To the Editor: Over the past few years, a completely new concept has been introduced into the medicines regulatory jargon — “biosimilar” medicinal products (or “biogenerics”). It refers to development of “essentially similar” medicinal products (“generics”) where the reference product is not a simple chemical entity, but a biotechnological product, ie, a therapeutic recombinant (glyco)protein (1). The concept has emerged in parallel with the expiry of the patent and other forms of protection of the first such proprietary drugs introduced into human medicine around 20 years ago, such as recombinant human growth hormone and recombinant human erythropoietin (epoetin). This latter molecule has attracted the most attention. First, epoetin is a rather complex molecule, where the protein (coded) part (the primary sequence, as well as the protein folding) is decisive for the receptor interactions, but the sugar moiety, which is approximately 40% of the molecule and is susceptible to huge variations depending on a variety of factors (expression system, culture conditions, pooling and downstream purification procedures, etc.), largely affects the biological activity of the molecule in vivo (2). Second, epoetin has been shown as a practically life-saving drug for chronic renal failure patients (and with expanding use in other indications), with high efficacy and very good safety (3). Third, the worldwide market for epoetin is expected to exceed 20 billion US dollars in the next few years (4). Clearly, even a small portion of this “cake” would be an attractive bite for almost any pharmaceutical company.

The legislation in European Union requires that any biotechnological product intended for use in humans needs to be evaluated (and eventually approved) through the so-called “centralized procedure” at the European Medicines Evaluation Agency (EMEA). This applies to biosimilar products as well (1). While the manufacturing/pharmaceutical (ie, quality) requirements that a biosimilar product needs to meet are identical to those asked for any proprietary (“new”) biotechnological medicinal product and are clearly set forth in the “quality” guidance documents that have evolved over the past 25 years, until recently there has been no clear-cut guidance or “list of requirements” related to non-clinical and, especially, clinical development program that needs
to be met in the case of biosimilar products. In June 2005, EMEA Committee for Human Medicinal Products (CHMP) released a draft guidance document with specific regulatory requirements for non-clinical and clinical development of biosimilar medicinal product containing recombinant human erythropoietin (5). Recently, this matter has been brought to the attention of Croatian public through a series of articles in public media, which were related to the fact that Croatia was the first country in Europe to approve a medicinal product containing epoetin on the grounds of “biosimilarity” claim. These articles clearly demonstrated a lack of understanding of the issue that might be confusing for the patients, as well as for medical practitioners dealing with the use of epoetin. The following few items need to be clearly stated regarding “biosimilar” epoetin(s):

The applicant needs to provide evidence that the product is manufactured in full compliance with the relevant quality requirements. By this, the “manufacturing/pharmaceutical” similarity of the test and the reference product would be demonstrated. This requirement is practically identical as for any (“conventional”) essentially similar drug product.

The applicant needs to provide evidence of “molecular comparability” of the test vs the reference active substance. This requirement represents a major difference between “biosimilar” and essentially similar conventional drugs. In the case of “conventional generics”, proving the identity of the molecule, together with the qualitative/quantitative impurities profile (against, eg, a pharmacopoeial monograph), is an evidence that the proposed test product contains molecularly exactly the same active substance as the reference product. In the case of “biosimilar” glycoproteins, such as epoetin, this is not possible. It is recognized that, due to the variability of the sugar moiety, it is practically impossible to produce “identical” epoetin molecules. Also, it is recognized that certain molecular differences (glycosilation) between two epoetins coming from different manufacturers (although using the same expression system) may not be detectable even with the state-of-the-art analytical procedures, and could be relevant in terms of efficacy and safety, in particular regarding immunogenicity. Immunogenicity of epoetin is related to possible hypersensitivity reactions and also to the generation of neutralizing antibodies which may lead to pure red cell aplasia. This is why the “comparability” concept has been introduced. In practical terms, proving molecular comparability of a biosimilar epoetin would require two steps. First, the applicant needs to prove not only that the proposed active substance is indeed human erythropoietin, but also that it fully complies with the molecular (physico-chemical) properties of recombinant erythropoietin as set forth in the pharmacopoeial monograph (6). The methodology and criteria for proving “comparability” are included in the monograph. Second, through a series of in vitro and in vivo (animal models) assays (methodology and criteria specified in the monograph), the applicant needs to show that the proposed substance performs comparably to the reference active substance (direct comparisons between the test and the reference compound).

Proving quality and “molecular comparability” are only the first steps in demonstrating that the proposed test product is equivalent to the reference one in terms of efficacy and safety. This, again, is a major difference vs the “essential similarity” concept applicable to conventional medicines. For a conventional product which is a solution for subcutaneous (with no sustained-release mechanism) or intravenous injection, and assuming qualitatively and quantitatively similar composition of the inactive ingredients, proving quality and identity would automatically result in a self-evident equivalent safety and efficacy vs the reference product. In the case of epoetin (which is a solution for s.c. or i.v. administration), this yet needs to be proven due to the “molecular uncertainties” mentioned ad 2), and includes toxicological and clinical testing. The toxicological program should include at least a 12-week repeated-dose toxicity study (by both routes of administration) in rodents, directly comparing the test and the reference product, with a focus on immunogenic responses, and a local tolerance comparison between the two products in at least one species. These studies need to be conducted in line with the specific guidance documents, and should demonstrate qualitative and quantitative similarity between the products.
Assuming that all the previous steps have been fulfilled in a satisfactory way, both methodologically and in terms of the outcomes, the clinical development program would be a final step in demonstrating “biosimilarity”. A usual misconception is that it is intended to prove that the test product “works”, ie, in the case of epoetin, that it cures renal anemia (renal anemia is used as a standard setting for clinical evaluation of any epoetin product). In fact, the clinical part needs to formally prove that the test product is EQUIVALENT to the reference product in terms of safety and efficacy. It should comprise at least the following elements:

Relative pharmacokinetic (PK) properties of the test and the reference product after a single-dose administration in a crossover mode need to be determined based on area under the concentration-time curve (AUC), maximum serum concentration (Cmax) and terminal elimination half-life (T1/2).

Although this looks similar to the requirement of “equivalent bioavailability” (bioequivalence) in the case of conventional generics, this requirement is much more complex. The test and the reference product both need to be administered by the i.v. and the s.c. routes – test i.v. vs reference i.v. (determination of relative bioavailability of test vs reference by the i.v. route); test s.c. vs reference s.c. (determination of relative bioavailability of test vs reference by the s.c. route – “compatibility” in the i.v. setting does not necessarily translate into “compatibility” in the s.c. setting). In practical terms, and assuming that only one dose-level is to be tested (therefore, skipping a comparative assessment of dose-linearity), this would require 2 separate 2x2 crossover studies, or one study with 4 treatments, 4 sequences and 4 periods. The important moment here is the fact that the current draft guidance DOES NOT automatically accept the “conventional” margins of “bioavailability equivalence” applicable for conventional drugs, ie, the 0.8 to 1.25 range for the geometric means ratios of test over reference product. In fact, the recommendation is that the “equivalence margins should be justified primarily on clinical grounds” (5). Here is an example of a potential practical meaning of this. A 2x2 crossover study aiming to prove bioequivalence of two epoetin products under the assumptions of the true test/reference ratio of 1, within-subject coefficient of variation of 30% (applicable for epoetin), power of 90% and with the standard acceptance range of 0.8 to 1.25 for the 90% confidence interval (CI) around the test/reference ratio would require 34 subjects. If the acceptance range is only slightly narrowed, ie, from 20% difference as acceptable, to 15% difference as acceptable, then the study would require 74 subjects for a power of 90% for demonstrating the equivalence of two truly equivalent products. On the other hand, if the acceptance range is to be 0.8 to 1.25, but the CI around the ratio is to be 95%, then 48 subjects would be needed. Finally, a narrower acceptance range with 95% CI would require 88 subjects (power 90%). And last, but not least, it might be required that the absolute s.c. bioavailabilities of two products are demonstrated comparable. Namely, certain differences between the two products might result in acceptable results of the i.v. and s.c. comparative studies, but may “sum-up” to result in different absolute s.c. bioavailabilities between the two products. In practical terms, such differences might indicate that dose-adjustments needed after switching a patient from an i.v. to s.c. (and vice versa) administration of the same epoetin product, might be different for the test and the reference product. This could be perceived as a “deviation from therapeutic equivalence.”

Therapeutic equivalence of test and the reference product needs to be demonstrated for “anemia correction” and “maintenance of hemoglobin levels,” and also for the s.c. and i.v. routes of administration. In practice, this could be managed by 2 separate studies – one “titration study,” where anemic patients would be included (for example, pre-dialysis patients), and products would be administered s.c.; and one “maintenance study,” where patients kept on “optimal” treatment with the reference product would be included and the products would be administered i.v. Both studies should be parallel-group, double-blind (or, at least the staff responsible for epoetin dose adjustments and data analysis should be blinded) and randomized. The “titration study” does not need to be longer than 12 weeks, while the “maintenance study” would need to last for at least 6 months (and both should probably have an open-label extension with test product of up to 12...
months – see below). Both studies should be focused on proving "efficacy equivalence" based on (primarily) per-protocol population, but also the intent-to-treat population ("robustness") and should have 2 co-primary outcomes – a "hemoglobin responder rate" (ie, a proportion), and "epoetin dose" (ie, a continuous variable illustrating epoetin utilization). In both studies, the equivalence evaluation should be based on the position of the 95% CI around the test-reference difference relative to a pre-defined acceptance range. The current draft guidance, however, DOES NOT define the "responder rates", nor does it define the "acceptance range" – except that both should be defined on the grounds of clinical acceptability. For the "titration study," and according to the common practice in the use of epoetin, the "responder rate" would be a proportion of patients who, over a 12-weeks period, stably attain a certain pre-defined hemoglobin level (typically 11-13 g/dL). With the "standard approved epoetin products", such a goal is typically achieved in around 90%-95% of the patients (7). Therefore, the study should be conceived as a "non-inferiority" study – with such a high reference rate, there is no point in proving equivalence, which is a "two-sided" concept, but the goal should rather be to prove "non-inferiority," which is a "one-sided" concept. In other words, there is no point in proving that the responder rate to the test product is "not lower and not higher" than for the reference product, but rather that it is not "relevantly lower" for the test vs the reference product. In practical terms, this means that the one-sided (lower) 97.5% confidence limit of a test-reference difference should not be more negative than a certain pre-defined (acceptable) margin. The key question, of course, is the definition of the "acceptable margin." Let's assume that the margin is defined as -10%, ie, the test product is to be considered non-inferior to the reference product if the 97.5% confidence limit of the test-reference difference in proportion of responders is not more than -10% (the proportion of test responders is not lower than 90%) A study with 90% power to prove non-inferiority of two truly non-inferior products (actual difference = 0) would require 410 patients, 205 assigned to the test and 205 assigned to the reference product in per-protocol population. Accounting for the drop-out rate of around 10% (patients not entering the per-protocol set due to various reasons), around 450 patients would actually need to be recruited. Reducing the acceptance margin to, let's say, 7% would almost double the required number of patients. For the "maintenance study," the definition of the "responder rate" is not so obvious. Some studies (7) have defined it as a proportion of patients in whom the Hb level does not drop from the pre-approval safety evaluation period. With such a definition, the reference responder rate may fall below 90% (ie, to 80% or lower) – and this means that the number of patients required for a certain power for demonstrating non-inferiority (or equivalence) would further increase.

Comparable safety of the test and the reference products needs to be demonstrated. It is considered that the pre-approval safety "data package" is sufficient when comprising data on around 300 patients treated with the test product for at least 12 months. This could be achieved by "extending" test treatments from the "titration" and "maintenance" studies in an open-label fashion. The safety evaluation has two steps. The first one is demonstration that the adverse reactions to the test product are "comparable" (not significantly different – the whole program, as described, is not "powered" for "safety equivalence" evaluation) to that of the reference product – by type, by severity and by incidence. This is achieved from the parallel group parts of the studies. The second one is demonstration of non-immunogenicity of the test product – this is achieved by evaluation of the data on the test product from the parallel-group periods and open-label extensions of the trials. It includes specific tests for neutralizing antibodies that need to be done for up to 12 months during the treatment, presumably every 3 months, and for all patients treated with the test.
product, i.e., at least for 300 patients. Finally, the applicant is obliged to provide a detailed pharmacovigilance plan for the post-approval period, since the side-effects of concern (related to immunogenicity) are rare and could be reasonably expected to be evaluated only with a broader exposure to the product.

In conclusion, clinical development of a biosimilar product containing epoetin is much more complex than in the case of "conventional generic drugs." There are attempts to standardize the requirements for such a procedure, and those appear to be on the "right track." Simplified discussion about the "number of patients that need to be included" have become common even in the public media, but those are typically not based on scientific facts and are likely to introduce confusion to all interested parties, primarily the patients, but also practicing medical professionals.

**Conflict of interest:** Vladimir Trkulja is a member of the Committee for Human Medicines at the Croatian Agency for Medicinal Products

**Vladimir Trkulja**  
vladimir.trkulja@zg.htnet.hr  
Department of Pharmacology  
Zagreb University School of Medicine  
Zagreb, Croatia


