Clinical trials of biosimilars should become more similar

Jonathan Kay,1 John D Isaacs2

Since the approval of the infliximab biosimilar, CT-P13, by the Korean Ministry of Food & Drug Safety (MFDS) on 23 July 2012,1 biosimilars to treat inflammatory diseases have become available to patients in many countries around the world.2 The European Commission (EC) approved CT-P13 on 10 September 2013,3 and, subsequently, it has been granted marketing authorisation by regulatory agencies in many other countries, including the USA. CT-P13 is now marketed in >70 countries worldwide as Remsima, Inflectra and Flammegis.4

Several other biosimilars of tumour necrosis factor (TNF) inhibitors have also received regulatory approval and are commercially available. HD203 was the first etanercept biosimilar granted marketing authorisation by the Korean MFDS on 11 November 2014 and was marketed as Davictrel in South Korea.5 However, because the facility where it was manufactured was sold, its licence could not be retained and it was withdrawn from the market upon the request of its manufacturer on 30 September 2015. Another etanercept biosimilar, SB4, which received approval by the Korean MFDS on 8 September 2015 and by the EC on 14 January 2016, is marketed as Brenzys in South Korea and as Benepali in the European Union (EU) and in the European Economic Area (EEA) member states of Norway, Iceland and Liechtenstein.6 SB2, another infliximab biosimilar, was granted marketing authorisation by the Korean MFDS on 4 December 2015 and by the EC on 26 May 2016, and is sold as Remslexis in South Korea and as Flixabi in the EU and EEA member states.7,8 In this issue, Bae et al8 report the results of the phase III clinical trial of HD203 (etanercept biosimilar), Emery and colleagues report the results of the phase III clinical trial of SB4 (etanercept biosimilar)9 and Choe et al10 report the results of the phase III clinical trial of SB2 (infliximab biosimilar).

The European Medicines Agency (EMA) was the first to establish a pathway for the regulatory approval of biosimilars in 2005, in which a biosimilar is compared with its reference product or originator biological, hereafter referred to as bio-originator.11 This occurs in a series of analytical, in vitro, in vivo, pharmacokinetic, pharmacodynamic and clinical studies according to a stepwise approach. Other regulatory agencies have followed suit and established similar regulatory pathways. In 2012, the US Food and Drug Administration (FDA) issued draft guidance regarding its pathway for the review and approval of biosimilars and articulated a ‘totality-of-the-evidence’ approach to evaluating the data generated by all of these studies.12

The purpose of a clinical trial comparing a biosimilar with its bio-originator is to reduce residual uncertainty following extensive analytical, in vitro and pharmacokinetic analyses. Efficacy of the bio-originator already has been proven in the pivotal clinical trials that were conducted to gain regulatory approval and by subsequent experience in clinical practice. Thus, if equivalence of the biosimilar to its bio-originator can be demonstrated, there is no need to re-establish its clinical benefit. The clinical trial of a biosimilar therefore can be viewed as a bioassay to demonstrate that it exhibits a clinical effect comparable to that of the bio-originator in patients with a disease for which the bio-originator is approved. Similar principles permit subsequent regulatory extrapolation to other licensed indications of the bio-originator, provided that therapeutic efficacy relies on a similar mechanism of action in the extrapolated indications.

Phase III clinical trials comparing biosimilar TNF inhibitors with their bio-originators have employed different designs. Although their primary end points were similar, the phase III clinical trial of CT-P13 evaluated efficacy only at 12, 24 and 54 weeks,8 each of which is a time point during the plateau phase of the time–response curve. In contrast, the studies of SB2 and SB4 also evaluated efficacy at several earlier time points.9,10 Since potential differences in efficacy are more likely to be detected during the rapid rise phase of the time–response curve compared with the plateau phase, assessment of efficacy at early time points is a more sensitive way of comparing a biosimilar with its bio-originator.11 This aspect of clinical trial design should be standardised for future studies of biosimilars. Indeed, it could be argued that a ‘standard’ clinical trial design be adopted for all biosimilars of a particular bio-originator in a given disease.

To demonstrate two-sided therapeutic equivalence of a biosimilar to its bio-originator in a clinical trial, the 95% CI for the mean absolute difference in the primary end point between the biosimilar and the bio-originator must fall within a predefined equivalence margin (δ).12 This equivalence margin is often derived from a meta-analysis of the therapeutic effect of the bio-originator in the original placebo-controlled clinical trials, calculated as the risk difference in the end point of interest between active drug (a) and placebo (p), often referred to as the ‘delta’ (δa-p). Whereas the EMA suggests use of 95% CI,13 the US FDA prefers use of the narrower 90% CI for demonstration of therapeutic equivalence.14 A one-sided equivalence (non-inferiority) study would be inadequate to demonstrate biosimilarity since it cannot exclude the possibility that the test treatment might be a ‘bio-better’.

In order to preserve a proportion (1−ε) of the therapeutic effect of the bio-originator, δ should be a relatively small fraction (ε) of the difference between bio-originator and placebo in the meta-analysis of placebo-controlled trials (δa-p).16 For example, the study comparing HD203 to bio-originator etanercept used an equivalence margin of ±20%,8 whereas that of SB4 employed one of ±15%.3 A meta-analysis of the pivotal placebo-controlled clinical trials of bio-originator etanercept estimated δa-p = 0.4049 (an absolute risk difference of 40.49%) in the proportion of American College of Rheumatology (ACR)20 responders between bio-originator etanercept and placebo.17 Thus, an equivalence margin of ±20% preserves 50% of this therapeutic effect (1−ε, where ε = ±δa-p = 0.5), whereas an equivalence margin of ±15% preserves 62.5% of this therapeutic effect (1−ε, where
state, type of assay, assay sensitivity and interference by free drug. Assays for ADAs must also avoid interference by rheumatoid factor and heterophile antibody. In randomised controlled trials comparing a biosimilar with its bio-originator, differences in ADA incidence are likely to accurately reflect relative immunogenicity when ADAs are measured in an identical manner and if all other end points (particularly pharmacokinetic) demonstrate equivalence. In addition to their incidence, the titre and specificity of ADAs are important. Differences in immunogenicity may reflect, for example, subtle differences in glycosylation or other post-translational modifications, aggregates, impurities, and formulation and packaging effects. Although greater immunogenicity of a potential biosimilar compared with its bio-originator would preclude biosimilarity, regulatory agencies will accept a biosimilar with lower immunogenicity. In such a circumstance, it is essential to ascertain whether ADAs influence pharmacokinetics, adverse events or efficacy by comparing outcomes at relevant time points in both ADA-positive and ADA-negative patients.

In the clinical trial of SB4, ADAs were observed in only 0.7% of subjects receiving the biosimilar compared with 13.1% of subjects who received the bio-originator. The additional ADAs detected to bio-originator etanercept were transient, of low titre and detected mostly at early time points between weeks 4 and 8. There was no difference between biosimilar and bio-originator in either safety or efficacy among ADA-positive and ADA-negative patients, suggesting that ADAs did not interfere with clinical activity. Indeed, only one patient receiving the bio-originator developed neutralising antibodies. Importantly, the higher incidence of ISRs with bio-originator etanercept was not explained by ADAs.

The SB4 study used a highly sensitive electrochemiluminescence bridging assay to detect ADAs, perhaps explaining the higher incidence of ADAs than that historically detected by conventional ELISAs or radioimmunoassays in previous clinical trials of bio-originator etanercept. Nonetheless, the assay used to measure ADAs in the SB4 study had a low tolerance to free drug, and, at week 8, a pharmacokinetic sub-study suggested a greater area under the concentration-time curve between doses for SB4 than for bio-originator etanercept, although this was possibly explained by higher inter-subject variability. Furthermore, potential drug-ADA complexes were acid-dissociated, thereby improving free drug tolerance. Perhaps critically, assays of ADAs were performed more frequently and at earlier time points than in previous studies—it was at those earlier time points that differences in ADA incidence were detected. Notably, a lower incidence of ADAs was also demonstrated in a phase I pharmacokinetic study of SB4 compared with bio-originator etanercept in healthy subjects. The EMA presented its overall findings in an EPAR and, regarding immunogenicity, concluded: “Possible explanations for the differences in ADA incidence between BENEPALI [SB4] and ENBREL [bio-originator etanercept] could be the slightly different drug concentrations in samples or differences in the sensitivities of the corresponding analytical methods … Therefore as the observed differences with respect to ADA formation … appeared to be transient, with almost no differences after 8 weeks of treatment, their clinical significance was considered minimal.”

An ELISA was used to study immunogenicity in the HD203 study, wherein eight patients receiving the biosimilar developed ADAs (three of which had neutralising ADAs) and three patients receiving bio-originator etanercept developed ADAs (one of which had neutralising ADAs). These numbers are too low to identify any potential clinical consequences of ADAs. However, in the SB2 study, in which 55.1% of patients treated with SB2 and 49.7% of patients treated with bio-originator infliximab developed ADAs, ADA-positive patients were more likely to experience infusion or hypersensitivity reactions and less likely to achieve the primary outcome of an ACR20 response. Nonetheless, SB2 and bio-originator infliximab remained equivalent in analyses stratified for the presence or absence of ADAs. Since equivalence trials of a particular biopharmaceutical always compare a biosimilar with the same bio-originator, they should be based on the same meta-analysis of double-blind, randomised-controlled clinical trial data comparing the bio-originator with placebo. Thus, it would be appropriate to standardise clinical trial design as much as possible. This would simplify making indirect comparisons among novel biosimilars, in terms of pharmacokinetics, efficacy and potential adverse events. Box 1 lists the various aspects of clinical trial design that could be standardised in this way.

The availability of biosimilars has reduced the cost of treating an individual patient since biosimilars typically are marketed at a lower price than their bio-originators. This is especially important in developing markets where access to biopharmaceuticals is restricted by cost: access to a lower-priced biosimilar might allow a patient to receive a treatment that previously was difficult to obtain or was
unavailable. In some countries, such as South Korea, competition introduced by a biosimilar has driven down the price of the bio-originator.26 As many more biosimilars are in development, we recommend that clinical trial design be standardised. This standardisation could be agreed upon and overseen by regulatory agencies around the world. The introduction of consistency across clinical trials should increase confidence in these more affordable biopharmaceuticals, both within the healthcare community and among patients.

Handling editor Tore K Kvien

Contributors Both authors meet the criteria for authorship: JK and JDJ conceived of and wrote this editorial and thereby take responsibility for this work. Both authors agree to publish this work.

Competing interests JK has provided expert advice to Amgen; AbbVie; Boehringer Ingelheim GmbH; Bristol-Myers Squibb Company; Eli Lilly and Company; Epirus Biopharmaceuticals; Genentech; Hospira; Janssen Biotech; Merck Sharp & Dohme; Novartis Pharmaceuticals Corporation; Pfizer; Samsung Bioepis; Sandoz; Roche Laboratories; and UCB and has received research grant support paid to the University of Massachusetts Medical School from AbbVie; Pfizer; Genentech; Roche Laboratories; and UCB. JDJ has received honoraria, research grants or consulting fees from AbbVie; Baxalta; Biogen; Boehringer Ingelheim GmbH; Bristol-Myers-Squibb Company; Celltrion; Healthcare; Chugai; Pharmaeutica; Eli Lilly and Company; Epirus Biopharmaceuticals; Hospira; Janssen Pharmaceuticals; Merck Serono S.A.; Pfizer; and Roche Laboratories.

Provenance and peer review Commissioned; externally peer reviewed.

REFERENCES


