REVIEW

Biosimilars in rheumatology: what the clinician should know

Gilberto Castañeda-Hernández,1,2 Rodrigo González-Ramírez,1 Jonathan Kay,3 Morton A Scheinberg4,5

ABSTRACT

Biosimilars are now a reality in rheumatology. Although analytical and non-clinical procedures to establish similarity have evolved significantly, clinical trials demonstrating equivalent efficacy and safety are absolutely required for all biosimilars. The design of such trials, including equivalence and non-inferiority statistical approaches, are discussed. Clinical evidence on biosimilars that have been approved recently or are presently being developed for use in rheumatology is also reviewed and compared with that available for biometrics (or intended copies), which are non-innovator biologies that are marketed in several countries but have not undergone review according to a regulatory pathway for biosimilars.

BACKGROUND

There is now considerable interest in biosimilars among rheumatologists, although the distinction between a true biosimilar and a biomimic (or intended copy) may not be clear to most. Many countries have changed their regulatory requirements to accommodate this new class of medicinal products and to distinguish them from generics.1,2 Currently, only one biosimilar is approved by the European Medicines Agency (EMA) for the treatment of rheumatological diseases: an infliximab biosimilar, which is commercialised as Remsima/Inflectra. Following the lead of the EMA, regulatory agencies in other countries, including South Korea, Canada, Japan, Turkey and Colombia,3 have approved this infliximab biosimilar. However, the approved indications differ among these countries. For example, EMA allowed the results of clinical trials conducted in rheumatological diseases trials to be extrapolated to inflammatory bowel diseases, while Health Canada did not.4 Recently, BOW015, an infliximab biosimilar with the commercial name Infimab,5 and ZRC-3197, an adalimumab biosimilar with the commercial name Exemptia,6 were approved in India, while HD203, an etanercept biosimilar, was approved in South Korea.7 Whether other countries will approve these products with the currently available data, however, remains to be seen.

Currently, numerous biosimilars are in development and it is likely that some of them will be commercialised in the near future.2 Although analytical and non-clinical procedures to establish similarity have evolved significantly, clinical trials demonstrating equivalent efficacy and safety are an absolute requirement for regulatory approval of all biosimilars.8 Understanding the principles by which these trials are designed and analysed will help the clinician to evaluate and use these drugs in practice.
DESIGN OF CLINICAL TRIALS FOR BIOSIMILARITY

Pharmacokinetic studies

Pharmacokinetic comparisons demonstrating equivalence of certain biosimilars with their corresponding innovators have been conducted in healthy volunteers and in patients with rheumatological diseases. Pharmacokinetic equivalence is necessary, but not sufficient, to demonstrate biosimilarity.

Biosimilars are not identical to innovators. Certain molecular differences, albeit minimal, can modify affinity for the target ligand without modifying pharmacokinetics. Hence, innovator and non-innovator products may exhibit differences in clinical efficacy and safety despite comparable pharmacokinetics. This is why, unlike generic small molecule drugs, bioequivalence of a biosimilar with its reference product cannot be established solely on pharmacokinetic grounds.

Small molecule generics are molecularly identical to the innovator small molecule drugs, such that all pharmacodynamic properties of both drugs are exactly the same. The same is not true of biosimilars with innovator biopharmaceuticals. Furthermore, a biosimilar and its reference biopharmaceutical can also differ in terms of immunogenicity. Thus, at least one clinical study documenting equivalent efficacy and safety is absolutely required.

Equivalence studies

The purpose of a clinical trial comparing a biosimilar to its innovator is to demonstrate equivalent efficacy and safety. Therefore, a suitable patient population must be chosen that is sensitive enough for differences in the measured end point, which might be due to dissimilarities between the products assayed, to be detected. The patient groups receiving the innovator and the biosimilar must be balanced in terms of demographics and clinical characteristics. The number of subjects studied should be sufficient to achieve statistical power. Finally, the primary end point and the study duration must be clinically relevant.

Hypothesis tests are designed to demonstrate differences between two samples and not equivalence. In a comparison performed using a hypothesis test, such as the Student's t test or analysis of variance, the absence of a significant difference (i.e., p>0.05) does not demonstrate equivalence. A sample is described by its mean value and its variability, which is frequently expressed as the SD. If two samples overlap, it is difficult to demonstrate a significant difference. Overlapping can be due to a narrow difference between means or to wide SDs. The SD (s) of a sample is calculated by equation 1, where x and \( \bar{X} \) are the individual and mean values of the measured parameter, and n the number of subjects.

\[
S = \sqrt{\frac{\sum (x - X)^2}{n - 1}}
\]  

If the difference between mean values is the same, overlapping depends on the SD values of both samples. This is shown in figure 1, where two comparisons are performed where the difference between mean values is the same in both situations. In figure 1A, the SD values are low. Hence, the samples do not overlap and comparison using a hypothesis test yields a significant difference. In contrast, in figure 1B, the SD values are high and the samples thus overlap. In this circumstance, it is not possible to show a significant difference between the two samples using hypothesis tests.

Hypothesis tests are used to demonstrate superiority of one treatment over another and are not valid to demonstrate equivalence. In equation 1, if n is small, the value of s is larger. Accordingly, it is impossible to show a significant difference between two samples, if the number of patients to be compared is small, because of high-SD values. This, of course, does not imply that the two samples being compared are equivalent. Therefore, clinical studies with small sample sizes are inadequate to demonstrate biosimilarity and can be misleading. To demonstrate that two treatments are equivalent, a clinical trial typically would require a larger sample size than if it were intended to demonstrate superiority of one over the other.

Equivalence analysis is less familiar to most clinicians than is hypothesis testing. However, the principles of equivalence analysis are not difficult to understand. The first consideration is that equivalence does not imply identity, but, rather, means that the results observed with a test product are within a pre-established range in relation to those obtained with the reference product.
The limits of this equivalence range are arbitrary and are established on the basis of past experience. For a candidate biosimilar that is being developed to treat rheumatological diseases, the equivalence range typically is based on a meta-analysis of placebo-controlled clinical trials of the reference product. The mean difference in the outcome measure of interest between the reference product and placebo is calculated, and then divided by an arbitrary factor (usually of 2) to narrow the equivalence range.\(^{13}\) Once this range has been established, the clinical trial comparing the biosimilar to its reference product is conducted, the observed difference between treatments is assessed, and the 95% CIs of this difference are calculated. If these 95% CI are contained within the pre-established equivalence range, the two treatments are equivalent.\(^{13}\)\(^{17}\)

How is equivalent efficacy determined in practice for a biosimilar compared to its reference (or innovator) product? As an example, let us consider the evaluation of an infliximab biosimilar in comparison to innovator infliximab (Remicade)\(^{13}\)\(^{18}\) First, a suitable patient population is chosen (ie, patients with rheumatoid arthritis taking methotrexate). Next, a suitable primary end point and treatment duration are chosen (ie, percentage of patients achieving the ACR20 response at week 30). Finally, an equivalence range is chosen. Since the difference in the ACR20 response between treatment with infliximab and placebo is about 30% in the several placebo-controlled clinical trials of infliximab treatment for rheumatoid arthritis, an equivalence range of half that value (15%) is considered to be reasonable. This means that the ACR20 response to treatment with the biosimilar can vary by \(\pm 15\%\) with respect to that to treatment with the innovator. A logarithmic transformation of the data is then performed. With log-transformed data, the \(\pm 15\%\) equivalence range is 0.85–1.18, so as to be symmetrical.

Now, let A and B represent mean values of an efficacy parameter for the candidate biosimilar and the innovator, respectively. The hypothesis to prove is shown in equation 2:

\[
0.85 \leq \frac{A}{B} \leq 1.18
\]

Once A and B have been determined for the two groups of treated patients studied, A/B is calculated and the corresponding 95% CIs are estimated. If the 95% CI for A/B lies entirely within the 0.85–1.18 range, the two treatments are considered to be equivalent. If this 95% CI exceeds the lower, higher, or both equivalence limits, the treatments are not considered to be. This is shown graphically in figure 2. White bars depict various 95% confidence limits for values of A/B that correspond to equivalence. The mean values can be low or high, and variability can be small or large. However, as long as the 95% CI lies within the pre-established range, equivalence of the two treatments is established. In each of these situations, the non-innovator is considered to be biosimilar to the innovator. Dark bars depict other 95% CIs for values of A/B that do not meet the definition of equivalence. Therefore, the candidate non-innovator biopharmaceutical cannot be considered to be a biosimilar. In each of these cases, the 95% CI goes beyond the lower, higher or both equivalence limits. The absolute magnitude of A/B does not matter when the 95% confidence limit exceeds either the lower or upper limit of the pre-established range. In that situation, equivalence cannot be established and the candidate non-innovator biopharmaceutical tested is considered to not be biosimilar to the reference product.\(^{17}\)

The infliximab biosimilar CT-P13 showed efficacy equivalent to that of innovator infliximab (Remicade) in the PLANETRA study.\(^{18}\) In this study, 606 patients with rheumatoid arthritis, inadequately responsive to methotrexate, were randomised 1:1 to receive either CT-P13 or innovator infliximab. The primary efficacy outcome was the proportion of subjects achieving an ACR20 response at 30 weeks. At this time point, the response rates for the biosimilar and the innovator were 60.9% (184/302) and 58.9% (178/304), respectively, in the intention-to-treat population. The difference between treatments was 2% with a CI 95% of −6% to 10%. Since the entirety of this CI lies within the pre-established range of \(\pm 15\%\) (0.85–1.18 after logarithmic transformation), the efficacy of CT-P13 was considered to be equivalent to that of innovator infliximab. A similar finding was observed in the per protocol analysis, where the response rates for the
biosimilar and the innovator were 73.4% (182/248) and 69.7% (175/251), respectively. The difference between treatments in this analysis was 4%, with a 95% CI of −4% to 12%. Since this 95% CI also lies entirely within the pre-established range of ±15%, the two treatments again were considered to be equivalent.

Non-inferiority approach
In some circumstances, a non-inferiority approach can be suitable to evaluate biosimilarity.17 19 The guidelines on biosimilar evaluation issued by the WHO consider a non-inferiority design to be acceptable.20 Accordingly, several countries will accept such a non-inferiority approach. In Canada, regulatory authorities have acknowledged that an equivalence trial design is preferred. However, if clearly justified, a non-inferiority approach may be acceptable under certain conditions. Evidence must be provided that an eventual superiority has no clinical meaning and that there is no increase in adverse reactions with regard to the reference product.21 In comparative effectiveness clinical trials designed using a non-inferiority approach, the test product can be superior to the reference product but it cannot be inferior. If a biosimilar is to be compared to an innovator using a non-inferiority approach, the hypothesis to be proven is shown in equation 3:

\[ 0.85 \leq \frac{A}{B} \]  

(3)

In this situation, all of the 95% CI for log-transformed data must be above the inferiority limit. However, here there is no superiority limit. Examples of this are depicted in figure 3. White bars represent 95% CIs for values of A/B that correspond to non-inferiority. The results may be close to or far above the inferiority limit and variability can be small or large. However, as long as the inferiority limit is not exceeded, the biosimilar is considered to be non-inferior to the reference product. Alternatively, the dark bars represent values of A/B that do not meet the definition of non-inferiority, since the 95% CIs for these values of A/B go below the inferiority limit, regardless of whether or not variability is low or high.

The non-inferiority approach has the advantage of requiring fewer patients to achieve statistical power.17 19 However, it is not adequate to assess biosimilarity of biopharmaceuticals used to treat rheumatological diseases, since it does not detect the situation in which a non-innovator is superior to the reference product and thus is bio-better and not biosimilar. A bio-better that produces a greater pharmacological response in some patients might have not only superior efficacy, but also increased toxicity.8 To date, no biosimilar has been approved using a non-inferiority approach to compare effectiveness to its reference product.

OTHER BIOSIMILARS
Presently, only one biosimilar has been approved by the EMA for use in rheumatology: the biosimilar infliximab CT-P13 (Remsima/Inflectra).2 This biosimilar monoclonal antibody has also been approved in Canada, Japan, Turkey, Colombia and South Korea,3 although with differences in the indications for which it has been authorised.4 CT-P13 has been submitted for review by the US Food and Drug Administration (FDA).3 There is worldwide interest regarding how the US FDA will consider this biosimilar. Will full extrapolation of indications be granted? Will additional clinical studies be requested? Such questions will soon be answered when the US FDA makes public its findings.

Other biosimilars to treat rheumatological diseases, including infliximab, adalimumab and etanercept, have been approved recently in India and South Korea.5–7 Other biosimilars are under development, many of which have been tested in clinical trials, and likely will reach the market over the coming years.2

The results of several trials of biosimilars in rheumatological diseases have been presented at major international rheumatology meetings. Bae et al22 compared an etanercept biosimilar, HD203, with innovator etanercept, Enbrel, in South Korean patients with rheumatoid arthritis inadequately responsive to methotrexate. A total of 294 patients were randomised 1:1 to receive either HD203 or innovator etanercept, each in combination with methotrexate. The primary end point was the proportion of patients achieving ACR20 response at week 24. Patients also were assessed at week 48. The difference in the ACR20 response rate between patients

Figure 3 Graphical representation of non-inferiority analyses. (A and B) are the efficacy parameters of biosimilar and innovator, respectively. Bars correspond to the 95% CIs of A/B. Non-inferiority is declared when the entirety of the 95% CI lies above the inferiority limit (white bars). Non-inferiority cannot be declared when the 95% CI goes beyond the inferiority limit.
treated with HD203 and innovator etanercept was 2.12% with a 95% CI of −7.65% to 11.89% after 24 weeks, and was 4.37% with a CI of −5.57% to 14.31% after 48 weeks. Equivalence was concluded, although the published abstract did not disclose the pre-established equivalence range. On the basis of this trial, HD203 was approved in South Korea. Whether regulatory agencies in other countries will request additional information before approving this biosimilar etanercept remains to be seen.

Kay and coworkers compared BOW015, an infliximab biosimilar, with innovator infliximab (Remicade) in patients with rheumatoid arthritis inadequately responsive to methotrexate. A total of 189 patients were randomised 2:1 to receive BOW015 or innovator infliximab. The primary end point was the proportion of patients achieving an ACR20 response at week 16, with an equivalence margin of ±23%. Data have been reported through 54 weeks. The differences in the ACR20, ACR50 and ACR70 response rates between patients treated with BOW015 and innovator infliximab were within this equivalence range, both by intention-to-treat and per protocol analysis. Furthermore, both biopharmaceuticals exhibited similar kinetics of these responses at several time points before the plateau phase of the time-response curve. The authors concluded therapeutic equivalence and BOW015 was recently approved as a biosimilar in India on the basis of this clinical trial. Whether other countries will grant approval to BOW015 as an infliximab biosimilar with the presently available clinical trial data remains to be seen.

As is evident, there is no standardised clinical assay for biosimilarity in rheumatological diseases at present. To date, studies of biosimilars to treat rheumatoid arthritis have used the ACR20 response as the primary end point. However, these clinical trials vary in their randomisation strategies, equivalence ranges and duration of treatment. Since additional biosimilars are being developed, the design of clinical trials to gain regulatory approval should be standardised.

BIOMIMICS OR INTENDED COPIES

Biomimics, also known as intended copies, are non-innovator biologics that had received marketing approval before biosimilar regulations were put in place. They currently are available in some Asian and Latin American countries. Biomimics cannot be considered to be biosimilars, since they have not been subjected to review by an appropriate regulatory body according to a prespecified regulatory pathway for biosimilar approval. Limited and flawed clinical information is available regarding biomimics used to treat rheumatological diseases.

Yisaipu is an etanercept biomimic that is manufactured and marketed by the Mexican company, Probiomed. An abstract describing a study comparing Yisaipu to innovator etanercept (Enbrel) was presented in the 2015 EULAR Annual Scientific meeting. The design of this study is confusing. Three groups of patients were studied. In the first group, 12 patients received Yisaipu and methotrexate for 24 weeks. In the second group, 12 patients initially received Yisaipu and methotrexate for 12 weeks, followed by innovator etanercept and methotrexate for the subsequent 12 weeks. Patients in these two groups were participating in a pharmacokinetic study. In the third group, 30 patients received Yisaipu and methotrexate for 24 weeks. End points were DAS28 at weeks 12 and 24, but a primary end point was not specified. The authors stated that DAS28 improved and that drug safety was similar in all treatment groups. However, it appears that only patients enrolled in the first and third groups received the same treatment regimen. The only distinction between these groups was that patients in the first group were participating in a pharmacokinetic study, whereas those in the third group were not. The authors concluded that “none [sic] significant difference was observed in the pharmacokinetic groups (p=0.355)”. As mentioned earlier, failure to detect a significant difference is evidence of neither equivalence nor of biosimilarity. This is particularly true when the sample size is evidently small (12 patients per group), yielding high-SD values. Thus, based on this study, Yisaipu cannot be considered to be an etanercept biosimilar. Whereas the primary reason for a biosimilar to be used instead of the innovator is its lower price, the cost of this biomimic in some acquisition contracts by the Social Security in Mexico has been the same as that of innovator etanercept.

As can be appreciated, there is an enormous difference in quality between the studies described in the abstracts about the biosimilars HD203 and BOW015, and those about the biomimics Etanar and Yisaipu. The data on HD203 and BOW015 likely will be published in peer-reviewed journals in the near future, allowing readers to assess the advantages as well as the limitations of these biosimilars. On the other hand, no
peer-reviewed publication documents the efficacy or safety of Yisaipu/Etanar/Endar or of Infinitam, despite these products having been available and used to treat patients with rheumatological diseases in several countries for several years.

Kikuzubam is a rituximab biosimilar that was manufactured and marketed in Mexico by ProbioMed.\(^2\)\(^8\)\(^25\) In 2012, the Mexican program of Pharmacovigilance issued a communication to health professionals warning them of anaphylactic reactions that occurred in several patients who were switched from innovator rituximab (Mabthera) to the biosimilar, or vice versa.\(^29\) This was surprising, since innovator rituximab had exhibited a very favourable peri-infusion safety profile among Mexican patients.\(^30\) Owing to these anaphylactic reactions and the lack of clinical data documenting the efficacy and safety of Kikuzubam, approval to market Kikuzubam in Mexico was withdrawn by the regulatory authority on 28 March 2014.\(^31\)

Reditux is a rituximab biomimic manufactured in India and marketed in India and in several Latin American countries.\(^2\)\(^8\)\(^9\) To the best of our knowledge, no clinical trial has been performed to demonstrate equivalent efficacy and safety of Reditux with innovator rituximab (Mabthera) in patients with a rheumatological disease. However, analytical studies have demonstrated significant differences in physicochemical properties between Reditux and innovator rituximab.\(^32\) Thus, Reditux cannot be considered to be a rituximab biosimilar.

**CONCLUDING REMARKS**

True biosimilars should not be confused with biomimics that have not undergone rigorous clinical testing or regulatory review according to a pathway for approval of biosimilars. There are only two options for biomimics: either document their efficacy and safety in well-designed comparative effectiveness clinical trials, and subject them to regulatory review so that they can become true biosimilars, or remove them from the market altogether.

The perception of biosimilars in rheumatology has evolved over the past several years. Distrust about the use of biosimilars in clinical practice has waned as regulatory pathways have been implemented and applied to the approval of a biosimilar infliximab. Although analytical and non-clinical evidence plays a considerable role, clinical trials that demonstrate comparable efficacy and safety of the biosimilar with its reference product are of paramount importance to establish biosimilarity. Familiarity of clinicians with issues regarding study design and the methods of analysing data obtained in comparative effectiveness clinical trials undoubtedly will improve decision-making surrounding the prescription of biosimilars to patients with rheumatological diseases.

**Acknowledgements**

The authors would like to thank Dr Ricardo Garcia (CLAPBio) for his useful comments and suggestions.

**Contributors**

GC-H, RG-R and MAS conceived and drafted, this manuscript. GC-H, RG-R, JK and MAS revised this review critically for important intellectual content and approved the version submitted.

**Funding**

This work was supported by CINVESTAV and CLAPBio.

**Competing interests**

GC-H has received consultancy fees from Amgen, Abbvie, Astra-Zeneca, Bayer, Boehringer-Ingelheim, Eli-Lilly, Janssen-Cilag, Laboratories Liomont, Laboratories Sophia, Medix, Merck-Serono, Merck, Sharp and Dohme, Novartis, Pfizer, Roche, Sanofi and UCB. RG-R has received consultancy fees from Amgen and Sanofi. JK has received research support (paid to the University of Massachusetts Medical School) from AbbVie Inc; Ardea Biosciences, Inc; Eli Lilly and Company; Pfizer Inc; and Roche Laboratories, Inc; and has received consultancy fees from Alexion Pharmaceuticals; Amgen, Inc; AbbVie Inc; AstraZeneca; Boehringer Ingelheim GmbH; Bristol-Myers Squibb Company; Crescendo Bioscience, Inc; Eli Lilly and Company; Epirus Biopharmaceuticals, Inc; Genentech Inc; Hospira, Inc; Janssen Biotech, Inc; Merck Sharp and Dohme Corp; Nippon Kayaku Co, Ltd; Novartis Pharmaceuticals Corporation; PanGenetics, BV; Pfizer Inc; Samsung Bioepis; Roche Laboratories, Inc; and UCB, Inc. MAS has received consultancy fees from Abbvie, Bioepis, Boehringer Ingelheim, Epirus, Janssen, Orygen, Pfizer and Samsung.

**Ethics approval**

This is a review article. No ethical approval was needed.

**Provenance and peer review**

Commissioned; externally peer reviewed.

**Data sharing statement**

No additional data are available.

**Open Access**

This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/

**REFERENCES**

21. Canadian Agency for Drugs and Technologies in Health. 


