current methods. The results of the simulation studies show that one of the two nonparametric methods outperforms all the other methods and provides a satisfactory coverage probability even with moderate sample sizes. In addition, it is simple and straightforward to implement. Therefore, it is a preferred method for immunogenicity assay cut point determination.

1. Introduction

Biotechnology-derived therapeutics including monoclonal antibodies, proteins, and peptides hold great promise for treating various diseases such as cancer and inflammatory diseases. However, biological drug products can lead to immunogenic responses, resulting in the formation of anti-drug antibodies (ADAs) because of their large molecule size, complex structure, and complicated manufacturing process. These antibody responses have the potential to negatively affect product safety and efficacy. In humans, ADAs of some biological therapeutics have been shown to cause a variety of clinical consequences ranging from relatively mild hypersensitivity to serious adverse events. Therefore, mitigation of immunogenicity risk is of great interest to patients, clinicians, manufacturers, and regulatory agencies (Yang et al., 2015). However, since immunogenicity is a complex phenomenon, owing to myriad factor, immunogenicity risk assessment and control can be challenging.

In order to assess the immunogenic potential of biological therapeutics, it is important to develop reliable laboratory test methods (immunogenicity assays) that provide valid assessments of anti-drug antibody (ADA) responses in both nonclinical and clinical studies. Validated assays for the detection and characterization of ADAs are routinely used in clinical trials to monitor patient immunogenic responses. Unlike other bioassays, the most challenging aspect of immunogenicity assay development and validation is the lack of reference standards. For this reason, regulatory guidelines such as those from the U.S. Food and Drug Administration (FDA) (FDA, 2009) and the U.S. Pharmacopeial Convention (USP) (USP, 2014) and industry white paper (Shankar et al., 2008) suggest a multi-tiered approach to differentiating ADA-negative samples from ADA-positive samples by means of statistically derived cut points. In the screening assay, sample responses are compared against the screening cut point, which is determined from negative samples. Samples with responses below the screening cut point are declared to be negative and excluded from further testing; whereas samples with responses at or above the screening cut point are defined as potentially positive and subjected to additional testing in the confirmatory assay. The FDA guidance (FDA, 2009) recommends that 5% false positive rate be preserved for the screening assay. As a result, it is suggested that the 95th percentile of the negative subject population be used as the screening cut point. Because the population 95th percentile is unknown, it is often estimated using the 95th sample percentile calculated from either a parametric or nonparametric model. Detailed description on how to estimate the screening cut point can be found in the white paper by Shankar et al. (2008) and the book by Yang et al. (2015).

Recognizing that the sample percentile is a point estimate, which can be either bigger or smaller than the population percentile because

* Corresponding author.
E-mail address: yangh@medimmune.com (H. Yang).

http://dx.doi.org/10.1016/j.jim.2017.01.001
0022-1759/

of sample variability, recently FDA (2016) recommends use of a 90% (80%) one-sided lower confidence limit for the 95th (99th) percentile of the negative subject population as the screening (confirmatory) cut point. This interval estimator of the unknown population percentile is, in essence, the lower tolerance limit with a content of 5% (1%) and a confidence level of 90% (80%) for the negative population. It is intended to assure that the cut point will give rise to a false positive rate of at least 5% with high confidence (Shen et al., 2015) in the case of screening assay. Two parametric methods for determining such tolerance limit are investigated by Shen et al. (2015) based on the assumption that the test results are independently and identically distributed (IID) according to a normal distribution. Separately, method of nonparametric tolerance limit is available (Krishnamoorthy and Mathew, 2009; Young and Mathew, 2014). However, in a typical immunogenicity assay quantification/validation experiment, data consist of results of samples which are repeatedly measured. As a result, sample responses are not IID as repeated measurements of samples from the same subject are correlated. As shown in this paper, methods without taking data structure into account have a higher chance to over-estimate the cut point. As a matter of fact, the tolerance limit can be constructed under random effect models. Such methods are readily available in published literature under normal distribution; see, for example, Hoffman (2010), Vangel (1992), Krishnamoorthy and Mathew (2004). More recently, Hoffman and Berger (2011) considered the "confidence-level" cut point based on random effect model under normal distribution.

All of the above methods assume that distribution of immunogenicity data is normal. However, in reality, this normality assumption seldom holds due to heterogeneous nature of ADA responses of different subjects. Therefore, it is desirable to develop cut point determination methods that are insensitive to departure from the normality assumption. Zhang et al. (2015) considered cut point estimation using random effect models with skew-t and log-gamma distribution. Schaarschmidt et al. (2015) explored cut point estimation using mixture models in the presence of pre-existing antibodies. But all these methods are focused on point estimation. In keeping with the recent regulatory requirements on cut point analysis, it is necessary to develop interval-based statistical procedures for cut point estimation.

It is also worth noting that in general it is often difficult to specify a distribution that well characterizes the negative subject population in random effect models. For this reason, nonparametric methods are appealing as they do not rely on distributional assumptions. As performance of nonparametric methods depends largely on sample size, nonparametric methods suggested for cut point analysis would only be useful if they perform well based on data from a typical immunogenicity assay qualification/validation experiment.

Under one-way repeated measure data structure, Olsson and Rootzén (1996) suggested two nonparametric quantile estimators for balanced and unbalanced data following any continuous distributions. The two methods are shown to have good asymptotic properties under large sample sizes. In this article, based on Olsson and Rootzén's results, we propose two nonparametric procedures for cut point determination. As shown later, one of the two methods outperforms all the other methods discussed above. Moreover, the method is computationally simple and can be implemented easily. The statistical robustness of computational ease makes the method a preferred choice for immunogenicity cut point determination.

The paper is organized as follows. Section 2 describes current available methods and the two non-parametric methods for calculating the lower tolerance limit as cut point. In Section 3, performance of the two non-parametric methods characterized by their coverage probability is studied and compared with methods by Shen et al. (2015), Hoffman and Berger (2011) and Young and Mathew (2014), under various known distributions through simulations. Section 4 discusses the discrete behaviour of coverage probability of the non-parametric method and performance of a bootstrap alternative, which is widely used for constructing confidence intervals. Section 5 concludes the article with a recommendation of a non-parametric method for determining the lower tolerance limit of the cut point.

2. Method

2.1. Model

Consider the balanced one-way random effect model, where it is assumed that there are I subjects with each subjects being tested J times with a validated immunogenicity assay, possibly by different analysts on different days. Let

$$Y_{ij} = \mu + \alpha_i + \varepsilon_{ij},\ \ i = 1,\ldots,I; j = 1,\ldots,J \tag{1}$$

where Y_{ij}s are the (normalized) sample responses, which are identically distributed with a continuous cumulative distribution function $F(x)$. The random effect α_i represents the ith subject's true effect and ε_{ij} represents the error at jth measurement for the ith subject. It is further assumed that data from different subjects are independent. Under the assumption of the random effect model (1), where the subject effects α's are IID normal with mean 0 and variance σ_α^2; and the error terms ε_{ij}'s are IID normal with mean 0 and variance ${\sigma_\varepsilon}^2$, the immunogenicity assay responses Y_{ij} are normally distributed with mean μ and variance ${\sigma_T}^2 = \sigma_\alpha^2 + {\sigma_\varepsilon}^2$.

2.2. Current methods

2.2.1. Method 1

Shen et al. (2015) investigated two tolerance limit approaches for screening cut point estimation. One is the exact lower confidence limit of a normal percentile, which was originally proposed by Chakraborti and Li (2012) and the other is an approximate lower confidence limit of a normal percentile. The former is recommended by Shen et al. (2015) as it performs better than the latter, as evidenced by the fact that its coverage probability is much closer to the nominal level. We denote this method as Normal1.

Normal1 provides a lower $(1-\alpha)$ confidence limit for the p-quantile as follows:

$$\hat{\theta}_{\text{Normal1}} \equiv \overline{Y} + Cz_pS - bS\sqrt{1 + nz_p^2\left(C^2 - 1\right)}/\sqrt{n} \tag{2}$$

where $n = I*J$, $\overline{Y}$ is the sample mean; and S is sample standard deviation. C is a bias-correcting factor given in Shen et al. (2015) and Chakraborti and Li (2012). z_p is the lower pth quantile of standard normal distribution. b is a solution of $F_T(b) = 1 - \alpha$ (see Shen et al., 2015 for detail).

One apparent shortcoming of the above method is that its performance or coverage probability is influenced by the assumption that the assay results Y_{ij} are independently identically distributed (IID) according to a normal distribution. In a typical immunogenicity assay quantification/validation study, the data usually follow random effect models such as the model in (1). It can be readily verified that observations from the same subject/patient are correlated, with a correlation coefficient given by $\rho \equiv \sigma_\alpha^2/(\sigma_\alpha^2 + \sigma_\varepsilon^2)$. When IID assumption no longer holds, the sample standard deviation S in (2) is given by

$$S = \sqrt{\frac{1}{n-1}\sum_i\sum_j \left(Y_{ij} - \overline{Y}\right)^2}$$

The grand sample mean $\overline{Y}$ is an unbiased estimator of μ while the sample variance S^2 under-estimates the total variance σ_T^2. To see this, note that S^2 can be written as $\frac{1}{n-1}\left(\sum_i\sum_j Y_{ij}^2 - n\overline{Y}^2\right)$. It is straightforward