9.4.3 Data Analysis and Interpretation of Trial Results

Two methods are typically used for the statistical analysis of data from equivalence trials, and both methods can be applied to biosimilars. The most commonly used method is based on the indirect confidence interval comparison that requires specification of the equivalence limits (Chow and Liu, 2004; FDA, 2010). Equivalence is demonstrated when the confidence interval for the selected metric of the treatment effect falls entirely within the lower and upper equivalence limits. If a *p*-value approach is used, then the *p*-values should be computed based on the two-one sided *t*-tests (TOST) procedure testing simultaneously the null hypotheses of inferiority and superiority (Schuirmann, 1987). In using the TOST procedure, equivalence is demonstrated when the *p*-values obtained are less than the significance level used. An error that is commonly made is concluding equivalence based on a statistically nonsignificant *p*-value from testing the null hypothesis of no difference. In general, concluding equivalence based on observing a statistically nonsignificant *p*-value from testing the null hypothesis of no difference between the biosimilar and the RBD is incorrect (ICH E9, 1998). This observation is especially applicable to secondary endpoints for which comparability margins are not typically predefined. In such cases, it is erroneous to conclude similarity on the secondary endpoints based on statistically nonsignificant *p*-values from the test of a difference between the biosimilar and the RBD.

The second method that is sometimes used for analyzing equivalence trial data is the virtual comparison method that synthesizes the estimate of the treatment effect relative to the control in the current equivalence trial with the estimate of the treatment effect of the control relative to placebo from historical trial (Chow and Liu, 2004; FDA, 2010). The resulting statistics are treated as if they were derived from the same trial. This method is unique as it does not require the prespecification of equivalence limits. The method can also be used to provide an estimate of the fraction of the effect size of the active control that is retained by the test drug product, and this can be useful in the overall interpretation of the study results. To date, the biosimilars that have been approved by Health Canada used the indirect confidence interval comparison method because there appears to be more familiarity with this approach, in general, relative to the synthesis method.

With regard to analysis sets, analysis based on the intention to treat (ITT) approach in an equivalence trial might bias the results toward equivalence (ICH E9, 1998). The per-protocol (PP) approach, which excludes subjects who either do not meet study eligibility criteria or have major protocol violations, can also lead to biased results. Bias is especially a concern when a large proportion of subjects are excluded from analysis. As a result, equivalence trials should be analyzed using ITT and per-protocol approaches, and both approaches should support equivalence (Snapinn, 2000). In using each approach, a careful assessment of the potential for bias should be performed. Health Canada's approach to this issue has been to assess the results from analyses based on both approaches in order to ensure that both sets of results support the conclusion of equivalence of the biosimilar to the RBD.

Finally, proper interpretation of the results from an equivalence trial is critical. Overall, the lack of a placebo concurrent control creates challenges in the interpretation of the results of any equivalence trial, and this is no exception for biosimilar