In the context of a global development program, it has been necessary for sponsors to include the reference product from different regional sources (EU and US), as well as testing the biosimilar product formulation–container combination that is intended for commercialization.

12.22 THERAPEUTIC EQUIVALENCE STUDIES

The primary evidence for the relative immunogenicity of the EU-approved biosimilar products has been obtained from therapeutic equivalence studies, rather than from the comparative PK studies discussed in the preceding subsection (Chamberlain, 2014). Current EU regulatory guidance (EMA, 2014c) recommends that therapeutic equivalence studies be performed for products for which there are no surrogate markers for efficacy. The study population should be sensitive for detecting the clinical impact of potential differences between the biosimilar and reference products, as well as representing an approved therapeutic indication of the reference product.

As stated above, FDA guidance indicates that a separate clinical study from the comparative PK/PD (Phase 1) study may be required to evaluate the immunogenicity of the biosimilar candidate in direct comparison to the reference product. Nevertheless, the ADA response should be measured in all clinical studies.

The protocol for therapeutic equivalence studies should specify a descriptive data analysis of ADA formation; this could include a description of the incidence and median titer of ADAs, the proportion of subjects with neutralizing antibodies (ideally, subclassified by low/medium/high titer), and a correlation of ADA positive versus ADA negative status with drug trough concentration, efficacy, and frequency of adverse events. To check for systematic bias, the proportion of false positives (positive in screening assay but negative in confirmatory assay) in each treatment group should be compared. If the residual drug concentration exceeds the validated drug tolerance limit in some subjects, ADA results should also be evaluated separately for the subpopulations with drug concentration below or above the drug tolerance level.

Sample time-points should be selected to indicate the dynamics of antibody formation in relation to treatment outcome, taking into account the need to minimize the potential for interference in the ADA assay by residual circulating drug. The ADA sampling schedule used for the comparative study of Remsima versus Remicade in ankylosing spondylitis is summarized in Figure 12.2. A similar ADA sample schedule was used in the therapeutic equivalence study, CT-P13 1.3, performed in rheumatoid arthritis patients (EPAR for Remsima).

This sampling schedule enabled a demonstration of a highly similar profile of neutralizing ADA incidence and magnitude across the two treatment groups in both clinical studies (EPAR for Remsima), as illustrated in Figure 12.3a and b.

A biosimilar version of etanercept, SB4 (Benepali), was reported to have a lower incidence of detectable ADA than for the reference product (Enbrel) during the initial 24-week treatment period of a Phase 3 comparative efficacy study (Emery et al., 2015). CHMP questioned whether insufficient drug tolerance of the ADA assay, allied to higher drug levels in some subjects, may have contributed to a higher number of false negatives in the biosimilar treatment arm at the 8-week time-point (EPAR for Benepali). Nevertheless, because no difference in incidence of ADA signals detected after the