

Biosimilars entering the clinic without animal studies

A paradigm shift in the European Union

Leon AGJM van Aerts^{1,2,3,*}, Karen De Smet^{4,5,6}, Gabriele Reichmann^{3,5,7}, Jan Willem van der Laan^{1,8}, and Christian K Schneider^{9,10}

¹Medicines Evaluation Board (CBG-MEB); Utrecht, Netherlands; ²Member of the Safety Working Party (SWP); EMA; London, UK; ³Expert of the Working Party on Similar Biological Medicinal Products (BMWP) of the Committee for Medicinal Products for Human Use (CHMP); EMA; London, UK; ⁴Federal Agency for Medicines and Health Products (FAMHP); Brussels, Belgium; ⁵Expert of the SWP of the CHMP; EMA; London, UK; ⁶Member of the BMWP; EMA; London, UK; ⁷Paul Ehrlich Institute (PEI); Langen, Germany; ⁸Chair of the SWP; EMA; London, UK; ⁹Danish Health and Medicines Authority (DHMA); Copenhagen, Denmark, and Twincore Centre for Experimental and Clinical Infection Research; Hanover, Germany; ¹⁰Chair of the BMWP; EMA; London, UK

The concept of biosimilars has spread from Europe to other regions throughout the world, and many regions have drafted regulatory guidelines for their development. Recently, a paradigm shift in regulatory thinking on the non-clinical development of biosimilars has emerged in Europe: In vivo testing should follow a step-wise approach rather than being performed by default. To not require animal testing at all in some instances can well be seen as a revolutionary, but science-based, step. Here, we describe the internal discussions that led to this paradigm shift. The mainstay for the establishment of biosimilarity is the pharmaceutical comparability based on extensive physicochemical and biological characterization. Pharmacodynamic comparability can be evaluated in in vitro assays, whereas pharmacokinetic comparability is best evaluated in clinical studies. It is considered highly unlikely that new safety issues would arise when comparability has been demonstrated based on physicochemical and in vitro comparative studies.

Introduction

Decisions on marketing authorization of medicinal products, and to even a larger extent approval of clinical trials, are partly based on non-clinical data, including animal studies. Contrarily to applications for generic medicinal products, these studies have traditionally been requested for biosimilars (see

definition below) from early on. However, in the European Union (EU), the need for animal studies in the development of biosimilars has recently been internally discussed extensively by regulators, which led to a shift in paradigm as visible in currently emerging guidance documents. To explain the background of the change in the EU, we present the main arguments that were at the basis of the change in paradigm, which can well be seen as a revolutionary, but science-based step. The strategy is revolutionary, because it implies that there can be biosimilar developments with no animal testing at all, and because it implies that regulators may even discourage developers from performing such studies. It is also science-based because, rather than requiring a “tick box” approach for non-clinical testing, the new strategy implements a step-wise, knowledge and science-driven approach, thus further guiding non-clinical development of biosimilars toward the goal of employing suitable test systems, especially those that will give the best results for establishing biosimilarity.

A generic medicinal product is a medicinal product that has *the same* qualitative and quantitative composition in active substances and the same pharmaceutical form as the reference medicinal product. In other words, chemically it is an exact copy of the reference product, although minor variations such as different salts are allowed, as long as pharmacokinetic (PK) bioequivalence can be

Keywords: biosimilar, follow-on biologicals, comparability, safety, regulatory expectations

© Leon AGJM van Aerts, Karen De Smet, Gabriele Reichmann, Jan Willem van der Laan, and Christian K Schneider

*Correspondence to: Leon AGJM van Aerts; Email: la.v.aerts@cbg-meb.nl

Submitted: 06/01/2014

Revised: 07/05/2014

Accepted: 07/07/2014

<http://dx.doi.org/10.4161/mabs.29848>

This is an Open Access article distributed under the terms of the Creative Commons Attribution-Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited. The moral rights of the named author(s) have been asserted.

demonstrated. Biological medicinal products similar to a reference medicinal product (i.e., biosimilars) do not usually meet all the conditions to be considered for a generic medicinal product mainly due to molecular characteristics, manufacturing process characteristics, and the raw materials used.¹

Scientific guidance on the development of biosimilars² is based on the concept that an exact copy of a biological reference product cannot be produced due to technical and inherent limitations (as also no batch of a given biological can be an exact copy of the previous one). Also, due to the complexity of biologicals, there have been limitations to the extent to which these products could be characterized using physicochemical methods. Today, these methods have advanced to a considerable extent, and now many are suitable to detect subtle differences between the biosimilar and the reference product. Yet, the clinical relevance of the observed subtle structural and compositional differences (i.e., the effect on efficacy and safety) is often not clear based on the analytical data alone. Therefore, these differences are being further evaluated in non-clinical and clinical studies, as reflected in the overarching biosimilar guidelines.^{2,3}

In the EU, a new Directive on the protection of animals used for scientific purposes was issued in 2010,⁴ which updates and replaces the 1986 Directive 86/609/EEC. The aim of the new Directive is to strengthen legislation, and improve the welfare of those animals still needed to be used, as well as to firmly anchor the principle of the “Three Rs,” to Replace, Reduce and Refine the use of animals, in EU legislation. Directive 2010/63/EU has taken full effect from 1 January 2013. According to this Directive, the use of animals for scientific or educational purposes should only be considered where a non-animal alternative is unavailable (preamble 12) and Member States shall ensure that, wherever possible, a scientifically satisfactory method or testing strategy, not entailing the use of live animals, is used instead (Article 4.1). Moreover, non-human primates (NHP) are exempted from use in animal studies whenever possible. This is reflected in Article 8.1(b) as there

should be scientific justification that the purpose of the procedure (animal study) cannot be achieved by the use of species other than NHPs.

Concurrently, a guideline on the non-clinical and clinical issues of the development of biosimilar monoclonal antibodies (mAbs) was being drafted by the Working Party on Similar Biological Medicinal Products (BMWP) of the Committee for Medicinal Products for Human Use (CHMP).⁵ The discussions around this guideline made clear that the conventional paradigms regarding toxicity testing for biosimilar mAbs, and thus potentially for other classes of biosimilars, had severe limitations related to the high species and target specificity of mAbs (thus in many cases rendering any other species than NHP non-relevant).

The main issues that have arisen during the discussions in the BMWP on the use of animal studies to address efficacy and safety of claimed biosimilars and how these have led to a shift in paradigm in the regulatory thinking in the EU on the non-clinical development of biosimilars are schematically depicted in **Figure 1** and will be addressed below. Where previously a non-clinical package for a biosimilar development was expected to consist of comparative studies, including a pharmacodynamic study (bioassay) and a repeated dose toxicology study, a new paradigm has emerged from these internal discussions, in which the use of animals is obviated in most cases by a thorough step-wise approach of testing.

The main arguments for this paradigm shift are not only important for biosimilar mAbs, but can be applied to other biosimilar medicinal products as well.

The discussions in this paper may not be only relevant for marketing authorization applications for biosimilars, but may also assist ethics committees or national regulatory authorities deciding on the acceptability of animal experiments and clinical trials for biosimilars or even biologicals in general. We aim to clarify that safety of patients will not be compromised when clinical trials with biosimilars are initiated without preceding animal studies.

The concept of biosimilar development was first adopted in the EU. Meanwhile,

this concept has spread and now other regions throughout the world have issued or drafted regulatory guidance for the development of biosimilars. Generally, a similar approach to the non-clinical development is taken as initially was done in the EU, i.e., requesting animal testing; sometimes even more extensive animal studies are demanded. By explaining the paradigm shift in the EU, we hope to contribute to a more global consideration of the 3Rs principles where the development of biosimilars is concerned.

Key Aspects of Biosimilar Development that Limit the Suitability of Animal Studies

As for all biologicals, the pharmaceutical quality of a biosimilar has to be demonstrated, including a complete description of the manufacturing process and full characterization of the quality attributes. In addition, comparability of these quality attributes between biosimilar and reference product has to be shown. The molecular nature with equal amino acid sequence of the biosimilar and its reference product, demonstrated by comparative physicochemical and biological assays, forms a solid basis for establishing biosimilarity. A critical feature of biosimilars development is that instead of demonstrating efficacy and safety of the product per se, it should be demonstrated that the product is (*highly*) *similar* to the reference product. In essence, this means that, next to a high level of physicochemical similarity, no clinically-relevant differences in safety or efficacy are detected as a consequence of potential subtle differences in quality attributes between the products. To evaluate whether there are clinically meaningful differences, both clinical and non-clinical studies can be considered. The relevance of animal studies for evaluating the clinical relevance of differences in quality attributes is discussed in more detail in the sections on process-related impurities and product-related substances. When comparative clinical and non-clinical studies have to be powered to detect all relevant differences in efficacy and safety, these studies would have to be of considerable size. While sensitivity can

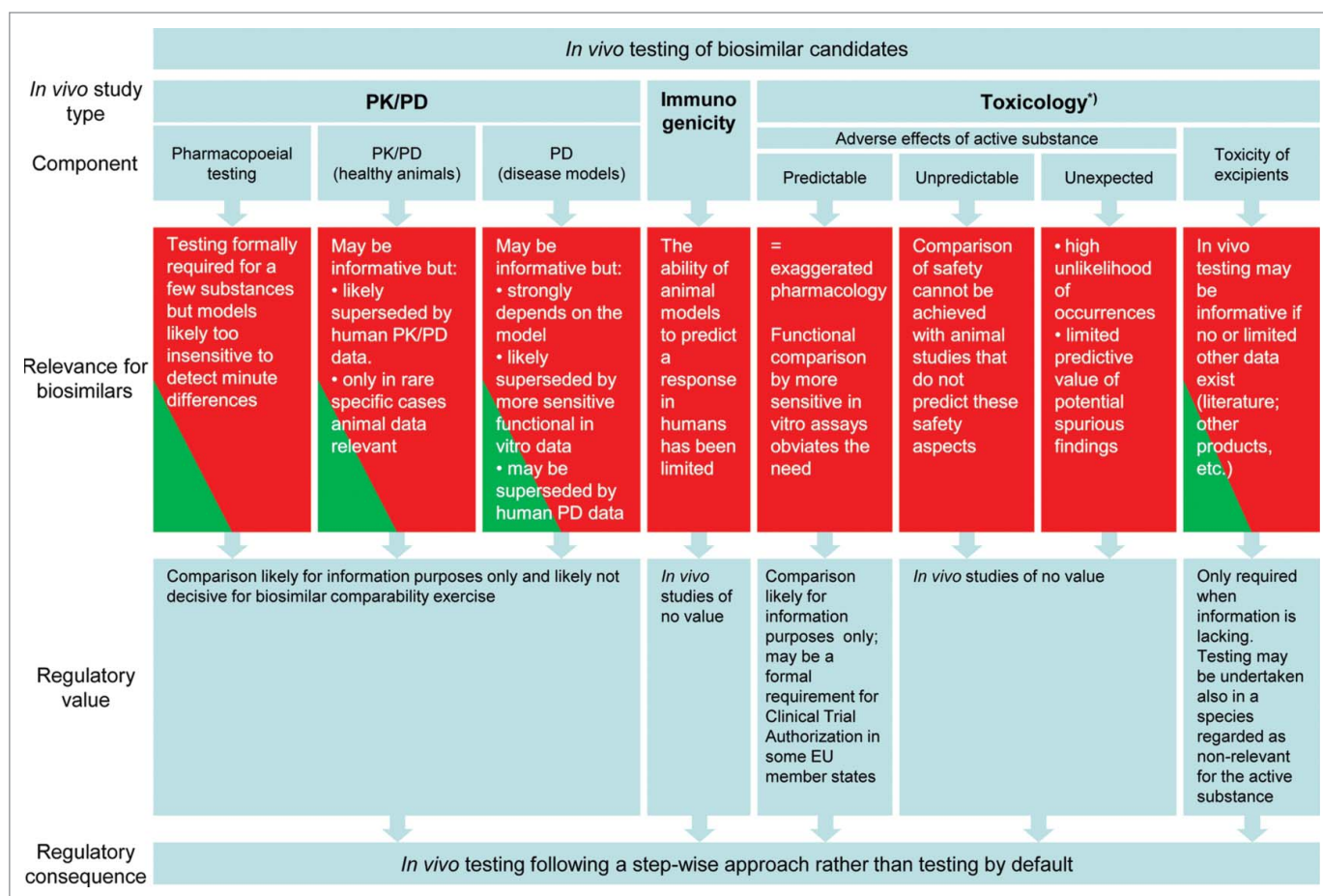


Figure 1. Overview on the relevance and regulatory value of in vivo studies for biosimilar candidates. ^{*)} specific studies like carcinogenicity, reproductive toxicity etc. are not required for biosimilars.

be an issue for clinical studies, this is even more the case for animal studies. The usually small group size in animal studies (especially when non-rodents are being used) limits the sensitivity of these studies to detect relevant differences in safety and efficacy.

Furthermore, species differences between animals and humans create another hurdle. When a difference is seen in an animal study, it needs to be considered whether or not the observed difference is relevant for humans, but also vice versa, when no differences are observed in an animal study, there may still be remaining uncertainty whether or not there will be any meaningful differences in humans.

Another aspect regarding biologicals, especially when mAbs are concerned, is that, due to the high specificity of these molecules, only species closely related to humans are pharmacologically responsive.

In many cases that means that only NHPs are suitable to detect a pharmacological response of the investigational medicinal product. Studies in NHPs have notably small group sizes, and interindividual variability further reduces the sensitivity of these studies to detect differences in pharmacological response. Thus, the low sensitivity of animal studies to detect differences between biosimilar and reference product and species differences between animals and humans limit the suitability of this approach to evaluate biosimilarity.

Pharmacodynamics in vitro or in vivo

Suitability of an in vitro approach

Most biologicals are designed to interact with the body in a very specific way,

e.g., by binding to a receptor, ligand or substrate. The resulting functional effects are often detectable at a molecular or cellular level. Both the binding to the target and the subsequent functional effects can be assessed in in vitro assays using human cells or human receptors. As a matter of fact, such assays are often already available as part of the pharmaceutical characterization. In the case of mAbs, besides the binding of the complementarity-determining region (CDR) to its primary target, the Fc portion of the molecule also contains binding sites to different receptors, which may elicit several effector functions, notably complement activation, complement-dependent cytotoxicity (CDC) and antibody-dependent cell-mediated cytotoxicity (ADCC). These Fc-related binding properties and effector functions can also be evaluated in vitro.

Because the biological properties of a biological can be characterized *in vitro*, there is little – if any – further information that would be gained by *in vivo* models. Thus, there is no need to re-establish the pharmacodynamic response in an *in vivo* model. Moreover, using cellular systems, more extensive, precise and thus more sensitive comparisons can be made in *in vitro* assays, which further strengthens the *in vitro* approach in the evaluation of biosimilarity compared with the use of animal studies.

Pharmacopoeial *in vivo* bioassays

For some biologicals, pharmacopoeial bioassays exist in which the potency is determined in animals. Such assays are employed to express the biological activity of the product in International Units (IU). For those products where the drug substance is extracted from a biological matrix and is characterized only to a limited extent, this approach appears sensible. However, poorly characterized products are not very good candidates for a biosimilar development. On the other hand, well-characterized products such as recombinant proteins under development as biosimilars may be compared with their reference product using *in vitro* assays and the dose declared on a mass content basis (i.e., µg instead of IU). Thus, there is no *a priori* need to determine their activity using a pharmacopoeial *in vivo* bioassay. If such assays are performed for commercial or traditional reasons, companies may compare the activities measured. Yet, variability in the assays may lead to considerable ranges in activity for both biosimilar and reference product.^{6,7} Although the ranges would be expected to be similar, the variability of the *in vivo* assays limits their use in a comparability exercise, and the results would not significantly contribute to establishing biosimilarity. As far as the pharmacodynamic component of biological activity is concerned, *in vitro* alternatives to the pharmacopoeial *in vivo* assays can be employed. For example, measuring proliferation of an erythropoietin-dependent cell line instead of using the normocythaemic assay in mice to compare biological activity of erythropoietin;⁸ or using homologous (e.g., granulosa cells or Sertoli cells) or heterologous

(transfected cell lines expressing the target receptor) assay systems instead of using the Steelman-Pohley assay to compare biological activity of a human follicle stimulating hormone (FSH).⁹ Homologous systems have the advantage of being similar to the natural environment, but the disadvantages are the need to harvest cells from animals with a limited yield, the intrinsic variability and the non-human nature of the cells. When using heterologous systems, advantages are the possibility to use the human receptor as target and the limitless availability of a cell line. On the other hand these heterologous constructs express the target only and may lack other relevant components that interact *in vivo*. Novel approaches such as the use of differentiated human stem cells may have potential for the development of comparative cell assays, combining the advantages and avoiding the disadvantages of homologous and heterologous assays in an optimal way.¹⁰

We acknowledge that quality attributes of a biological may also affect its PK behavior, which is not covered by the *in vitro* assay. However, it is preferable to establish PK similarity in human volunteers or patients, as will be discussed later on. Furthermore, even when a choice is made to determine activity in an *in vivo* bioassay, manufacturers should be aware that there is no obligation to use this assay routinely for every batch because there are possibilities to correlate biological activity with physicochemical determinants that can be used to predict the activity of subsequent batches.^{6,7}

Pharmacodynamic models of disease

For some diseases, animal models exist in which the pharmacological activity of a pharmaceutical can be shown. For instance, SCID mice with xenotransplants of tumors can be used for oncology products, or, for rheumatoid arthritis, transgenic mouse models, such as Tg197 carrying a modified human TNF gene construct,¹¹ could be employed. Although it is possible to compare pharmacological activity of a biosimilar and a reference product in such models, the outcome would depend on the robustness of the model. Variable growth of xenotransplants and semi-quantitative scoring of

pathological features would, however, decrease the usefulness of the models to detect differences in biological activity. As long as *in vitro* pharmacodynamic assays are available and there is a need to perform studies in humans where pharmacodynamic (if they exist) and other parameters can be compared, it seems that the contribution of animal models of disease to the totality of evidence for establishing biosimilarity is rather limited.

Pharmacokinetics

It may be possible to compare PK properties of a biosimilar and a reference product in an animal study to some extent, but these studies will always have their limitations. The PK of a biological are dependent on many factors, such as the presence and preponderance of the binding target, whether this target is soluble or not, the mechanism of clearance (e.g., receptor-complex mediated or not), and, in the case of mAbs, on Fc-dependent mechanisms. Species differences in these factors limit the suitability of animal models for PK investigations and limit the relevance of animal data with regard to PK in humans. Consequently, the PK of biologicals can best be appreciated when the compound is administered to humans. *In vitro* methods to determine the PK of a biological have limited use, if any at all. As so many factors are important and even differences between healthy volunteers and patients may be relevant, it is obviously necessary to assess and compare the PK properties of a biosimilar and the reference product in the whole organism, i.e., in humans. Therefore, comparative PK assessment in animals contributes little to the comparability exercise undertaken for evaluation of biosimilarity, and, for this aspect, judgment will have to rely on data obtained in plasma samples from humans.

An exception to the rule that human PK data prevail over non-clinical data could be when there is a need to establish if biosimilar and reference product distribute similarly to tissues. This could be the case when differences in glycosylation affect the uptake in target tissues, for example, different levels of mannosylation

of recombinant proteins used for replacement therapy could potentially affect distribution to target tissue. Whereas plasma PK data are easily obtainable in humans, this could prove more difficult when target tissue data are needed. In a non-clinical model, the actual tissue concentrations could be measured, and the pharmacodynamic response could be evaluated at the tissue level.

Differences in Formulation

A biosimilar product may contain excipients that are different from the reference product. It is therefore conceivable that the *in vivo* PK behavior of the active substance is affected. For widely used excipients employed to maintain the right osmolality and pH, there would in general be no concern at all, but less well-known substances may raise a concern. In case a completely new excipient is introduced or when there is no experience with the intended route of administration for the excipient, safety could be an issue. This has to be evaluated in accordance with relevant guidance and may include toxicology studies when insufficient information is available. However, when there are no safety concerns, a potentially different PK behavior can in principle be evaluated in humans, and there would be no need to first test this in animals. Hypothetically, when the product is a pharmaceutical with a steep dose response curve with a poor safety profile, making it ethically difficult to study the compound in volunteers and yielding suboptimal doses to patients being equally unacceptable, there could be a reason to compare the PK behavior of biosimilar and reference product in animals first.

Safety Evaluation

Toxicity studies are used for the safety evaluation of pharmaceuticals, including biologicals. However, the value of extensive NHP use in routine safety studies of mAbs has been questioned, as toxicity of these products is characterized as exaggerated pharmacology.¹² Regarding safety evaluation of biosimilars, we will explain

in this section that there is no need for toxicology studies. To do this, we would like to address three different areas of adverse effects: predictable, unpredictable and unexpected.

Like efficacy, *predictable* adverse effects of biologicals are considered to be related to the pharmacology and are referred to as exaggerated pharmacology. For this area of adverse effects, animal models have shown to be predictive for adverse outcomes in humans.^{13–17} Yet, as argued before, once comparable pharmacological activity has been established *in vitro*, there is no need to confirm these properties in a less sensitive animal model. This accounts also for the exaggerated pharmacological properties (to which, for any biological, many “adverse effects” belong).

Some types of adverse effects are not predicted by animal studies and can be designated *unpredictable* adverse effects. For example, progressive multifocal leukoencephalopathy due to a manifestation of a latent JC virus after administration of an immunosuppressive drug is not predicted by animal studies because these animals do not carry the same virus. Even in humans, the manifestation only occurs at a low rate, which would make it unlikely to be picked up in an animal study due to the small numbers used, even when they would be sensitive to this effect. Another example is toxicity based on cytokine release syndrome. This could occur as a consequence of the disease state of the patient, e.g., in certain oncology conditions where lymphocytes are targeted, or as a result of human-specific response due to differences in T-cell reactivity to a drug, as was the case in the TGN1412 event.^{18,19} Infusion reactions, which are also associated with the release of cytokines, may be triggered by different causes, including process-related impurities acting through TOLL-like receptors (TLRs), which are part of the innate immune system. Although the innate immune system is evolutionary an old system and TLRs are widely distributed,²⁰ the reactivity to the different triggers varies a lot between species, and even intraspecies variability due to polymorphisms may determine the response.²¹ Taken together, this makes animal models poor predictors of cytokine release in

humans. Tailor-made *in vitro* assays could be employed to assess this issue when needed.²² When there is a concern related to the presence of process-related impurities, *in vitro* TLR assays may have potential to detect and identify such contaminants.²³

A final example of unpredictable adverse effects is the formation of anti-drug antibodies (ADA) which are cross-reactive with an endogenous protein. The most well-known example is probably the development of pure red cell aplasia after administration of erythropoietin.²⁴ As animals are prone to produce antibodies against human proteins or humanized biologicals, which may or may not be cross-reactive with the animal's endogenous molecule, this sequence of events in humans is not predicted in animal studies.

The common denominator for the examples above is that adverse effects in humans are not predicted by animal studies. When an animal model has been proven to be unreliable with respect to safety evaluation of the biological concerned, it may be obvious that there is no good reason to employ animal studies for assessing biosimilarity in terms of comparable safety.

The third area of adverse effects, and seemingly the most elusive one, is the type of adverse effects to which we refer here as *unexpected* toxicity. Sometimes the terms ‘off-target toxicity’ and ‘non-specific toxicity’ are used in a similar way. For new chemical entities, it can be a sound reason to perform toxicology studies. However, for biosimilars we think this is not the case. Philosophically, one might argue that it is not possible to exclude unexpected toxicity because non-anticipation of the adverse effects is inherent to the unexpectedness of these effects. Yet, we should consider this aspect more pragmatically. When discussing this issue at a more open playing field, e.g., at European Medicines Agency (EMA) workshops where several interested parties were present,²⁵ the examples of unexpected toxicity presented were scarce. None of these examples actually concerned biosimilars, but rather biologicals that were still in development. Relatively few details were provided on these cases, which makes it difficult to critically assess their relevance

for the issue. One example that has been given some more attention, and for which data have been published, is a case where thrombocytopenia occurred after administration of a mAb under development, whereas this did not happen with four other mAbs directed at the same pharmacological target.²⁶ In this case, the mAbs had the same target, but were not biosimilars. Subsequently it was shown that the functional differences between these antibodies had a structural basis: The unexpected toxicity was shown to be driven by one to three amino acid differences in the light chain.²⁷ Obviously, such structural differences are not within the scope of a biosimilar development. Although the observed toxicity came unexpectedly for the company developing the drug, it is not an example that could be applied to biosimilars when structural and functional similarity have already been shown by analytical and in vitro methods. Thus, unexpected toxicity may be encountered in animal studies in rare cases during the development of new biological entities; it has never been shown to occur during the development of a biosimilar. Also, it was stated during aforementioned discussions at the EMA workshop that, following changes to the production process of an already marketed biological, unexpected toxicity by the post-change product was never encountered in animal studies.

Process-Related Impurities

As the raw materials used and the cell lines expressing the biosimilar and the original biological may be different and the purification steps for biosimilar and reference product are unique for each process, a potential difference between both products is the different level and type of process-related impurities, such as host cell DNA or host cell proteins. These latter molecules may constitute a “danger signal” to the immune system triggering an immune response in the patient through TLR or may cause a hypersensitivity reaction in sensitized individuals (or cause a sensitization). Yet, as these types of reactions are by and large unpredictable by animal studies it is not meaningful to conduct animal studies for this purpose.

As already stipulated in the ICH S6 guideline on preclinical safety evaluation of biotechnology-derived pharmaceuticals,²⁸ it is preferable to rely on purification processes to remove impurities and contaminants rather than to establish a non-clinical testing program for their qualification. The levels of process-related impurities should be kept to a minimum, which is the best strategy to minimize any associated risk.

Product-Related Substances

As for any biological medicinal product, the biosimilar medicinal product is defined by the molecular composition of the drug substance resulting from its process, which may introduce its own molecular variants, isoforms or other product-related substances. Deamidated and oxidated forms may be present and terminal amino acid truncations or modifications may occur. Post-translational modifications also lead to a range of different glycoforms and sialylation patterns. Also, dimers, oligomers or heteromolecules may be formed and result in a different mixture of product-related substances. The relative proportion of all these substances in the final product needs to be determined in order to establish consistency of the product.²⁹ When a biosimilar is produced, the final composition of the product is likely to be slightly different from the reference product regarding the relative contribution of all the product-related substances. It is therefore important to determine that the differences in composition of biosimilar and reference product are only minor and to establish that these minor differences do not affect safety and efficacy.

The pharmacological activity of biologicals is determined by their biological properties. Due to the high pharmacological specificity of biologicals, a change in a relevant part of the structure of the molecule is likely to lead to a change of biological activity, whereas minor variations in other parts are most likely not going to affect the functional properties of the molecule. For instance, when deamidated and oxidated forms of interferon (IFN)- α 2b were evaluated in a HEK293 cell line stably transfected with luciferase gene under

the control of interferon-stimulated response element promoter, a modest loss of bioactivity in oxidized interferon was observed, but the deamidated form essentially retained its activity.³⁰

Another example is the contribution of different glycoforms of a mAb to its function since parts of the sugar moiety in the Fc region of IgG molecules contribute to the binding to Fc gamma receptors (Fc γ R).³¹ For example, the binding affinity of a mAb for Fc γ RIIIa is influenced by the presence or absence of fucose, where a higher degree of afucosylation leads to an increased binding affinity for the receptor. Subsequently, the difference in binding may be functionally reflected by an increased potency in an ADCC assay.³² Other examples where different glycoforms affect functionality are IgGs containing glycans lacking terminal sialic acid and galactose and terminating in N-acetyl glucosamine (GlcNAc) being capable of binding mannose-binding lectin, which leads to activation of the lectin pathway of complement activation.³³ Alternatively, the addition of terminal sialic acid to the N-linked glycan reduces Fc gamma receptor binding and converts IgG antibodies to anti-inflammatory mediators through the acquisition of novel binding activities.³⁴

Differences in structure, like in the examples above, are detected when comparability of quality characteristics is evaluated, including characterization of glycosylation. It is expected that any potential change in pharmacological activity occurring as a consequence of the slight differences in structure, will be detected when biosimilar and reference product are compared in binding affinity and functional in vitro assays.

It is not anticipated that minor differences in the levels of product-related substances will lead to the introduction of completely new and unexpected properties, except for a potential effect on immunogenicity (see below) and hypersensitivity. Concerning hypersensitivity, a well-known example is the presence of a glycan with a terminal galactose- α -1-3 galactose configuration in cetuximab causing hypersensitivity reactions in individuals with IgE antibodies against this structure.³⁵ Where knowledge is available

that such structures have the propensity to trigger hypersensitivity reactions, greater levels of these structures in a biosimilar than in the reference product should be avoided. For structures for which it is not known if hypersensitivity is an issue, animal studies do not provide a predictive model to find out. To prevent safety issues in patients, the presence of glycoforms not observed in the reference medicinal product is best avoided in the development of a biosimilar.

Immunogenicity

Formation of ADA in patients is an important issue that needs to be evaluated for all biologicals - both originator products and biosimilars - before they enter the market. Binding antibodies may affect the PK and neutralizing antibodies can lead to loss of efficacy. When ADA are cross-reactive with endogenous proteins, this may even constitute a serious safety issue, as is the case for erythropoietin-induced pure red cell aplasia.²⁴

The ability of animal models to predict a response in humans has been limited by differences in primary sequence and structure between the human protein and the particular animal ortholog, differences in MHC-peptide binding and presentation, differences in the T-cell repertoire and other aspects of the immune response.³⁶ The propensity of animals, including NHPs, to produce ADA against human-targeted biologicals is often greater than it is for humans.³⁷ When a biosimilar and a reference product would induce different levels of ADA in an animal experiment, this does generally not provide a clue for the propensity of both products to induce ADA in patients. For an interferon (rhIFN β -1a), a transgenic mouse model has been generated in which the animals are immune tolerant for human IFN.³⁸ In this transgenic mouse model, it was shown that oxidation-mediated aggregation increased the immunogenicity of rhIFN β -1a, whereas aggregated preparations devoid of measurable oxidation levels were hardly immunogenic. Although this model appears to discriminate between immunogenic and non- or less immunogenic forms of IFN, the predictive value

for humans is unknown. But even when a correlation with ADA formation in patients could be anticipated, the observed differences in the mice appeared to be related to considerable differences in the levels and types of aggregates present in these preparations. Knowing that the presence of aggregates may increase the immunogenicity, their levels should be similar (or lower) in a biosimilar compared with the reference product. This can be determined by analytical methods. It may thus be concluded that animal studies have no role in a comparability exercise to establish the biosimilarity of a product with regard to its propensity to induce ADA in humans.

Conclusion and Outlook

The molecular nature with equal amino acid sequence of the biosimilar and its reference product, demonstrated by comparative physicochemical and biological assays, forms a solid basis for establishing biosimilarity. The arguments provided in this paper show that in vivo animal studies rarely provide decisive information to this end. Schematically, these arguments are shown in **Figure 1**. The functional properties of the biosimilar can be tested and compared in vitro and these assays are generally more sensitive than animal studies. These in vitro assays would cover both pharmacodynamic aspects and predictable adverse effects since the latter are in fact (exaggerated) pharmacology as well. For those adverse effects that cannot be predicted by animal studies, animal toxicology studies are obviously not suitable for a comparability exercise. Furthermore, a physicochemically and biologically well-characterized biological for which close similarity with a well-known reference product has been demonstrated based on extensive analytical and in vitro data are highly unlikely to pose a safety concern different from the reference product, with the exception of immunogenicity issues. For the latter, animal studies have no predictive value. To minimize any risk associated with the presence of impurities, these impurities should be kept to a minimum instead of setting up a non-clinical evaluation

program. For the PK comparability, human data will be more informative and hence these data would supersede the animal data. Only in very rare cases (e.g., when comparative tissue distribution data would be needed) animal PK data could be relevant.

The need or not for in vivo studies for biosimilar candidates will therefore be driven not only by availability of specific animal models for testing, but also on what level of evidence one can feasibly generate to inform the biosimilar comparability exercise. It will be interesting to see if animal testing may or may not become more prominent when biosimilars for orphan conditions would one day be developed (e.g., for enzyme replacement therapy). The limited number of patients available for studying comparability of biosimilar and reference product may pose a problem as regards clinical comparability. The absence of sufficiently powered studies in humans could increase the relative weight of non-clinical pharmacodynamic and PK data in the totality of evidence when biosimilarity needs to be established for this kind of products.

The change in paradigm presented in this paper reflects the need to search for alternative methods to obtain scientifically valid data before an animal experiment is conducted, which is now firmly expressed in EU legislation. We encourage a more global discussion on these aspects and it will be interesting to see if this could lead to a more harmonized approach in which the use of animals for scientific purposes is restricted to those instances where necessary data can only be obtained by animal experiments.

Disclosure of Potential Conflicts of Interest and Financial Disclosure Statements

The authors declare that there are no conflicts-of-interest and they have no financial relationship that would require disclosure in connection with the work presented in this paper.

Disclaimer

The views expressed in this article are the personal views of the authors and may not be understood or quoted as being

made on behalf of or reflecting the position of the European Medicines Agency or one of its committees or working parties.

References

- Consolidated Directive 2001/83/EC of the European Parliament and of the Council of 6 november 2001 on the community code relating to medicinal products for human use. *Official Journal L* – 311, 67-128 (28/11/2004).
- Committee for Medicinal Products for Human Use (CHMP). Guideline on Similar Biological Medicinal Products (Draft). CHMP/437/04 Rev 1. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2013/05/WC500142978.pdf (2013).
- Committee for Medicinal Products for Human Use (CHMP). Guideline on similar biological medicinal products containing biotechnology-derived proteins as active substance: non-clinical and clinical issues (Draft). EMEA/CHMP/BMWP/42832/2005 Rev. 1. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2013/06/WC500144124.pdf (2013).
- The European Parliament and the Council of the European Union. Directive 2010/63/EU of the European Parliament and of the Council of 22 September 2010 on the protection of animals used for scientific purposes. *Official Journal of the European Union L* 276, 33-79 (20/10/2010). <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2010:276:0033:0079:EN:PDF>
- Committee for Medicinal Products for Human Use (CHMP). Guideline on similar biological medicinal products containing monoclonal antibodies – non-clinical and clinical issues. EMA/CHMP/BMWP/403543/2010. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2012/06/WC500128686.pdf (2012).
- Mulders JWM, Wijn H, Theunissen F, Machielsen P, Jansen P. Prediction of the in-vivo biological activity of human recombinant follicle-stimulating hormone using quantitative isoelectric focusing, optimization of the model (Conference Paper) *Pharm Pharmacol Comm* 1999; 5:51-55.
- Zimmermann H, Gerhard D, Hothorn LA, Dinger-mann T. An alternative to animal testing in the quality control of erythropoietin. *Pharmeur Bio Sci Notes* 2011; 2011:66-80; PMID:21619857
- Miyazaki Y, Kuriyama K, Higuchi M, Tsushima H, Sohma H, Imai N, Saito M, Kondo T, Tomonaga M. Establishment and characterization of a new erythropoietin-dependent acute myeloid leukemia cell line, AS-E2. *Leukemia* 1997; 11:1941-9; PMID:9369430; <http://dx.doi.org/10.1038/sj.leu.2400838>
- Zambrano E, Zariñán T, Olivares A, Barrios-de-Tomasi J, Ulloa-Aguirre A. Receptor binding activity and in vitro biological activity of the human FSH charge isoforms as disclosed by heterologous and homologous assay systems: implications for the structure-function relationship of the FSH variants. *Endocrine* 1999; 10:113-21; PMID:10451219; <http://dx.doi.org/10.1385/ENDO:10:2:113>
- Lan C-W, Chen M-J, Jan P-S, Chen H-F, Ho H-N. Differentiation of human embryonic stem cells into functional ovarian granulosa-like cells. *J Clin Endocrinol Metab* 2013; 98:3713-23; PMID:23884780; <http://dx.doi.org/10.1210/jc.2012-4302>
- Keffer J, Probert L, Cazlaris H, Georgopoulos S, Kaslaris E, Kiousis D, Kollias G. Transgenic mice expressing human tumour necrosis factor: a predictive genetic model of arthritis. *EMBO J* 1991; 10:4025-31; PMID:1721867
- van Meer PJ, Kooijman M, van der Laan JW, Moors EH, Schellekens H. The value of non-human primates in the development of monoclonal antibodies. *Nat Biotechnol* 2013; 31:882-3; PMID:24104750; <http://dx.doi.org/10.1038/nbt.2709>
- Polson AG, Fuji RN. The successes and limitations of pre-clinical studies in predicting the pharmacodynamics and safety of cell-surface-targeted biological agents in patients. *Br J Pharmacol* 2012; 166:1600-2; PMID:22364106; <http://dx.doi.org/10.1111/j.1476-5381.2012.01916.x>
- Bugelski PJ, Martin PL. Concordance of preclinical and clinical pharmacology and toxicology of therapeutic monoclonal antibodies and fusion proteins: cell surface targets. *Br J Pharmacol* 2012; 166:823-46; PMID:22168282; <http://dx.doi.org/10.1111/j.1476-5381.2011.01811.x>
- Martin PL, Bugelski PJ. Concordance of preclinical and clinical pharmacology and toxicology of monoclonal antibodies and fusion proteins: soluble targets. *Br J Pharmacol* 2012; 166:806-27; PMID:22168335; <http://dx.doi.org/10.1111/j.1476-5381.2011.01812.x>
- Baldrick P. Safety evaluation of biological drugs: what are toxicology studies in primates telling us? *Regul Toxicol Pharmacol* 2011; 59:227-36; PMID:20937341; <http://dx.doi.org/10.1016/j.yrtph.2010.10.005>
- Kooijman M, van Meer PJK, Moors EHM, Schellekens H. Thirty years of preclinical safety evaluation of biopharmaceuticals: Did scientific progress lead to appropriate regulatory guidance? *Expert Opin Drug Saf* 2012; 11:797-801; PMID:22861668; <http://dx.doi.org/10.1517/14740338.2012.712110>
- Suntharalingam G, Perry MR, Ward S, Brett SJ, Castello-Cortes A, Brunner MD, Panoskatsis N. Cytokine storm in a phase 1 trial of the anti-CD28 monoclonal antibody TGN1412. *N Engl J Med* 2006; 355:1018-28; PMID:16908486; <http://dx.doi.org/10.1056/NEJMoa063842>
- Eastwood D, Findlay L, Poole S, Bird C, Wadhwa M, Moore M, Burns C, Thorpe R, Stebbings R. Monoclonal antibody TGN1412 trial failure explained by species differences in CD28 expression on CD4+ effector memory T-cells. *Br J Pharmacol* 2010; 161:512-26; PMID:20880392; <http://dx.doi.org/10.1111/j.1476-5381.2010.00922.x>
- Leulier F, Lemaitre B. Toll-like receptors—taking an evolutionary approach. *Nat Rev Genet* 2008; 9:165-78; PMID:18227810; <http://dx.doi.org/10.1038/nrg2303>
- Ferwerda B, McCall MB, Verheijen K, Kullberg BJ, van der Ven AJ, Van der Meer JW, Netea MG. Functional consequences of toll-like receptor 4 polymorphisms. *Mol Med* 2008; 14:346-52; PMID:18231573; <http://dx.doi.org/10.2119/2007-00135.Ferwerda>
- Vidal JM, Kawabata TT, Thorpe R, Silva-Lima B, Cederbrant K, Poole S, Mueller-Berghaus J, Pallardy M, Van der Laan JW. In vitro cytokine release assays for predicting cytokine release syndrome: the current state-of-the-science. Report of a European Medicines Agency Workshop. *Cytokine* 2010; 51:213-5; PMID:20471854; <http://dx.doi.org/10.1016/j.cyto.2010.04.008>
- Huang L-Y, Dumontelle JL, Zolodz M, Deora A, Mozier NM, Golding B. Use of toll-like receptor assays to detect and identify microbial contaminants in biological products. *J Clin Microbiol* 2009; 47:3427-34; PMID:19726599; <http://dx.doi.org/10.1128/JCM.00373-09>
- Casadevall N, Nataf J, Viron B, Kolta A, Kiladjian JJ, Martin-Dupont P, Michaud P, Papo T, Ugo V, Teyssandier I, et al. Pure red-cell aplasia and antierythropoietin antibodies in patients treated with recombinant erythropoietin. *N Engl J Med* 2002; 346:469-75; PMID:11844847; <http://dx.doi.org/10.1056/NEJMoa011931>
- European Medicines Agency (EMA). Closed Workshop on biosimilar monoclonal antibodies and immunogenicity of monoclonal antibodies. EMA/97951/2012. 24th October 2011, London. http://www.ema.europa.eu/docs/en_GB/document_library/Report/2012/06/WC500128739.pdf
- Santostefano MJ, Kirchner J, Vissinga C, Fort M, Lear S, Pan WJ, Prince PJ, Hensley KM, Tran D, Rock D, et al. Off-target platelet activation in macaques unique to a therapeutic monoclonal antibody. *Toxicol Pathol* 2012; 40:899-917; PMID:22552394; <http://dx.doi.org/10.1177/0192623312444029>
- Everds N, Li N, Bailey K, Fort M, Stevenson R, Jawando R, Salyers K, Jawa V, Narayanan P, Stevens E, et al. Unexpected thrombocytopenia and anemia in cynomolgus monkeys induced by a therapeutic human monoclonal antibody. *Toxicol Pathol* 2013; 41:951-69; PMID:23475561; <http://dx.doi.org/10.1177/0192623312474727>
- Committee for Medicinal Products for Human Use (CHMP). ICH guideline S6 (R1) – preclinical safety evaluation of biotechnology-derived pharmaceuticals. EMA/CHMP/ICH/731268/1998. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500002828.pdf (2011).
- Committee for Medicinal Products for Human Use (CHMP). Guideline on similar biological medicinal products containing biotechnology-derived proteins as active substance: quality issues. Revision 1 (draft) EMEA/CHMP/BWP/247713/2012. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003953.pdf (2012).
- Larocque L, Blui A, Xu R, Diress A, Wang J, Lin R, He R, Girard M, Li X. Bioactivity determination of native and variant forms of therapeutic interferons. *J Biomed Biotechnol* 2011; 2011:174615; PMID:21403871; <http://dx.doi.org/10.1155/2011/174615>
- Radaev S, Sun P. Recognition of immunoglobulins by Fc gamma receptors. *Mol Immunol* 2002; 38:1073-83; PMID:11955599; [http://dx.doi.org/10.1016/S0161-5890\(02\)00036-6](http://dx.doi.org/10.1016/S0161-5890(02)00036-6)
- Shields RL, Lai J, Keck R, O'Connell LY, Hong K, Meng YL, Weikert SH, Presta LG. Lack of fusion on human IgG1 N-linked oligosaccharide improves binding to human Fc gamma RIII and antibody-dependent cellular toxicity. *J Biol Chem* 2002; 277:26733-40; PMID:11986321; <http://dx.doi.org/10.1074/jbc.M202069200>
- Malhotra R, Wormald MR, Rudd PM, Fischer PB, Dwek RA, Sim RB. Glycosylation changes of IgG associated with rheumatoid arthritis can activate complement via the mannose-binding protein. *Nat Med* 1995; 1:237-43; PMID:7585040; <http://dx.doi.org/10.1038/nm0395-237>
- Anthony RM, Ravetch JV. A novel role for the IgG Fc glycan: the anti-inflammatory activity of sialylated IgG Fcs. *J Clin Immunol* 2010; 30(Suppl 1):S9-14; PMID:20480216; <http://dx.doi.org/10.1007/s10875-010-9405-6>
- Chung CH, Mirakhor B, Chan E, Le QT, Berlin J, Morse M, Murphy BA, Satinover SM, Hosien J, Mauro D, et al. Cetuximab-induced anaphylaxis and IgE specific for galactose-alpha-1,3-galactose. *N Engl J Med* 2008; 358:1109-17; PMID:18337601; <http://dx.doi.org/10.1056/NEJMoa074943>
- Johnson R, Jiskoot W. Models for evaluation of relative immunogenic potential of protein particles in biopharmaceutical protein formulations. *J Pharm Sci* 2012; 101:3586-92; PMID:22736238; <http://dx.doi.org/10.1002/jps.23248>
- Brinks V, Jiskoot W, Schellekens H. Immunogenicity of therapeutic proteins: the use of animal models. [Review]. *Pharm Res* 2011; 28:2379-85; PMID:21744171; <http://dx.doi.org/10.1007/s11095-011-0523-5>
- van Beers MM, Sauerborn M, Gilli F, Brinks V, Schellekens H, Jiskoot W. Oxidized and aggregated recombinant human interferon beta is immunogenic in human interferon beta transgenic mice. *Pharm Res* 2011; 28:2393-402; PMID:21544687; <http://dx.doi.org/10.1007/s11095-011-0451-4>