Clinical trials of biosimilars should become more similar

Jonathan Kay,[1] John D Isaacs[2]

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However, because the facility where it was manufactured was sold, its licence could not be retained and it was withdrawn from the market upon the request of its manufacturer on 30 September 2015.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.[5 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 Renflexis in South Korea and as Flixabi in the EU and EEA member states.[5 7] In this issue, Bae *et al*[8] 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 al*[10] 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 14, 30 and 54 weeks,[13 14] and that of HD203 evaluated efficacy only at 12, 24 and 48 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.[15] 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).[16] 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' (δ_{ap}). Whereas the EMA suggests use of 95% CI,[17] the US FDA prefers use of the narrower 90% CI for demonstration of therapeutic equivalence.[18] 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 (δ_{ap}).[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%.[9] A meta-analysis of the pivotal placebo-controlled clinical trials of bio-originator etanercept estimated $\delta_{ap}=0.4049$ (an absolute risk difference of 40.49% in the proportion of American College of Rheumatology (ACR)20 responders between bio-originator etanercept and placebo).[19] Thus, an equivalence margin of ±20% preserves 50% of this therapeutic effect (1−ε, where $\varepsilon=\delta/\delta_{ap}=0.5$), whereas an equivalence margin of ±15% preserves 62.5% of this therapeutic effect (1−ε, where

[1]Division of Rheumatology, Department of Medicine, UMass Memorial Medical Center and University of Massachusetts Medical School, Worcester, Massachusetts, USA; [2]Musculoskeletal Research Group, Institute of Cellular Medicine, Newcastle University and National Institute of Health Research Newcastle Biomedical Research Centre, Newcastle upon Tyne, UK

Correspondence to Professor Jonathan Kay, Division of Rheumatology, Department of Medicine, UMass Memorial Medical Center and University of Massachusetts Medical School, Memorial Campus, 119 Belmont Street, Worcester, MA 01605, USA; jonathan.kay@umassmemorial.org

$\varepsilon = \delta/\delta_{ap} = 0.375$). Both equivalence margins were agreed upon in discussion with the relevant regulatory agencies but, in our opinion, it would be ideal to standardise the equivalence margin used in future clinical trials of all biosimilars of the same bio-originator since δ_{ap} is derived by analysing historical data from the same placebo-controlled trials of the bio-originator. The EMA suggests that the choice of this δ_{ap} should be "supported by evidence of what is considered an unimportant difference in the particular disease area".[17] Interestingly, HD203 demonstrated superiority to bio-originator etanercept when ACR50 responses were compared at weeks 24 and 48.[8] Importantly, however, biosimilar and bio-originator demonstrated equivalence at all other time points, including at the primary end point of ACR20 at week 24. Therefore, this result did not preclude biosimilarity since the study met its primary end point.

In all three studies, minor differences were detected in the occurrence of adverse events. The incidence of injection-site reactions (ISRs) with HD203 and SB4 was lower than with bio-originator etanercept (2.0% vs 5.5% and 3.7% vs 17.2%, respectively), although with slightly different definitions.[8 20] The incidence of alanine aminotransferase elevations with SB2 was higher than that with bio-originator infliximab (7.9% vs 2.7%).[10] However, the EMA European Public Assessment Report (EPAR) for Flixabi concluded that "most of the liver enzyme elevations were transient, and there was no difference in prolonged enzyme elevation between the SB2 and the EU Remicade treatment groups".[21] A biosimilar must have the same primary amino acid sequence as its bio-originator. However, there may be differences in the vehicle and excipients in which the drug substances are formulated; and some of the relevant bio-originator information may remain proprietary. Such differences may account for the varying incidences of these adverse events. For example, a lack of L-arginine in the formulation and of latex in the needle shield may help to explain the reduced incidence of ISRs with SB4 compared with bio-originator. Such differences do not preclude biosimilarity from a regulatory standpoint and are factors that might differentiate between products.

Biopharmaceuticals, both bio-originators and biosimilars, are immunogenic and induce anti-drug antibodies (ADAs) that have the potential to influence therapeutic efficacy. The incidence of ADAs depends upon a number of factors, including disease state, type of assay, assay sensitivity and interference by free drug.[22] 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.[23] Although greater immunogenicity of a potential biosimilar compared with its bio-originator would preclude biosimilarity, regulatory agencies will accept a biosimilar with lower immunogenicity.[24] 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.[9] 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.[20] Nonetheless, the assay used to measure ADAs in the SB4 study had a low tolerance to free drug, and, at week 8, a pharmacokinetic substudy 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.[25] 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".[19]

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).[8] 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.[10] 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

Box 1 Aspects of clinical trial design to assess biosimilarity that could be standardised

- Healthy subjects versus patients (in phase I)
- Inclusion and exclusion criteria
- Equivalence margins
- Primary end point (including timing of assessment)
- Secondary end points (including timing of assessment)
- Pharmacokinetic assays (end points compared and timing of assessment)
- Immunogenicity (assays used and timing of testing)
- Analysis of effects of immunogenicity on pharmacokinetics, efficacy and safety
- Definition of adverse events, for example, injection-site reactions
- Statistical analyses
- Cross-over designs beyond primary end point (in phase III)

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 JDI 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. JDI has received honoraria, research grants or consulting fees from AbbVie; Baxalta; Biogen; Boehringer Ingelheim GmbH; Bristol-Myers-Squibb Company; Celltrion Healthcare; Chugai Pharmaceutica; 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.

To cite Kay J, Isaacs JD. *Ann Rheum Dis* 2017;**76**:4–6.

Received 27 June 2016
Revised 18 July 2016
Accepted 31 July 2016
Published Online First 25 August 2016

Linked

▶ http://dx.doi.org/10.1136/annrheumdis-2015-207588
▶ http://dx.doi.org/10.1136/annrheumdis-2015-207613
▶ http://dx.doi.org/10.1136/annrheumdis-2015-207764

Ann Rheum Dis 2017;**76**:4–6.
doi:10.1136/annrheumdis-2015-208113

REFERENCES

1 GaBI Online—Generics and Biosimilars Initiative. Biosimilar monoclonal antibody approved in Korea. 2012. http://www.gabionline.net/Biosimilars/News/Biosimilar-monoclonal-antibody-approved-in-Korea (accessed 22 Jun 2016).
2 Dörner T, Strand V, Cornes P, *et al*. The changing landscape of biosimilars in rheumatology. *Ann Rheum Dis* 2016;75:974–82.
3 European Medicines Agency. Remsima. 2016. http://www.ema.europa.eu/ema/index.jsp?curl=pages/medicines/human/medicines/002576/human_med_001682.jsp&mid=WC0b01ac058001d124 (accessed 22 Jun 2016).
4 Celltrion Inc. Approved products: Remsima, the world first biosimilar mAb. 2015. http://www.celltrion.com/en/products/license.do (accessed 22 Jun 2016).
5 GaBI Online—Generics and Biosimilars Initiative. Biosimilars approved in South Korea. 2016. http://www.gabionline.net/Biosimilars/General/Biosimilars-approved-in-South-Korea (accessed 22 Jun 2016).
6 European Medicines Agency. Benepali. 2016. http://www.ema.europa.eu/ema/index.jsp?curl=pages/medicines/human/medicines/004007/human_med_001944.jsp&mid=WC0b01ac058001d124 (accessed 22 Jun 2016).
7 European Medicines Agency. Flixabi. 2016. http://www.ema.europa.eu/ema/index.jsp?curl=pages/medicines/human/medicines/004020/human_med_001980.jsp&mid=WC0b01ac058001d124 (accessed 22 Jun 2016).
8 Bae SC, Kim J, Choe JY, *et al*. A phase III, multicentre, randomised, double-blind, active-controlled, parallel-group trial comparing safety and efficacy of HD203, with innovator etanercept, in combination with methotrexate, in patients with rheumatoid arthritis: the HERA study. *Ann Rheum Dis* 2017;76:65–71.
9 Emery P, Vencovsky J, Sylwestrzak A, *et al*. A phase III randomised, double-blind, parallel-group study comparing SB4 with etanercept reference product in patients with active rheumatoid arthritis despite methotrexate therapy. *Ann Rheum Dis* 2017;76:51–7.
10 Choe JY, Prodanovic N, Niebrzydowski J, *et al*. A randomised, double-blind, phase III study comparing SB2, an infliximab biosimilar, to the infliximab reference product Remicade in patients with moderate to severe rheumatoid arthritis despite methotrexate therapy. *Ann Rheum Dis* 2017;76:58–64.
11 Committee for Medicinal Products for Human Use. Guideline on similar biological medicinal products. 2005. http://www.emea.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003517.pdf (accessed 22 Jun 2016).
12 United States Department of Health & Human Services, Food & Drug Administration. Guidance for industry: Scientific considerations in demonstrating biosimilarity to a reference product, draft guidance. 2012. http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM291128.pdf (accessed 22 Jun 2016).
13 Yoo DH, Hrycaj P, Miranda P, *et al*. A randomised, double-blind, parallel-group study to demonstrate equivalence in efficacy and safety of CT-P13 compared with innovator infliximab when coadministered with methotrexate in patients with active rheumatoid arthritis: the PLANETRA study. *Ann Rheum Dis* 2013;72:1613–20.
14 Yoo DH, Racewicz A, Brzezicki J, *et al*. A phase III randomized study to evaluate the efficacy and safety of CT-P13 compared with reference infliximab in patients with active rheumatoid arthritis: 54-week results from the PLANETRA study. *Arthritis Res Ther* 2015;18:82.
15 Kay J, Smolen JS. Biosimilars to treat inflammatory arthritis: the challenge of proving identity. *Ann Rheum Dis* 2013;72:1589–93.
16 Ng T-H. Choice of delta in equivalence testing. *Drug Inf J* 2001;35:1517–27.
17 Committee for Medicinal Products for Human Use. Guideline on the choice of the non-inferiority margin. 2005. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003636.pdf (accessed 18 Jul 2016).
18 United States Department of Health & Human Services, Food & Drug Administration. Guidance for industry: Statistical approaches to establishing bioequivalence. 2001. http://www.fda.gov/downloads/Drugs/…/Guidances/ucm070244.pdf (accessed 18 Jul 2016).
19 Committee for Medicinal Products for Human Use. Assessment report: Benepali. 2015. http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Public_assessment_report/human/004007/WC500200380.pdf (accessed 22 Jun 2016).
20 Emery P, Vencovský J, Kang JW, *et al*. Confirmation on the immunogenicity assay used in the SB4 phase III study: response to the comments by Meacci *et al*. *Ann Rheum Dis* 2016;75:e40.
21 Committee for Medicinal Products for Human Use. CHMP assessment report: Flixabi. 2016. http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Public_assessment_report/human/004020/WC500208358.pdf (accessed 22 Jun 2016).
22 van Schouwenburg PA, Rispens T, Wolbink GJ. Immunogenicity of anti-TNF biologic therapies for rheumatoid arthritis. *Nat Rev Rheumatol* 2013;9:164–72.
23 United States Department of Health & Human Services, Food & Drug Administration. Guidance for industry: Immunogenicity assessment for therapeutic protein products. 2014. http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm338856.pdf (accessed 22 June 2016).
24 Committee for Medicinal Products for Human Use. Guideline on similar biological medicinal products. 2014. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2014/10/WC500176768.pdf (accessed 22 Jun 2016).
25 Lee YJ, Shin D, Kim Y, *et al*. A randomised Phase I pharmacokinetic study comparing SB4 and etanercept reference product (Enbrel®) in healthy subjects. *Br J Clin Pharmacol* 2016;82:64–73.
26 Kim SC, Choi NK, Lee J, *et al*. Brief report: utilization of the first biosimilar infliximab since its approval in South Korea. *Arthritis Rheumatol* 2016;68:1076–9.