Because of the pivotal nature of the pharmacokinetic study, the study is usually performed with clinical material made at the final manufacturing scale so that regulatory authorities can be assured that additional manufacturing changes will not be needed that could alter product characteristics or clinical responses. When performing pharmacokinetic (PK) studies to determine if a biosimilar is bioequivalent to the reference product, the least variability is encountered with normal healthy volunteers. However, with some biological drugs with a relatively long half-life or prolonged pharmacodynamic effects, it is not possible to use normal healthy volunteers. For example, normal volunteers have been used in PK studies of filgrastim, erythropoietin, anti-TNF biologics, and even bevacizumab (Sorgel et al., 2009; Elmeliegy et al., 2015; Sorgel et al., 2015). However, it is not possible to perform PK studies in normal volunteers with rituximab (Maloney et al., 1997) due to the prolonged depletion of B lymphocytes that would make these studies unethical. Therefore, the initial PK study of biosimilarity may be fairly straightforward with some biologics (except for pegfilgrastim as mentioned above), but for others like rituximab, PK studies have to be performed in patients. This adds variability and difficulty standardizing the multiple blood sample draws required for dense PK profiling in routine clinical centers. Limiting patient variability to have a completely homogeneous population is antithetical in the development of novel biologics. For novel biologics the evaluation of PK while assessing the impact of patient factors on the exposure and use of the product is actively sought.

Traditional approaches to statistical justification of bioequivalence margins hold true for biosimilar PK studies just as with all other molecules (FDA, 2013). Regulatory authorities generally agree that using a 90% confidence interval is acceptable. Thus, it is possible to reasonably harmonize biosimilar development between regulatory regions. This is not the case for efficacy equivalence designs discussed below. Some regulators consider it more sensitive to pick up differences between a reference product and a proposed biosimilar when lower doses than those used in routine clinical treatment are tested. However, when conducting PK/PD studies in patients, it is difficult to use lower doses than are indicated for that disease state, as this would be ethically inappropriate. When using normal healthy volunteers, it is possible to use lower doses than specified on the label for therapeutic use. As mentioned above, dosing on the steep part of the dose response curve may be more sensitive to pick up differences between the biosimilar and the reference product if differences exist. Once analytical or PK differences are detected, whether statistically significant or not, these differences have to be evaluated to determine if these differences would be clinically meaningful in practice.

7.4.3 Confirmatory Efficacy Study

The typical nomenclature for a pivotal trial for a novel drug uses the "phase III" terminology. However, in the development of biosimilars, the PK/PD study is also pivotal (as mentioned above), while the confirmatory efficacy study addresses residual uncertainty about the biological function of the proposed biosimilar compared to the reference product. Therefore, biosimilar sponsors often simply refer to these studies as confirmatory safety and efficacy studies and avoid the phase III moniker. Even