FINAL POST-ECBS version
ENGLISH ONLY

Guidelines on evaluation of biosimilars

Replacement of Annex 2 of WHO Technical Report Series, No. 977

Adopted by the Seventy-fifth meeting of the World Health Organization Expert Committee on Biological Standardization, 4–8 April 2022. This is the final edited version which will be published in the WHO Technical Report Series.

Annex 3

Guidelines on evaluation of biosimilars

Replacement of Annex 2 of WHO Technical Report Series, No. 977

Guidelines published by the World Health Organization (WHO) are intended to be scientific and advisory in nature. Each of the following sections constitutes guidance for national regulatory authorities (NRAs) and for manufacturers of biological products. If an NRA so desires, these WHO Guidelines may be adopted as definitive national requirements, or modifications may be justified and made by the NRA. It is recommended that modifications to these Guidelines are made only on condition that such modifications ensure that the product is at least as safe and efficacious as that prepared in accordance with the guidance set out below.

Abbreviations

ADA	anti-drug antibody
ADCC	antibody-dependent cellular cytotoxicity
ADCP	antibody-dependent cellular phagocytosis
CDC	complement-dependent cytotoxicity
C1q	complement component 1q
Fab	antigen-binding fragment
Fc	fragment crystallizable
FIIa	activated blood coagulation factor II
FXa	activated blood coagulation factor X
G-CSF	granulocyte-colony stimulating factor
ICH	International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use
Ig	immunoglobulin
INN	international nonproprietary name
IS	international standard(s)
IU	International Unit(s)
mAb	monoclonal antibody
NRA	national regulatory authority
PD	pharmacodynamic(s)
PK	pharmacokinetic(s)
RP	reference product
SD	standard deviation
TNF	tumour necrosis factor

1. Introduction

Biotherapeutic products (biotherapeutics) have a successful record in treating many life-threatening and chronic diseases. The expiry of patents and/or data protection periods for a number of such biotherapeutics has ushered in an era of products that are designed to be highly "similar" to the corresponding licensed "originator" product. Based on a comprehensive head-to-head comparison and demonstrated high similarity, such products can partly rely for their licensing on safety and efficacy data obtained for the originator products. A variety of terms have been used to describe these products, including "biosimilars", "similar biotherapeutic products", "similar biological medicinal products" and "biosimilar products" *(1)*.

The term "generic medicine" is usually used to describe chemical, small-molecule medicinal products that are structurally identical to an originator product whose patent and/or data protection period has expired. Demonstration of the analytical sameness and bioequivalence of the generic medicine to a reference product is usually appropriate and sufficient proof of therapeutic equivalence between the two. However, the approach established for generic medicines is not suitable for the development, evaluation and licensing of relatively large and complex proteins such as biosimilars.

As part of its mandate to assure the global quality, safety and efficacy of biotherapeutics, WHO provides globally accepted norms and standards for their evaluation. WHO written standards adopted on the recommendation of the WHO Expert Committee on Biological Standardization serve as a basis for setting national requirements for the production, quality control and overall regulation of biological medicines. In addition, WHO international measurement standards established by the Committee are essential for assessing the potency of biological medicines worldwide.

By 2007 an increasingly wide range of biosimilars were under development or were already licensed in many countries and a need for guidance on their evaluation and overall regulation was formally recognized by WHO *(2)*. In 2009, the WHO Guidelines on evaluation of similar biotherapeutic products (SBPs) were adopted on the recommendation of the Committee *(3)*. This document provided the scientific principles and stepwise approach to be applied during the demonstration of similarity between a similar biotherapeutic product and its reference biotherapeutic product. The document also provided guidance on the development and evaluation of such biotherapeutics; it was however viewed as a "living" document that would be further developed in line with advances in scientific knowledge and experience. It was also anticipated that the increasing availability of biosimilars worldwide would lead to increased competition between manufacturers, thus bringing down prices and improving access to such products.

In line with World Health Assembly resolution WHA67.21 on access to biotherapeutics *(4)*, the Committee at its meeting in October 2020 recommended that a review should be undertaken of current scientific evidence and experience in this field to inform the updating and revision of the 2009 WHO Guidelines. This review would provide an opportunity to evaluate new developments and identify areas where the current guidance could be more flexible without compromising its basic principles, and allow for the provision of additional explanation of the possibility of tailoring the amount of data needed for regulatory approval *(5)*. At its subsequent meeting in December 2020 the Committee was informed that the review had taken into account a range of national and regional guidelines, and a number of sections in the 2009 WHO Guidelines had been identified for potential updating and revision *(6)*. Having been updated on progress in this area, the Committee expressed the opinion that the review of existing national and regional guidance had been comprehensive and indicated its support for the continuation of the proposed revision process *(5)*. It was intended that the revision of the

2009 WHO Guidelines would result in greater flexibility and reduced regulatory burden, while continuing to ensure the quality, safety and efficacy of such products.

Following international consultations *(7)*, the present document represents the outcome of the above revision process and replaces Annex 2 of WHO Technical Report Series, No. 977 *(3)*. The main changes made include:

- Updating the Introduction to reflect the discussions held on the revision process.
- Expanding the scope of the document to include the evaluation of biological products other than biotherapeutics and a corresponding shift to the use of the term "biosimilar" rather than "similar biotherapeutic product".
- Use of the term "reference product (RP)" rather than "reference biotherapeutic product (RBP)" and updating of the considerations regarding the use of non-local RPs.
- Extensively revising the sections on quality, and nonclinical and clinical evaluation to make them more consistent with current practices, and with other guidelines, as well as to provide more clarity and flexibility – specific topics addressed include but are not limited to:
  - the use of WHO international standards and reference reagents;
  - analytical considerations in quality evaluation;
  - considerations in establishing similarity ranges for quality comparisons, and in determining similarity;
  - new guidance on determining the need for in vivo animal studies and on the implementation of the 3Rs principles ("Replace, Reduce, Refine") to minimize the use of animals in testing; and
  - consideration of the amount and type of clinical data required.
- Updating the sections on pharmacovigilance, prescribing information and label, and the role and responsibilities of national regulatory authorities (NRAs) with additional details and references.

For public health purposes, it is essential that the standard of evidence supporting the decision to license a biosimilar is sufficiently high to ensure that the product meets acceptable levels of quality, safety and efficacy. Elaboration of the data requirements and considerations for the licensing of such products is expected to facilitate the development of and worldwide access to biological products of assured quality, safety and efficacy at more affordable prices. It is expected that these WHO Guidelines on the scientific principles for evaluating biosimilars will help to harmonize global requirements, and lead to easier and speedier approval and assurance of the quality, safety and efficacy of these products. It is important to note that biological products that are not shown to be similar to an RP as set out in these Guidelines should not be described as "similar" and should not be termed "biosimilars".

It is recognized that a number of important issues associated with the use of biosimilars, including but not limited to the following, need to be defined by the individual NRA:

- intellectual property issues;
- interchangeability modalities, including switching (physician-led) and substitution (pharmacy-led) of an originator product with a biosimilar; and
- labelling and prescribing information.

For this reason, these issues are not elaborated upon in detail in this document.

2. Purpose and scope

These WHO Guidelines are intended to provide globally acceptable principles for the licensing of biological products that are claimed to be similar to biological products of assured quality, safety and efficacy that have been licensed based on a full licensing dossier. On the basis of proven similarity, the licensing of a biosimilar would in part rely on nonclinical and clinical data generated for an already licensed originator product. These Guidelines can be adopted by NRAs worldwide or used as a basis for establishing national regulatory frameworks for the licensure of such biosimilars.

The Guidelines apply to biological products that can be well characterized, such as recombinant DNA-derived therapeutic peptides and proteins *(8)*. Some of the principles provided in these Guidelines may also apply to low molecular weight heparins and recombinant analogues of plasma-derived products. Vaccines and plasma-derived products are excluded from the scope of these Guidelines.

3. Terminology

The definitions given below apply to the terms as used in these Guidelines. These terms may have different meanings in other contexts.

Biosimilar: a biological product that is shown to be highly similar in terms of its quality, safety and efficacy to an already licensed **reference product**.

Comparability/similarity exercise: direct head-to-head comparison of a biological product with a licensed **reference product** with the goal of establishing similarity in quality, safety and efficacy.

Comparability margin: the largest difference that can be judged as being clinically acceptable.

Comparability/similarity range: predefined allowable differences in physicochemical properties and biological activity level.

Drug product: a pharmaceutical product that typically consists of a **drug substance** formulated with **excipients**.

Drug substance: the active pharmaceutical ingredient and associated molecules that are typically formulated with excipients to produce the **drug product**. This may also be referred to as the "active substance" in other documents.

Efficacy study: a clinical trial to compare the efficacy of the biosimilar to the reference product.

Excipient: a constituent of a medicine other than the **drug substance**, added in the formulation for a specific purpose. While most excipients are considered inactive, some can have a known action or effect in certain circumstances (for example, hyaluronidase). The excipients may differ for a biosimilar and its reference product and need to be declared in the labelling and package leaflet of the medicine to ensure its safe use.

Equivalent: equal or highly similar in the parameter of interest. Equivalent quality, safety and efficacy of two medicinal products denotes that they can be expected to have similar (no better and no worse) quality, safety and efficacy, and that any observed differences are of no clinical relevance.

Generic medicine: a medicine that is structurally identical to an **originator product** (comparator) for which the patent and/or data protection period has expired.

Head-to-head comparison: direct comparison of the properties of a biosimilar with its corresponding **reference product**. Comparison based on historical data is not acceptable.

Immunogenicity: the ability of a substance to trigger an immune response or reaction (for example, development of specific antibodies, T-cell response, or allergic or anaphylactic reaction).

Impurity: any component present in the **drug substance** or **drug product** that is not the desired product, a product-related substance or **excipient** (including buffer components). Impurities may be either process or product related.

Marketing authorization holder: any person or legal entity that has received a marketing authorization or licence to manufacture and/or distribute a medicine. It also refers to a person or legal entity allowed to apply for a change to the marketing authorization or licence. Under the same licence, the marketing authorization holder could have several manufacturing sites registered. Therefore, several manufacturers could be involved.

Non-inferior: not clinically inferior to a comparator in the parameter studied. A non-inferiority clinical trial is one that has the primary objective of showing that the response to the investigational product is not clinically inferior to that of a comparator within a pre-specified margin.

Originator product: a medicine that has been licensed by an NRA on the basis of a full registration dossier – that is, the approved indication(s) for use were granted on the basis of full quality, efficacy and safety data.

Pharmacodynamic study: a clinical study that measures a pharmacodynamic (PD) response that effectively demonstrates the characteristics of the product's target effects. PD biomarkers for biosimilars do not need to be surrogate end-points for clinical efficacy outcomes.

Pharmacovigilance: the science and activities relating to the detection, assessment, understanding and prevention of adverse effects caused by medical drugs.

Posology: dosage for each indication and each method/route of administration. Information includes dose recommendation (for example, in mg, mg/kg or mg/m^2), frequency of dosing (for example, once or twice daily, or every 6 hours) and treatment duration.

Reference product (RP): a biological product used as the comparator in a direct head-to-head **comparability exercise** with a **biosimilar** in order to demonstrate similarity in terms of quality, safety and efficacy. Only an **originator product** licensed on the basis of a full registration dossier and marketed for a suitable period of time with proven quality, safety and efficacy can serve as an RP.

Reference standard: a measurement standard such as an international, pharmacopoeial or national standard – it should be noted that reference standards are distinct from reference products and serve a different function.

Similarity: absence of any relevant difference in the parameter(s) of interest.

4. Scientific considerations and concept for licensing biosimilars

The regulatory framework for the licensing of generic medicines is well established in most countries. Demonstration of structural sameness and bioequivalence of the generic medicine and the reference product (RP) is usually sufficient for therapeutic equivalence between the generic and reference products to be inferred. However, the generic approach is not suitable for the licensing of biosimilars since biological products usually consist of relatively large and complex proteins that are more complicated to characterize and manufacture than small molecules.

Characterization and evaluation of the quality attributes of the RP should be the first step in guiding the development of the biosimilar. This is followed by a comparability exercise applying sensitive orthogonal analytical methods and assays to demonstrate structural, functional and clinical similarity. Comprehensive characterization and comparison showing similarity at the quality and nonclinical (in vitro) level are the basis for establishing

comparability, with a tailored confirmatory clinical data package required for licensure. If differences between the biosimilar and the RP are found, the underlying reasons for them should be investigated. Unless such differences are explained and justified in terms of lack of clinical impact, additional data (for example, on safety) may be required. The standalone development of biological products is not discussed in the current Guidelines.

In addition to quality and nonclinical (in vitro) data, clinical data are generally required for any biosimilar. The type and amount of such data considered to be necessary will depend on the product or class of products, on the extent of characterization possible using state-of-the-art analytical methods, on observed or potential differences between the biosimilar and the RP, and on clinical experience with the RP (for example, safety/immunogenicity concerns in a specific indication). A case-by-case approach will be needed for each class of products.

A biosimilar is intended to be highly similar to a licensed biological product for which substantial evidence exists of its safety and efficacy. Manufacturers should demonstrate both a full understanding of their product and consistent and robust manufacture, and should submit a full quality dossier that includes a complete characterization of the product. Comparison of the biosimilar and the RP with respect to quality represents an additional element to the "traditional" full quality dossier. Such comparison will include a comprehensive comparison of biological function at the in vitro level. A reduction in data requirements is therefore possible for the nonclinical in vivo and/or clinical parts of the development programme. The posology and route of administration of the biosimilar should be the same as for the RP.

Studies must be comparative in nature and must employ state-of-the-art analytical methods capable of detecting potential differences between the biosimilar and the RP. The main clinical studies should use the final formulation of the biosimilar (that is, derived from the final process material); if not, then additional evidence will be required to demonstrate that the biosimilar to be marketed is comparable to that used in the main clinical studies *(9)*.

If similarity between the biosimilar and the RP has been demonstrated, the biosimilar may be approved for all clinical indications of the RP supported by appropriate scientific data and justification (see section 9.7).

5. Key principles for the licensing of biosimilars

- Characterization of the quality attributes of the RP should be the first step in guiding the development of the biosimilar. The subsequent comparability exercise should demonstrate structural, functional and clinical similarity.
- Demonstration of similarity of a biosimilar to an RP in terms of structural and functional aspects is a prerequisite for establishing comparability, with a tailored clinical data package required as needed.
- A clinical bioequivalence trial with pharmacokinetic (PK) and pharmacodynamic (PD) parameters (if available), and including an assessment of immunogenicity in human subjects, will typically be a core part of the clinical comparability assessment, unless scientifically justified.
- The decision to license a biosimilar should be based on evaluation of the whole data package generated during the overall comparability exercise.
- If relevant differences between the proposed biosimilar and the RP are found at the structural, functional or clinical level, the product is unlikely to qualify as a biosimilar.
- If comparability exercises are not performed as outlined in this document then the final product should not be referred to as a biosimilar.
- Biosimilars are not "generic medicines" and the authorization process for such medicines generally does not apply.

- As with other biological products, biosimilars require effective regulatory oversight pre- and post-approval in order to manage the potential risks they pose and to maximize their benefits.

6. Reference products

Comprehensive information on the reference product (RP) provides the basis for establishing the quality, safety and efficacy profile against which the biosimilar will be compared. The RP also provides the basis for dose selection and route of administration, and is used in the similarity studies required to support the licensing application. Demonstration of a high level of analytical and functional similarity between the biosimilar and RP provides the rationale for a tailored nonclinical and clinical dataset to support the application for market authorization of the biosimilar.

The choice of RP is therefore critically important in the evaluation of a biosimilar. For licensing purposes for a specific biosimilar, a single biological product from one marketing authorization holder should be chosen and defined as the RP.

Traditionally, NRAs have required the use of a nationally licensed RP for the licensing of a generic medicine. In the case of biosimilars, this practice may not always be feasible or necessary, and several regulatory jurisdictions have allowed for the use of a non-local RP as comparator to enable faster development of and access to biological therapies. The use of an RP sourced from another jurisdiction with similar scientific and regulatory standards is therefore possible. The information needed to support the acceptability of an RP sourced from another jurisdiction will be determined by the NRA.

The posology and route of administration of the biosimilar should be the same as that of the RP. However, depending on the jurisdiction, the strength, pharmaceutical form, formulation, excipients and presentation (for example, use of a different medical device or number of syringes in a pack) of the biosimilar might differ from the RP, if justified. The acceptability of additional routes of administration following approval of the biosimilar will also depend upon the jurisdiction.

Since the choice of RP is crucial in the development of a biosimilar, the following should be considered:

- The RP should have been licensed on the basis of a full standalone set of quality, nonclinical, safety and efficacy data *(8)*. A biosimilar should therefore not be accepted as an RP.
- There should be sufficient information available to support the safe and efficacious use of the RP.
- For the licensing of a specific biosimilar, a single biological product from one marketing authorization holder should be chosen and defined as the RP. The entire comparability exercise should be performed against this RP. However, as outlined below, if allowed by the NRA it may be possible to use the same RP sourced from another jurisdiction in clinical studies.
- Where an RP marketed in another jurisdiction (non-local) is allowed by the NRA, the following should be considered:
  - the RP should be licensed in a jurisdiction that has a well-established regulatory framework, as well as experience with the evaluation of biological products and post-marketing surveillance activities; and
  - if the use of a non-local RP containing the same drug substance in clinical studies requires bridging between the local and non-local RPs, suitable analytical and functional bridging data should be provided to demonstrate the

representativeness of the non-local RP for the local RP – stringent similarity assessment should be applied for the analytical and functional bridging studies (following the principles provided in sections 7.3 and 7.4 below); additional PK bridging studies may be required, for example if the two RPs have different formulations that may affect PK.

- It is important to note that the acceptance of a non-local RP for the evaluation of a biosimilar in a particular country does not imply that the NRA of that country has approved the RP for use in the domestic market.

7. Quality

The comparison showing molecular similarity between the biosimilar and the RP provides the essential rationale for predicting that the clinical safety and efficacy profiles of the RP apply to the biosimilar. Therefore, a high degree of analytical and functional similarity between the biosimilar and the RP is the basis for developing a biosimilar.

Development of a biosimilar involves the thorough characterization of multiple RP batches in order to obtain an understanding of the overall quality profile as well as the range of variability of the RP batches on the market. Based on the knowledge gained from the RP characterization studies, as well as available in-house and public information, the manufacturing process of the biosimilar is developed to produce a product that is highly similar to the RP in all clinically relevant quality attributes (that is, attributes that may impact clinical performance).

The biosimilar documentation should comply with the standards required by NRAs for originator products. A full quality dossier for both drug substance and drug product is always required – see relevant guidelines for each class of product, such as those issued by the International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) and the WHO Guidelines on the quality, safety and efficacy of biotherapeutic protein products prepared by recombinant DNA technology *(8)* and the WHO Guidelines on evaluation of monoclonal antibodies as similar biotherapeutic products (SBPs) *(10)*. The manufacturer of the biosimilar should additionally carry out a comprehensive and comparative state-of-the-art physicochemical and biological characterization of the biosimilar and the RP and document the results in the submitted dossier.

7.1 International reference standards

WHO provides international standards (IS) and reference reagents, which serve as reference sources of defined biological activity expressed in International Units (IU) or Units (U). These materials are intended for use in the calibration of bioassays and are available for a wide range of substances including hormones (for example, erythropoietin, follicle-stimulating hormone) and cytokines – for example, granulocyte-colony stimulating factor (G-CSF) – as well as modified/long-acting proteins (such as pegylated G-CSF, darbepoetin and etanercept) and monoclonal antibodies (mAbs). IS for the latter product class are expanding and currently include standards for adalimumab, bevacizumab, infliximab, rituximab and trastuzumab.[1] These standards are produced according to defined criteria as per WHO recommendations *(11)* and often contain excipients which optimize the retention of biological activity and other important characteristics as well as ensuring stability, but which may also interfere with physicochemical methods. The standards are important for assay development, for qualifying

[1] For the full range of available WHO international biological reference materials please see: https://www.who.int/activities/providing-international-biological-reference-preparations

and validating assays for their intended use, for monitoring the potency of individual/diverse products, for calibrating bioassays (either directly or to calibrate national or pharmacopoeial standards) and for supporting assay performance throughout the life-cycle of a product. In addition, they can be used for the independent testing of falsified medicines and as an independent standard for tracing bioactivity horizontally (between products and batches) and longitudinally (over time) to support post-marketing surveillance activities and to assess any divergence that may occur as the product(s) evolve.

For biological medicines, expression of potency in units of bioactivity relative to an independent standard is an essential regulatory tool for harmonizing product dosing for patients globally. For many years, WHO IS have provided a mechanism for assigning and maintaining biological potency across diverse products. It should be noted however that with the development of innovative products the role of the IS in potency determination is changing and decisions on potency and labelling are likely to be made on a case-by-case basis depending on the product and the situation that exists when the biosimilar is developed. For example, for naturally derived proteins such as coagulation factors and hormones (for example, erythropoietin and follicle-stimulating hormone), where the establishment of the IS with an assigned IU preceded the development of versions derived from recombinant DNA (rDNA), the practice of using the IU for potency assignment, dosage and product labelling is well established, and where applicable this has continued for biosimilars. However, the situation is different for non-natural and engineered proteins such as mAbs. Since IS did not exist when the innovator products were developed, such products were licensed and marketed for clinical use with potency described by manufacturers in proprietary units relative to their in-house product-specific reference material, with product dosing and labelling given in mass units. The practice of determining potency relative to an in-house qualified reference material and of using mass units for dosing/labelling has also been implemented by biosimilar manufacturers and is expected to continue. In this situation, manufacturers should develop a well-characterized product-specific in-house reference material calibrated against the IS (where this exists) with a regulatory expectation that the implementation and management of this in-house reference material (two-tiered approach) will be conducted as per regulatory guidance. Consistent with the biosimilarity paradigm, the retrospective establishment of an IU value should not affect the potency of the biosimilar (which should be aligned with the RP) and should not affect the labelling or dosing regimens of existing or future products.

It is important to note that WHO IS and other WHO reference standards are not medicinal products (even though the drug substance in them may be derived from material that was produced at clinical grade) and are distinct (for example, in terms of protein content, formulation etc.) from the RP which has an established clinical history and is an essential component of the biosimilarity route to licensure. The RP defines the quality target product profile that a biosimilar must meet as per the principles of biosimilarity – a function that the reference standard does not serve. Instead, the IS defines the IU of bioactivity for the calibration of bioassays (either directly or through the calibration of manufacturer reference materials) and thus plays an essential role in the development of suitable assay methods. It should further be noted that the IS cannot be used to determine a product's specific activity, dictate the quality of acceptable biosimilars for regulatory purposes or demonstrate biosimilarity, and should therefore not be misused as a comparator for biosimilar development *(12–14)*. Importantly, the IS: (a) allows for an understanding of consistency in bioactivity across batches of a product throughout its life-cycle; (b) provides continuity with respect to the in-house reference material and supports transition (change) as the product evolves; (c) facilitates the harmonization of bioactivity across different products (both RPs and biosimilars); and (d) increases confidence in the quality of globally available biosimilars.

7.2 Manufacturing process

The manufacturing process of the biosimilar should be developed based on a comprehensive understanding of the RP gained through detailed characterization studies of a sufficient number of RP batches.

It is understood that a manufacturer developing a biosimilar will not normally have access to confidential details of the RP manufacturing process – thus, the process will differ from the licensed process for the RP. In order to produce a high-quality product as similar as possible to the RP, the biosimilar manufacturer should assemble all available knowledge on the RP regarding the type of host cell, product formulation and the container closure system used for marketing. Although the biosimilar does not need to be expressed in the same type of host cell as that used for the RP, it is recommended that a similar host cell type is used (for example, Escherichia coli, Chinese hamster ovary cells, etc.). This will reduce the potential for critical changes in the quality attributes of the protein, or in post-translational modifications, product-related impurities or the process-related impurity profile, that could potentially affect clinical outcomes and immunogenicity. If a different host cell is used (for example to avoid unwanted and potentially immunogenic glycan structures present in the RP) then changes introduced in terms of product-related substances, as well as product- and process-related impurities, need to be carefully considered.

The manufacturing process used can significantly affect the structure of the drug substance and thereby impact upon the potency of the product. For example, in the case of mAbs, when deciding upon the expression system to employ, manufacturers should be guided by the potential for both enzymatic and non-enzymatic modifications, such as incomplete disulfide bond formation, formation of aggregates, glycosylation, N-terminal pyroglutamine cyclization, C-terminal lysine processing, deamidation, isomerization and oxidation, modification of the N-terminal amino acids by maleuric acid, and amidation of the C-terminal amino acid.

The manufacturer must demonstrate the consistency and robustness of the manufacturing process by implementing state-of-the-art quality control and assurance procedures, in-process controls and process validation. The biosimilar manufacturing process should meet the same standards required for originator products, including manufacture under current good manufacturing practices *(15, 16)*.

As for any biological product, if process changes are introduced during the development of a biosimilar then the impact of the changes should be assessed through a comparability exercise *(9, 17)*. Although many of the same principles are followed, the assessment of manufacturing process changes should be addressed separately from the comparability exercise performed to demonstrate biosimilarity with the RP (see section 7.4 below). It is, however, strongly recommended that the pivotal data used to demonstrate biosimilarity are generated using biosimilar batches manufactured using the commercial manufacturing process and therefore representing the quality profile of the batches to be commercialized.

7.3 Analytical considerations

Thorough characterization of both the RP and the biosimilar should be carried out using state-of-the-art chemical, biochemical, biophysical and biological analytical techniques. The methods should be scientifically sound and demonstrated to be of appropriate sensitivity and specificity for their intended use.

Details should be provided on primary and higher-order structure, post-translational modifications (including, but not limited to, glycoforms), biological activity, purity, impurities, product-related (active) substances (variants) and immunochemical properties, where relevant.

Orthogonal methods should be used, as far as possible – that is, the variants and quality attributes of the product should be analysed using analytical methods with different underlying chemical, physical and biological properties. For example, ion exchange chromatography, isoelectric focusing and capillary electrophoresis all separate proteins based upon charge but do so under different analytical conditions and on the basis of different physicochemical properties of the biological product. As a result, one method may detect variants that another method does not. The goal of the comparability investigation is to be as comprehensive as possible in order to minimize the possibility of undetected differences between the RP and the biosimilar that may affect safety and clinical activity. The analytical limitations of each technique (for example, limit of detection or resolving power) should be considered when determining the similarity of a biosimilar to its RP.

Representative raw data should be provided for analytical methods (for example, high-quality reproductions of gels and chromatograms) in addition to tabular data summarizing the complete dataset and showing the results of all release and characterization analyses carried out on the biosimilar and the RP. Graphical presentation of datasets comparing biosimilar and RP analytical data should also be produced where possible. The results should be accompanied by sufficient interpretation and discussion of the findings.

The measurement of quality attributes in characterization studies (as opposed to batch release tests) does not necessarily require the use of validated assays, but the assays used should be scientifically sound and qualified – that is, they should provide results that are meaningful and reliable. The methods used to measure quality attributes for batch release should be validated in accordance with relevant guidelines, as appropriate. A complete description of the analytical techniques employed for release and characterization of the product, along with method validation or qualification data (as appropriate), should be provided in the licence application.

Due to the unavailability of drug substance for the RP, the biosimilar manufacturer will usually be using a commercial drug product for the similarity exercise. The commercial drug product will, by definition, be in the final dosage form containing the drug substance(s) formulated with excipients. It should be verified that these excipients do not interfere with the analytical methods used and thus have no impact on test results. If the drug substance in the RP needs to be purified from a formulated reference drug product in order to be suitable for characterization then studies must be carried out to demonstrate that product heterogeneity and relevant attributes of the active moiety are not affected by the isolation process. The approach used for isolating the drug substance of the RP and comparing it with the biosimilar should be justified and demonstrated (with accompanying data) to be appropriate for the intended purpose.

7.3.1 Physicochemical properties

The physicochemical characterization should include determination of primary and higher-order structure (secondary/tertiary/quaternary) and product variants using appropriate analytical methods (for example, mass spectrometry, circular dichroism, spectroscopy etc.) as well as other biophysical properties.

The amino acid sequence of a biosimilar should be confirmed to be the same as that of its RP. It is, however, further recommended that manufacturers pay special attention to any sequence variants present in the biosimilar. Although an identical primary sequence between the biosimilar and the RP is expected, low-level sequence variants may occur due to transcription and translation errors, especially through amino acid misincorporation during high-level expression, and should be identified if present. The presence of such variants could be acceptable if properly described and controlled to a reasonable level. An assessment of the potential clinical impact of such variants would also need to be considered.

An inherent degree of structural heterogeneity occurs in proteins as a result of biosynthesis processes. These include C-terminal processing, N-terminal pyroglutamation, deamidation, oxidation, isomerization, fragmentation, disulfide bond mismatch and free sulfhydryl groups, N-linked and O-linked oligosaccharide, glycation and aggregation. The structural heterogeneity present in the biosimilar should be evaluated relative to the RP. Experimentally determined disulfide bonding patterns should be compared to the predicted structure based on well-established structural data on the molecule.

7.3.2 Biological activity

Biological activity is the specific ability or capacity of the product to achieve a defined biological effect. It serves multiple purposes in the assessment of product quality and is required for characterization (see also section 8 below) and for batch analysis. Ideally, the biological assay used will reflect the understood mechanism of action of the drug substance of the RP and will thus serve as a link to clinical activity. A biological assay is a quality measure of the activity of the drug substance and can be used to determine whether a product variant is active (that is, a product-related substance) or inactive (and therefore defined as an impurity). Biological assays can also be used to confirm that small differences observed in the higher-order structure of a molecule have no influence on its biological activity. Thus, the use of relevant biological assay(s) of appropriate precision, accuracy and sensitivity provides an important means of confirming that there is no significant functional difference between the biosimilar and the RP.

For a product with multiple biological activities, manufacturers should perform, as part of product characterization, a set of relevant functional assays designed to evaluate the range of activities of the product. For example, certain proteins possess multiple functional domains that express enzymatic and receptor-binding activities. In such situations, manufacturers should evaluate and compare all relevant functional activities of the biosimilar and the RP.

Potency is the measure of the biological activity. The potency assay should be used together with an in-house qualified reference material that is representative of the biosimilar material. The use of the IS for determining potency depends on the prevailing practice for the product. Where appropriate, international or national standards and reference reagents should be used to determine product potency and to express results in IU – for other products, a suitable in-house reference material should be used (see section 7.1 above). In-house reference materials should be quantitatively calibrated against either an international or national standard or reference reagent, where available and appropriate.

Depending on the purpose of the method (batch release assay or characterization), the functional assays used may or may not be fully validated, but they must be scientifically sound and produce consistent and reliable results. The available information on these assays (including extent of validation, assessed parameters and available validation data) should be confirmed before they are applied to the testing and establishing of biosimilarity between a biosimilar and its RP. It should be noted that many biological assays may have relatively high variability that might preclude detection of small but significant differences between the biosimilar and RP. Therefore, it is recommended that assays are developed that are more precise and can detect changes in the intended biological activities of the product to be evaluated with adequate accuracy. Such assays can include target-binding assays (which are usually less variable) in addition to cell-based assays. Adopting automated laboratory equipment to help minimize manual operations, applying good analytical practices and appropriate control sampling, and using critical reagents calibrated against WHO or national reference standards where available (for example, tumour necrosis factor alpha (TNF-α) for potency assays for anti-TNF products) may help to reduce the variability of biological assays.

For a given method variability, the number of RP batches tested should be high enough to allow for a reliable assessment of similarity (see section 7.4.1 below).

When immunochemical properties are part of the activity attributed to the product (for example, antibodies or antibody-based products) analytical tests should be performed to characterize these properties and used in the comparative studies. For mAbs, the specificity, affinity and binding kinetics of the product to relevant fragment crystallizable (Fc) receptors (for example, neonatal Fc receptor, complement component 1q (C1q) and Fcγ receptors) should be compared using suitable methods such as surface plasmon resonance and biolayer interferometry. In addition, appropriate assays should be used to provide information on Fc-mediated functions – for example, antibody-dependent cellular cytotoxicity (ADCC), antibody-dependent cellular phagocytosis (ADCP) and complement-dependent cytotoxicity (CDC), where relevant.

The correlation between Fc-mediated effector functions, Fcγ receptor or C1q binding and physicochemical characteristics (for example, glycan pattern) should be considered and, whenever possible, established. Such analyses will facilitate the interpretation of subtle differences between the biosimilar and the RP and inform prediction of their clinical impact.

7.3.3 Purity and impurities

Product- and process-related impurities should be identified and quantified using orthogonal and state-of-the-art technologies.

Product-related substances and impurities, such as those caused by protein degradation, oxidation, deamidation, aggregation or potential post-translational modification of the protein, should be compared for the biosimilar and RP. If comparison reveals differences in product-related substances and impurities between the biosimilar and RP, the impact of the differences on the clinical performance of the drug product (including its biological activity) should be evaluated. Specifically, if the manufacturing process used to produce the proposed biosimilar introduces different impurities or higher levels of impurities than those present in the RP then additional functional assays to evaluate the impact of the differences may be necessary (see section 7.4.2 below). To obtain sufficient information of the product-related substances and impurities it is recommended that comparative stability studies under accelerated and/or stress conditions are conducted (see section 7.6 below).

Process-related impurities such as host cell proteins, host cell DNA, cell culture residues and downstream processing residues may be quantitatively and/or qualitatively different between the biosimilar and RP due to the different manufacturing processes used for their drug products. Nevertheless, process-related impurities should be kept to a minimum through the use of state-of the-art manufacturing technologies. The risk related to any newly identified impurities in the biosimilar should be evaluated.

7.3.4 Quantity

In general, a biosimilar is expected to have the same concentration or strength of the drug substance as the RP. Depending on the jurisdiction, concentration deviations not affecting the posology might be permissible, if justified (see section 8 below). The quantity of the biosimilar drug substance should be expressed using the same measurement system as that used for the RP (that is, mass units or units of activity). A description with appropriate justification should also be included to describe how the quantity was calculated (including, for example, the selection of the extinction coefficient).

7.4 Comparative analytical assessment

7.4.1 Considerations for the RP and the biosimilar

The number of RP batches needed for the comparative analytical assessment will be influenced by the criticality of the quality attribute(s) under investigation and the approach chosen for demonstrating similarity. The manufacturer of the biosimilar should include an appropriate and scientifically supportable number of batches of the RP in the comparability assessment. In order to characterize independent RP batches, it is recommended that the RP batches are sourced over an extended time period. These batches should also include the RP batches used in the clinical comparison studies of the biosimilar. In general, sampling a higher number of RP batches will provide a better estimate of the true batch-to-batch variability of the RP and allow for a more robust statistical comparison with the biosimilar.

Random sampling of RP batches is desirable but may be difficult to achieve in practice depending on the availability of such batches. However, the sourcing of RP batches should be carefully managed to generate a sample that captures the inherent variability of the RP (for example, collected over a sufficient timeframe with the aim of covering different manufacturing campaigns). The RP batches should be transported and stored under the recommended conditions and tested within their approved shelf-life. Any exception to this would have to be fully substantiated with experimental data. The shelf-life of the RP at time of characterization should be considered and it is expected that RP batches of different ages will be included in the similarity assessment.

The biosimilar batches included in the comparability assessment should be manufactured using the intended commercial manufacturing process and should preferably originate from different drug substance batches. Generally, each value for an attribute being assessed for a biosimilar should be contributed by an independent batch. For example, a single drug product batch produced from a single drug substance batch would be considered to be an independent batch while different drug product batches produced from the same drug substance batch cannot be considered to be independent. In addition, small- or pilot-scale batches can be included if comparability between the small- and commercial-scale batches has been properly demonstrated. Usually all commercial-scale batches produced – including process performance qualification batches and batches applied in the clinical trial(s) – should be included in the similarity assessment. As with the RP, the exact number of biosimilar batches required will be influenced by several factors, such as the criticality of the quality attribute(s) under investigation and the approach applied for similarity evaluation. In general, the risk of a false-positive conclusion on similarity will decrease with increasing number of batches. A robust manufacturing control system and demonstrated batch-to-batch consistency of the biosimilar (see section 7.2 above) are prerequisites for a successful similarity assessment.

7.4.2 Considerations for similarity assessment

Prior to initiating the comparability exercise, it is recommended that the quality attributes of the RP are identified and ranked according to their impact on the clinical performance of the product. For this purpose, a risk ranking tool could be developed. Such risk ranking tools should consider the impact of the quality attribute on safety, efficacy, PK and immunogenicity. Furthermore, the degree of uncertainty of impact should be taken into consideration. If it is known that a quality attribute will impact the clinical performance (that is, the uncertainty is low but the impact high) then that quality attribute should be prioritized and the overall risk score should be high. In cases where the clinical relevance of a certain quality attribute is unknown (that is, the uncertainty is high) then higher risk scores should be assigned even to lower impact quality attributes. Further guidance on the use of risk ranking tools can be found in national and international guidelines *(18)*.

The result of the risk ranking could then be used to guide the data analyses and the overall assessment of similarity. The most frequently used approach for similarity assessment relies on demonstrating that the quality attributes of the biosimilar batches lie within the predetermined similarity ranges established based on characterization data from multiple batches of the RP. Other approaches (such as equivalence testing of means) can also be used for similarity assessment. Each statistical approach has, however, specific strengths and weaknesses which should be appropriately discussed in the submission and considered in the similarity conclusion. In order to mitigate the risks inherent in employing statistical tests on limited samples (false-positive and false-negative conclusions), a comprehensive control strategy must be established for the biosimilar to ensure consistent manufacturing.

7.4.2.1 Statistical intervals for the establishment of similarity ranges

Where possible, quantitative similarity ranges should be established for the biosimilar comparability exercise. As the allowable differences in quality attributes between the biosimilar and the RP are usually difficult to establish based on clinical considerations alone, the batch-to-batch variability of the RP is typically used to further inform acceptable differences in quality attributes. The established similarity range should therefore tightly reflect the quality profile of the marketed RP batches. The ranges should normally not be wider than the batch-to-batch variability present in the RP unless it can be determined which differences would be acceptable (for example, less impurities is usually acceptable). Wide similarity ranges based on inappropriate use of statistical methods should not be used.

Different statistical intervals can be used to establish similarity ranges. Commonly used approaches include mean ± x SD, the min-max range and tolerance intervals:

- The most commonly applied approach for establishing similarity ranges is the x-sigma interval, that is, mean ± x SD of the RP batch data. The multiplier used (x) should be scientifically justified and could be linked to the criticality of the quality attribute tested, with a smaller multiplier applied for high criticality quality attributes.
- A conservative approach would be to establish the similarity ranges directly based on the min-max quality attribute data obtained from the characterization studies of RP batches. Such similarity ranges could be viewed as clinically qualified (since the RP batches are on the market and taken by patients). However, compared to other approaches the min-max approach is often associated with high risk of a false-negative conclusion (that is, a high risk of concluding non-similarity even though the underlying data distributions for the RP and biosimilar would support a similarity claim).
- Similarity ranges based on tolerance intervals would usually require a high number of RP batches for establishing meaningful ranges. With a limited number of RP batches characterized and/or inappropriate parameterization, the tolerance interval approach can result in an estimated range that is much wider than the actual min-max quality attribute ranges of the RP. The risk of a false-positive conclusion of similarity (that is, the risk of concluding similarity where the underlying data distributions do not support such a claim) may therefore be unreasonably high when the similarity ranges are based on inappropriately applied tolerance intervals.

The most frequently applied overall similarity criteria require that a certain percentage of the biosimilar batches (usually between 90% and 100%) fall within the similarity range. This figure should be determined prior to the initiation of the similarity assessment.

7.4.2.2 Analytical similarity evaluation

It is up to the manufacturer to justify the relevance of the established similarity ranges and criteria. Ideally, the data analyses should be robust and should as far as possible minimize the risk of a false-positive conclusion. In some jurisdictions, the use of a stringent similarity evaluation could also allow for discussion with the NRA on further tailoring of the clinical comparability programme. Although decreasing the risk of a false-positive conclusion is of primary importance from a patient and regulatory point of view, the risk of a false-negative conclusion also needs to be managed by the manufacturer and should be thoroughly considered during the planning of the similarity exercise.

Some minor differences between the RP and the biosimilar are expected. Nevertheless, any quality attributes not fulfilling the established similarity criteria should be considered as a potential signal for non-similarity and should be assessed for possible impact on clinical safety and efficacy. Confirmed differences in low criticality quality attributes also need to be adequately considered, but in the case of such differences reference to available information (which could, for example, originate from scientific publications) is usually sufficient. Lower impurity levels in the biosimilar (for example, of aggregates) or differences in quality attributes present at very low levels in both the RP and the biosimilar would in most cases be predicted to have no clinical relevance, and could therefore be accepted without further assessment. For differences in quality attributes with higher criticality, functional assays to thoroughly address their possible clinical impact are generally expected. Where there are confirmed differences in the most critical quality attributes it will be more challenging to justify the conclusion that the product is a true biosimilar. For example, if differences are found in quality attributes that alter the PK of the product and thereby change the dosing scheme then the product cannot be considered to be a biosimilar.

7.5 Specifications

Specifications are employed to verify the routine quality of the drug substance and drug product rather than to fully characterize them. As for any biological product, the specifications for a biosimilar should be set as described in established guidelines. Furthermore, a biosimilar should show the same level of compliance with a pharmacopoeial monograph as that required for the RP – however, compliance with a pharmacopoeial monograph is not sufficient to establish biosimilarity. It should also be noted that pharmacopoeial monographs may provide only a minimum set of requirements for a particular product, and specification of additional test parameters may be required. Reference to the analytical methods used and acceptance limits for each test parameter of the biosimilar should be provided and justified. All analytical methods referenced in the specification should be validated and the corresponding validation documented.

Specifications for a biosimilar may not be the same as for the RP since the manufacturing processes will be different, and different analytical procedures and laboratories will be used for the assays. Nonetheless, the specifications should capture and control important known product quality attributes. The setting of specifications should be based on: (a) the manufacturer's experience with the biosimilar (for example, with regard to its manufacturing history, assay capability and the quality profile of batches used for establishing similarity); (b) the experimental results obtained by testing and comparing the biosimilar and RP; and (c) attributes with potential impact on product performance. The manufacturer should take into consideration that the limits set for a given specification should not, unless properly justified, be significantly wider than the range of variability of the RP over the shelf-life of the product.

7.6 Stability

Stability studies should comply with relevant guidance as recommended by the NRA. Generally, stability studies should be summarized in an appropriate format (such as tables) and should include results from accelerated degradation studies and studies under various stress conditions (for example, high temperature, oxidation, freeze-thaw, light exposure, humidity and mechanical agitation). There are a number of specific reasons for performing stability studies:

- First, the stability data should support the conclusions reached on the recommended storage and shipping conditions, and on the shelf-life and storage period for the drug substance, drug product and process intermediates – which might be stored for significant periods of time. Real-time/real-temperature stability studies will determine the storage conditions and shelf-life for the biosimilar – which may or may not be the same as those for the RP. Results from studies conducted under accelerated and stress conditions may also show that additional controls should be used in the manufacturing process, and during shipping and storage, in order to ensure the integrity of the product.
- Secondly, stability studies should be carried out to show which release and characterization methods are stability-indicating for the product.
- Thirdly, comparative stability studies conducted under accelerated, and in some cases stress conditions (for example, freeze-thaw, light exposure and mechanical agitation), can be valuable in determining the similarity of the products by showing a comparable degradation profile and rate, with formulation, volume, concentration and/or container differences taken into account.

Stability studies on the drug substance should be carried out using containers and conditions that are representative of the actual storage containers and conditions. Stability studies on the drug product should be carried out in the intended drug product container closure system.

8. Nonclinical evaluation

This section addresses the pharmaco-toxicological assessment of the biosimilar. It is important to note that in order to design an appropriate nonclinical study programme a clear understanding of the characteristics of the RP is required.

The nature and complexity of the RP will have an impact on the extent of the nonclinical studies needed to confirm biosimilarity. In addition, any differences observed between the biosimilar and RP in the physicochemical and biological analyses will also guide the planning of the nonclinical studies. Other factors that need to be taken into consideration include the mechanism(s) of action of the drug substance (for example, the receptor(s) involved) in all authorized indications of the RP, and the pathogenic mechanisms involved in the disorders included in the therapeutic indications.

A stepwise approach should be applied during nonclinical development to evaluate the similarity of the biosimilar and its selected RP. At first, in vitro studies should be conducted and then a decision made on whether or not additional in vivo animal studies are required.

The following approach to nonclinical evaluation may be considered and should be tailored on a case-by-case basis to the biosimilar concerned. In all cases, the approach chosen should be scientifically justified in the application dossier.

8.1 In vitro studies

In order to assess any relevant difference in pharmaco-toxicological activity between the biosimilar and chosen RP, data from a number of comparative in vitro studies – some of which may already be available from the quality-related assays – should be provided. In light of this data overlap, it is suggested that the in vitro nonclinical studies related to characterization of the biological activity of the biosimilar be addressed alongside the related quality data in the corresponding quality module (see section 7.3.2 above). Any other nonclinical in vitro studies should then be addressed in the relevant nonclinical modules of the dossier where they should be reviewed and discussed from the point of view of potential impact on the efficacy and safety of the biosimilar.

Since experience has shown that in vitro assays are in general more specific and sensitive than in vivo studies in animals for detecting differences between the biosimilar and RP, the use of in vitro assays is of paramount importance in the nonclinical biosimilar comparability exercise.

For such in vitro studies, the following general principles apply:

- Typically, a battery of interaction studies addressing the primary binding events should be performed, along with cell-based or isolated-tissue-based functional assays (see below) in order to assess if any (clinically) relevant differences in reactivity exist between the biosimilar and RP and, if so, to determine the likely causative factor(s).
- Together, these assays should cover the whole spectrum of pharmaco-toxicological aspects with potential clinical relevance for the RP and for the product class. In the dossier, the manufacturer should discuss to what degree the in vitro assays used can be considered representative/predictive of the clinical situation according to current scientific knowledge.
- The studies should be comparative and designed to be sufficiently sensitive, specific and discriminatory to allow for the detection of (clinically) relevant differences in pharmaco-toxicological activity between the biosimilar and RP – or, conversely, to provide evidence that any observed differences in quality attributes are not clinically relevant.
- The studies should compare the concentration–activity/binding relationship of the biosimilar and the RP at the pharmacological target(s), covering a concentration range within which potential differences are most accurately detectable (that is, the ascending part of the concentration–activity/binding curve).
- A sufficient number of RP batches and biosimilar batches (preferably representative of the material intended for commercial use) should be evaluated. Assay and batch-to-batch variability will affect the number of batches needed. The number tested should be sufficient to draw meaningful conclusions on the variability of a given parameter for both the biosimilar and the RP and on the similarity of both products (see section 7.4.1 above).
- Where available, international reference standards can be used to support assay characterization, calibration and performance (see section 7.1 above). When no such reference standard exists, an in-house reference material should be established.

The nonclinical in vitro programme for biosimilars should usually include relevant assays for the following:

- **Binding studies**

Evaluation of the primary binding events – that is, binding of the biosimilar to cell membrane receptors or to other membrane-bound or soluble targets that are known/assumed to be involved in the pharmaco-toxicological effects of the RP in the clinically approved indications – for example, for immunoglobulin G (IgG)-based mAbs, antigen-binding fragment (Fab)-associated binding to the antigen and Fc-associated binding to representative isoforms of the relevant Fc receptors and to C1q – see *(10)*.

- **Functional studies/determination of biological activities**

  Studies should evaluate signal transduction and/or functional activity/viability of cells or isolated tissues known to be of relevance for the pharmaco-toxicological effects of the RP. Together these assays should broadly cover all the known mechanisms of action of the RP in the clinically authorized indications – for example, for IgG-based mAbs directed against membrane-bound antigens, evaluation of Fab-associated functions and of Fc-associated functions such as ADCC, ADCP and CDC – see *(10)*.

Such assays are often technically demanding and the experimental approach chosen should be appropriately justified by the manufacturer.

For additional guidance on these topics see section 7.3 above.

8.2 Determination of the need for in vivo animal studies

On the basis of the totality of quality and nonclinical in vitro data available and the extent to which there is residual uncertainty about the similarity of a biosimilar and its RP, it is at the discretion of the involved NRA to waive or not to waive a requirement for additional nonclinical in vivo animal studies. The decision of the NRA on whether or not to require such studies should take into account the following:

- If the quality comparability exercise and the nonclinical in vitro studies have shown high similarity and the level of residual uncertainty is considered acceptable to move to the clinical phase of the similarity exercise then an additional in vivo animal study is not considered necessary.
- If a need is identified to reduce remaining uncertainties concerning the similarity (including drug safety) of a biosimilar and its RP before the initiation of clinical evaluations then additional in vivo animal studies may be considered, if a relevant animal model is available – however this should only occur: (a) when it is expected that such studies would provide relevant additional information; and (b) if the needed additional information cannot be obtained using an alternative approach that does not involve in vivo animal studies. In this respect, the factors to be considered could include:
  - qualitative and/or quantitative differences in potentially or known relevant quality attributes between the biosimilar and its RP (for example, qualitative and/or quantitative differences in the post-translational glycosylation of proteins); and
  - relevant differences in formulation (for example, use of excipients in the biosimilar not widely used in medicinal products).
- On the basis of regulatory experience gained to date in marketing authorization applications for biosimilars, the need for additional in vivo animal studies would be expected to represent a rare scenario.

- If the quality and nonclinical in vitro comparability exercises indicate relevant differences between the biosimilar and the RP (thus making it unlikely that biosimilarity would eventually be established), then standalone development to support a full marketing authorization application should be considered instead (see section 5 above).

8.3 In vivo studies

8.3.1 General aspects to be considered

In the exceptional case that an in vivo evaluation is deemed necessary by the involved NRA, the focus of the study/studies (PK and/or PD and/or safety) will depend upon the type of additional information needed.

Animal studies should be designed to maximize the information obtained. The 3Rs principles for animal experiments (Replace, Reduce, Refine) should always be followed to minimize the use of animals in testing.

To address the residual uncertainties, the use of conventional animal species and/or of specific animal models (for example, transgenic animals or transplant models) may be considered.

Animal models are often not sensitive enough to detect small differences. If a relevant and sufficiently sensitive in vivo animal model cannot be identified, the manufacturer may choose to proceed directly to clinical studies, taking into account strict principles to mitigate any potential risk.

The effects of RPs are often species specific. In accordance with ICH S6(R1) *(19)* and the WHO Guidelines on the quality, safety and efficacy of biotherapeutic protein products prepared by recombinant DNA technology *(8)*, in vivo studies should be performed only in relevant species – that is, species which are known to be pharmacologically and/or toxicologically responsive to the RP.

The duration of the study/studies should be justified, taking into consideration the PK behaviour of the RP, the time to onset of formation of anti-drug antibodies (ADAs) in the test species and the clinical use of the RP.

8.3.2 Specific aspects

8.3.2.1 PK and/or PD studies

In cases where such studies are considered necessary, the PK and/or PD of the biosimilar and the RP should be compared quantitatively, when the model allows, using a dose–response assessment that includes the intended exposure in humans.

The studies may include animal models of disease to evaluate functional effects on disease-related PD markers or efficacy measures.

8.3.2.2 Safety studies

Where in vivo safety studies are deemed necessary, a flexible approach that follows the 3R principles to maximize the readout of relevant data and minimize the use of animals in testing should always be followed. If appropriately justified, a repeated dose toxicity study with refined design – for example, using just one dose level of biosimilar and RP, and/or just one gender and/or no recovery animals, and/or only in-life safety evaluations such as clinical signs, body weight and vital functions – may be considered. Depending on the chosen end-points, it may not be necessary to sacrifice the animals at the end of the study.

Repeated dose toxicity studies in non-human primates are not recommended and nor are toxicity studies in non-relevant species (for example, to assess unspecific toxicity due to impurities).

8.3.2.3 Immunogenicity studies

Qualitative or quantitative difference(s) in product-related variants (for example, in glycosylation patterns, charge, aggregates, and impurities such as host-cell proteins) may have an effect on immunogenic potential and on the potential to cause hypersensitivity. These effects are usually difficult to predict from animal studies and are better assessed in clinical studies.

However, determination of antibody formation against the study drugs may be required for the interpretation of PK/toxicokinetic data in cases where in vivo animal studies are needed.

8.3.2.4 Local tolerance studies

Studies on local tolerance are usually not required. However, if excipients are introduced for which there is little or no experience with the intended clinical route of application, local tolerance may need to be evaluated. If other in vivo animal studies are to be conducted, the evaluation of local tolerance may be integrated into the design of those studies.

8.3.2.5 Other studies

In general, safety pharmacology and reproductive and development toxicity studies – as well as genotoxicity and carcinogenicity studies; see *(8)* and *(19)* – are not warranted during the nonclinical testing of biosimilars.

9. Clinical evaluation

The main clinical data should be generated using the biosimilar product derived from the final manufacturing process, and which reflects the product for which marketing authorization is being sought. Any deviation from this recommendation needs to be justified and additional data may be required. For changes in the manufacturing process, relevant guidelines should be followed *(9, 17)*. Ideally, an RP from a single marketing authorization holder would be used as the comparator throughout the comparability programme of quality and clinical studies during the evaluation of the biosimilar in order to allow for the generation of coherent data and conclusions.

Clinical studies are a valuable step in confirming similarity. The goal of such studies is to confirm the absence of any clinically relevant differences between the proposed biosimilar and the RP.

Clinical studies should be designed to demonstrate confirmative evidence of the similar clinical performance of the biosimilar and the RP, and therefore need to use testing strategies that are sufficiently sensitive to detect any clinically relevant differences between the products.

If relevant differences between the biosimilar and the RP are detected at any stage of development, the reasons will need to be explored and justified. If this is not possible, the new product may not qualify as a biosimilar and a full licensing (standalone) application should be considered.

A comparative bioequivalence study involving PK and/or PD comparability is generally required for clinical evaluation. An adequately powered comparative efficacy and safety trial will not be necessary if sufficient evidence of biosimilarity can be drawn from other parts of the comparability exercise. The need for a comparative clinical efficacy and safety trial for the proposed biosimilar (and type of trial if required) will be influenced by factors such as:

- how well the biosimilar can be characterized;
- the availability of suitable, sensitive and orthogonal assays for adequate analytical and functional characterization;
- the degree of analytical and functional similarity between the biosimilar and RP;
- the existence of a relevant PD parameter;
- the degree of understanding of the mechanism(s) of action of the biological product in different indications and how well these can be investigated in binding and functional in vitro tests – the contribution of each mechanism of action to the observed clinical effect is not relevant as long as it can be measured;
- knowledge of any (potentially) unwanted immunogenicity – for example, ADA incidence and the magnitude of ADA response including level of neutralizing antibodies, and antibodies targeting endogenous substances (for example, erythropoietin and coagulation factors); and
- whether the impurity profile or the nature of excipients of the biosimilar gives rise to clinical concerns.

Current examples of biological products that can be comprehensively characterized and have a well-established mechanism of action include (but are not limited to) teriparatide, insulin, G-CSF and somatropin *(20, 21)*. The current data suggest that more-complex products such as mAbs can be sufficiently characterized by available suitable analytical methods, plus the structure–function relationships are well known and can be studied by sensitive orthogonal functional assays *(22)*.

9.1 Pharmacokinetic studies

The clinical comparability exercise should generally include a comparative PK study, if the drug substance can be measured in the blood, and should also include the measurement of PD markers if available and also immunogenicity data.

The PK study should be designed to demonstrate similar PK profiles for the biosimilar and the RP. When the RP and its proposed biosimilar have more than one route of administration (most commonly intravenous and subcutaneous) then carrying out the study/studies using the non-intravenous route of administration is preferred as this is usually the more immunogenic route and will provide more meaningful information for the comparability exercise. The omission of a PK study of other approved routes of administration needs to be justified for approval of all available options – for example, in cases when the molecule has an absorption constant that is much lower than the elimination constant (flip flop kinetics).

The sample size should be appropriate, taking into account PK variability in the study population, and consideration should be given to whether a cross-over or parallel group design would be the most adequate. If appropriate population PK or PK-PD models are available for the RP in the literature, modelling and simulation can be considered for optimizing study design – for example, justification of dose(s) and selection of the most sensitive study population to detect potential PK differences, and choice of sample size.

PK studies should preferably be performed in healthy volunteers (if considered ethical) and care should be taken to standardize the population with regard to factors that may influence variability (for example, ethnic origin, body weight and gender). If the drug substance under investigation is associated with risks or tolerability issues that are considered to be unacceptable for healthy volunteers, it will be necessary to perform the PK studies in patients.

The preferred design is a randomized, two-period, two-sequence, single-dose cross-over PK study using a dose within the therapeutic range at which the ability to detect

differences is sufficient to observe meaningful differences. The cross-over design eliminates inter-subject variability and therefore (compared with the parallel group design) reduces the sample size needed to show equivalent PK profiles of the biosimilar and RP. The treatment periods should be separated by a wash out phase that is sufficiently long to ensure that drug concentrations are below the lower limit of bioanalytical quantification in all subjects at the beginning of the second period – that is, at least 5 times the terminal half-life.

When a cross-over design is not suitable (for example, for biological products with a very long half-life or associated with immunogenicity affecting PK) then a parallel group study should be considered. In parallel group studies, care should be taken to avoid any imbalances between treatment groups that might affect the PK of the drug substance under investigation (for example, with regard to ethnic origin, body weight and gender).

A multiple-dose study in patients is acceptable as a pivotal PK study if a single-dose study cannot be conducted in healthy volunteers due to risks or tolerability reasons or if a single-dose study is not feasible in patients. Multiple-dose studies may also be acceptable in rare situations where problems with the sensitivity of the analytical method preclude sufficiently precise plasma or serum concentration measurements after a single dose administration. However, given that a multiple-dose study is less sensitive in detecting differences in Cmax than a single-dose study, this will only be acceptable with sound justification.

PK comparison of the biosimilar and the RP should not only include the rate and extent of absorption but also a descriptive analysis of elimination characteristics – that is, clearance and/or elimination half-life – which might differ between the biosimilar and the RP. Linear (nonspecific) clearance and nonlinear (target-mediated) clearance should be evaluated by assessment of partial areas under the curve (pAUCs). For further details on primary and secondary end-points for single- and multiple-dose PK studies, please refer to further guidance documents *(23)*.

Acceptance criteria for the demonstration of PK similarity between the biosimilar and the RP must be predefined and appropriately justified. It should be noted that the criteria used in standard clinical PK comparability studies (bioequivalence studies) may not necessarily be applicable to all biotherapeutic products. However, the traditional 80–125% equivalence range will in most cases be sufficiently conservative to establish similar PK profiles *(24)*. Correction for protein content may be acceptable on a case-by-case basis if pre-specified and adequately justified, with the assay results for the biosimilar and RP being included in the protocol. If adjustments for covariates are intended for parallel group studies (for example, in the case of adalimumab, stratification for body weight and gender), they should be predefined in the statistical analysis plan rather than being included in post hoc analyses.

Other PK studies, such as interaction studies (with drugs likely to be used concomitantly) or studies in special populations (for example, children, the elderly and patients with renal or hepatic insufficiency), are not required for a biosimilar.

Particular consideration should be given to the analytical method selected and its ability to detect and follow the time course of the protein in a complex biological matrix that contains many other proteins. The method should be optimized to provide satisfactory specificity, sensitivity and a range of quantification of adequate accuracy and precision. The same assay should be used to detect the serum concentrations of both the biosimilar and RP. A single PK assay (same binding reagents and a single analytical standard, usually a biosimilar) for determining biosimilar and RP concentration in a biological matrix can be adopted based on verification of the bioanalytical comparability of the two products within the method, with supporting data *(25)*.

In some cases the presence of measurable concentrations of endogenous protein may substantially affect the measurement of the concentration–time profile of the administered

exogenous protein. In such cases the manufacturer should describe and justify the approach taken to minimize the influence of the endogenous protein on the results (for example, baseline correction).

In some cases it may not be possible or meaningful to establish PK similarity due to the nature of the substance (for example, fractionated and unfractionated heparin cannot be measured in blood), the route of administration (for example, intraocular administration of aflibercept or ranibizumab) or unacceptably high PK variability (for example, romiplostim). In such cases clinical similarity should be supported by PD, immunogenicity and/or other clinical parameters.

9.2 Pharmacodynamic studies

PD parameters should preferably be investigated as part of the comparative PK studies. In some cases PK studies cannot reasonably be conducted and PD markers may then play a more important role. This is for example the case with heparins,[2] where serum concentrations cannot be measured and similarity needs to be established for the most important PD end-points; that is, at least anti-FXa and anti-FIIa activity.

PD effects should be investigated in a suitable population using a dose or doses within the steep part of the dose–response curve in order to maximize the chance of detecting potential differences between the biosimilar and the RP. PD markers should be selected on the basis of their clinical relevance.

9.3 Confirmatory PK and/or PD studies

If an adequately powered comparative efficacy trial is not necessary, comparative PK (see section 9.1 above) and/or PD studies (see section 9.2 above) may be sufficient for establishing confirmative evidence of the similar clinical performance of a biosimilar and its RP, provided that *(24)*:

- the acceptance ranges for confirmatory PK and/or PD end-points are predefined and appropriately justified;
- the PD biomarker reflects the mechanism of action of the biological product;
- the PD biomarker is sensitive to potential differences between the proposed biosimilar and the RP; and
- the PD biomarker assay is validated.

The applicant should consider the option of using additional PD measures (usually as secondary end-points) to assess the comparability of the PD properties of the RP and proposed biosimilar. Furthermore, even if relevant PD measures are not available, sensitive PD end-points may be assessed if such assessment may help to reduce residual uncertainty about biosimilarity.

An example of acceptable confirmatory PK/PD studies would be the use of euglycaemic clamp studies to compare the efficacy of two insulins. In addition, absolute neutrophil count and CD34+ cell count are the relevant PD markers for assessing the activity of G-CSF and could be used in PK/PD studies in healthy volunteers to demonstrate the similar efficacy of two medicinal products containing G-CSF.

The study population and dosage should represent a test system that is known to be sensitive in detecting potential differences between a biosimilar and the RP. In the case of

[2] Regulated as a biological in most countries *(1)*.

insulin, for example, the study population should consist of non-obese healthy volunteers or patients with type 1 diabetes rather than insulin-resistant obese patients with type 2 diabetes. Otherwise, it may be necessary to investigate more than one dose to demonstrate that the test system is discriminatory *(26)*.

The acceptance ranges for confirmatory PK and/or PD parameters (that is, for primary end-points) should be predefined and appropriately justified. If PD comparison is not essential for a conclusion of biosimilarity but the results are still expected to reasonably support biosimilarity then a purely descriptive analysis of the PD results may be justified. This may be the case for biological substances that have been extensively characterized and for which biosimilarity can already be concluded from the analytical, functional and PK comparisons. If appropriately designed and performed, such PK/PD studies are usually more sensitive in detecting potential differences in efficacy than trials using hard clinical end-points.

However, PD markers may also be used as end-points in clinical efficacy studies in patients.

Examples of appropriate markers include haemoglobin for measuring the efficacy of an epoetin, and lactate dehydrogenase (which is a sensitive biochemical marker of intravascular haemolysis) for evaluating the efficacy of a complex drug such as eculizumab. For denosumab, investigation of bone formation and resorption markers as part of the PK study may be useful or possibly sufficient. This would involve measurement of bone mineral density and bone turnover markers such as serum C-terminal telopeptide of type 1 collagen (CTX-1) and procollagen type 1 N-terminal propeptide (P1NP) after denosumab administration.

In certain cases (for example, when analytical similarity of the active ingredient in the biosimilar and the RP can be demonstrated to such a degree that clinical differences can be excluded) a comparative PK study may provide sufficient clinical evidence to support biosimilarity. However, a risk assessment (including for example, the impurity profile) should be conducted to determine the need for additional safety/immunogenicity data on the biosimilar (see sections 9.5 and 9.6 below).

9.4 Efficacy studies

A comparative efficacy trial may not be necessary if sufficient evidence of biosimilarity can be inferred from other parts of the comparability exercise. A comparative clinical trial, if necessary, should confirm that the clinical performance of the biosimilar and the RP is comparable. Demonstration of comparable potency, PK and/or PD profiles provide the basis for use of the RP posology in the comparative clinical trial.

If a comparative clinical trial of the biosimilar and RP is deemed necessary then it is expected that it will be an adequately powered, randomized and controlled clinical trial performed in a patient population that allows for sensitive measurement of the intended clinical parameters. The principles of such trials are laid down in relevant ICH guidelines *(26–28)*.

In principle, equivalence trial designs (requiring lower and upper comparability margins) are preferred for comparing the efficacy and safety of the biosimilar and RP. Non-inferiority designs (requiring only one margin) *(26)* or trials with asymmetrical margins may be considered if appropriately justified *(29)*. Regardless of which design is selected in a particular case, the comparability margin(s) must be pre-specified and justified on the basis of clinical relevance – that is, the selected margin should represent the largest difference in efficacy that would not matter in clinical practice. Treatment differences within this margin would therefore be acceptable as they would have no clinical relevance.

Similar efficacy implies that similar treatment effects can be achieved when using the same posology, and the same dosage(s) and treatment schedule should be used in clinical trials comparing the biosimilar and RP. In this regard, equivalence trials are again preferable to

ensure that the biosimilar is not clinically less or more effective than the RP when used at the same dosage(s).

A non-inferiority design could be acceptable, if justified by the applicant, for example:

- for biological products with high efficacy (for example, a response rate of over 90%), making it difficult to set an upper margin; or
- in the presence of a wide safety margin.

When using asymmetrical margins, the narrower limit should rule out inferior efficacy and the broader limit should rule out superior efficacy. The use of asymmetrical margins should be fully justified by the sponsor of the proposed biosimilar. Factors that would allow for the use of such margins in a clinical trial include:

- if the dose used in the clinical study is near the plateau of the dose–response curve; and
- there is little likelihood of dose-related adverse effects (for example, toxicity).

The final results obtained from the comparative clinical trial(s) along with comparative analytical, functional and PK data will determine whether the biosimilar and the RP can be considered to be clinically similar. If clinically relevant differences are found, a root cause analysis should be performed. If a plausible cause that is unrelated to the product (for example, inadvertent baseline differences between treatment groups despite randomization) cannot be found, the new product should not be considered to be similar to the RP.

Careful consideration should be given to the design of the comparative study/studies, including the choice of primary efficacy end-point(s). Studies should be conducted using a clinically relevant and sensitive end-point within an homogenous population that responds well to the pharmacological effects of the biological product of interest to show that there are no clinically meaningful differences between the biosimilar and RP. Clinical outcomes, surrogate outcomes (PD markers) or a combination of both can be used as primary end-points in biosimilar trials. The same study end-points used to establish the efficacy of the RP may be used because a large body of historical data would generally be available in the public domain for setting the comparability margin(s) and calculating the sample size. However, the primary end-point could be different from the original study end-point for the RP if it is well justified and relevant data are available to support its use as a sensitive end-point and its suitability for the determination of the comparability margin(s). A relevant PD end-point can be used as the primary end-point – for example, when it is a known surrogate of efficacy or when it can be linked to the mechanism of action of the product. The primary or secondary end-points can also be analyzed at different time points compared to those used in clinical trials with the RP if these are considered to be more sensitive in capturing the pharmacological action(s) of the biological product – for example, adalimumab efficacy could be measured by responses at week 12 or 16 in addition to week 24.

The sample size and duration of the comparative clinical study should both be adequate to allow for the detection of clinically meaningful differences between the biosimilar and RP. When a comparative clinical trial is determined to be necessary then adequate scientific justification for the choice of study design, study population, study end-point(s), estimated effect size for the RP and comparability margin(s) should be provided and may be discussed with regulators in order to obtain agreement at least in principle prior to trial initiation.

9.5 Safety

Safety data should be captured throughout clinical development from PK/PD studies and also in clinical efficacy trials when conducted. Knowledge of: (a) the type, frequency and severity of adverse events/reactions when compared with the RP; (b) whether these are due to exaggerated pharmacological actions; (c) the degree of analytical and functional similarity of the biosimilar and RP; and (d) the presence of novel impurities and novel excipients in the biosimilar will all inform the type and extent of data required to characterize the safety profile of the biosimilar.

If the clinical programme for the biosimilar is limited to confirmatory PK/PD studies, this will need to be adequately justified and a risk assessment should be conducted to determine the need to obtain additional safety data for the biosimilar. For example, for insulin the most relevant safety issue is hypoglycaemia which can be attributed to its pharmacological action. Highly similar physicochemical characteristics and PK/PD profiles of the biosimilar and RP could provide sufficient reassurance that the risk of hypoglycaemia is also similar, obviating the need for further safety data. Similar examples are teriparatide, filgrastim or somatropin. The current data suggest that more-complex products such as mAbs can be sufficiently characterized and also fall into this category *(22)*.

If the biosimilar contains impurities that are not present in the RP (for example, because of the use of a novel expression system) then the generation of further safety data may be necessary, or scientific justification should be provided as to why such data are not needed. Manufacturers should consult with regulators when proposing a clinical programme solely relying on PK/PD studies.

As for all medicinal products, further monitoring of the safety of the biosimilar will be necessary in the post-marketing phase (see section 10 below).

9.6 Immunogenicity

Immunogenicity should be investigated as part of the clinical evaluation package of the biosimilar relative to the RP unless the manufacturer can provide a scientific justification that human immunogenicity data are not needed. Such justification should be based on the degree of physicochemical similarity of the biosimilar and RP, and on a thorough risk assessment of any unwanted immunogenicity and clinical consequences known for the RP. Although published information will be useful in gaining knowledge of the immunogenicity risk of the RP and in planning the immunogenicity strategy, it is not generally sufficient to support approval of the biosimilar. The goal of the immunogenicity programme is to exclude an unacceptable/marked increase in the immunogenicity of the biosimilar when compared with the immunogenicity of the RP and to generate descriptive data in support of biosimilar approval and its clinical use. If conducted, the immunogenicity study report should include data on antibody incidence, magnitude of ADA response and neutralization ability, whether antibodies are transient or persistent, and their impact on PK and clinical correlates *(30)*.

The marketing authorization application should include an integrated immunogenicity summary comprising a risk assessment and, if appropriate, the results of testing using appropriately validated and characterized assays, along with details on the clinical study duration, sampling schedules and regimen, and the clinical immunogenicity assessment *(30–32)*.

The immunogenicity studies should be tailored to each product and require a multidisciplinary approach taking into account both quality and clinical considerations. The risk assessment should include:

- accumulated information on the immunogenicity of the RP (that is, on the nature, frequency and clinical relevance of the immune response);

- consideration of the quality aspects (including the nature and complexity of the drug substance, non-glycosylated/glycosylated, expression system, product- and process-related impurities, and aggregates);
- consideration of excipients and container closure system, and stability of the product, route of administration, dosing regimen; and
- consideration of patient- and disease-related factors (for example, immune competent/compromised and any concomitant immunomodulatory therapy).

Placing particular emphasis on any differences in product-related factors (for example, impurities arising from a novel expression system and/or novel excipients) that could modify immunogenicity will be crucial in the risk assessment of the biosimilar. Importantly, consideration of the type of product is also a critical element of the risk assessment, with the risk being higher for a product that has an endogenous non-redundant counterpart (for example, epoetin). In such cases, special attention should be paid to the possibility of the immune response seriously affecting the endogenous protein and its unique biological function, with serious adverse effects. Real-time testing for neutralizing ADAs is recommended for epoetins *(33)* and other high-risk products (for example, enzyme replacement therapies and coagulation factors). Conversely, for well-characterized biological substances (for example, insulin, somatropin, filgrastim, teriparatide), where an extensive literature and clinical experience indicate that immunogenicity does not impact upon product safety and efficacy, immunogenicity studies may not be necessary, provided that the biosimilar is highly similar to the RP and the risk-based evaluation indicates a low risk. This may also be applicable to other products, including mAbs. In such cases, manufacturers should consult with the regulatory authorities.

Appropriate scientific justification for not conducting a safety/immunogenicity study should always be provided.

9.6.1 Immunogenicity testing

A multi-tiered approach comprising screening and confirmatory immunoassays that detect binding ADAs followed by assays which determine ADA magnitude and neutralization potential is generally necessary and deviation from this requires justification.

Information on current assays and formats and on their benefits and limitations, along with the interpretation of results, has been extensively reviewed *(33–36)*. The manufacturer will need to justify the antibody-testing strategy and the choice of assays to be used. Attention should be given to the selection of suitable controls for assay validation and to the determination of cut-off points for distinguishing antibody-positive from antibody-negative samples. Aspects relating to potential interference by matrix components, including the pharmacological target and the residual drug in the sample, are also important. To mitigate such interference, corrective measures should be implemented. For example, for drug interference (which commonly occurs with samples taken from patients given mAbs) measures such as allowing time for clearance of the drug from the circulation prior to sampling, or incorporating steps for dissociating immune complexes and/or removal of the drug can be used. Care should be taken to ensure that the use of such measures does not compromise ADA detection or patient treatment.

Where required, comparative immunogenicity testing should be performed using the same assay format and sampling schedule. For immunogenicity assessment in new drug development, antibody testing is performed using the therapeutic given to the patient. In applying this concept to biosimilars, the development of screening assays with a similar sensitivity for the two patient groups (biosimilar and RP) within the same study is very challenging. Therefore, in the biosimilar scenario, relative immunogenicity is often assessed

by using a single assay which employs the drug substance of the biosimilar as the antigen for sample testing for both groups. This approach allows for the detection of all antibodies developed against the biosimilar. The manufacturer should demonstrate the suitability of the method(s) used and provide data assuring that the method(s) measure ADA to the RP and to the biosimilar to a similar extent *(25)*.

Neutralization assays reflecting the mechanism of action are usually based on the potency assay of the product. Non-cell ligand-based assays are relevant in cases where the therapeutic binds to a soluble ligand and inhibits its biological action. For products associated with high risk (for example, those with non-redundant endogenous homologs) and those for which effector functions are important, the use of functional cell-based bioassays is recommended. Where necessary, advice on the need for a neutralization assay and on the appropriate format to use (cell-based, ligand-based or based on enzyme activity) may be sought from regulatory authorities.

Further characterization of antibodies (for example, isotype) should be conducted if considered clinically relevant, or in special situations (for example, the occurrence of anaphylaxis or use of certain assay formats), taking into account the immunogenicity profile of the RP. For example, if the RP does not elicit an IgE response it is unlikely that the biosimilar would elicit one if the same expression system is used. The retention of patient samples under appropriate storage conditions will be necessary for retesting in cases where technical problems occurred with the original assay.

9.6.2 Clinical evaluation

ADAs can affect the PK, PD, safety and/or efficacy of the administered product. The immunogenic risk of a biological is determined by the ADA incidence in the treated population and the magnitude of the unwanted clinical effect, and influences the benefit–risk balance of the therapeutic.

If human immunogenicity data are needed, they should be generated in a comparative manner throughout the clinical programme. The sensitive patient population (that is, the population with the highest likelihood of mounting an immune response) is preferred for investigating immunogenicity. For example, if an epoetin is licensed for the treatment of renal anaemia and for patients with chemotherapy-induced anaemia, the selection of patients with renal anaemia is advised. Comparative PK and/or PD studies should be designed to also collect immunogenicity data regardless of the population to be included (for example, healthy volunteers and patients). A PK/PD cross-over design is possible for immunogenicity testing but if the exposure time until the switch does not provide sufficient immunogenicity data, the sponsor must ensure that a sufficient number of patients are treated without cross-over – for example, by extending the cross-over study with two parallel treatment arms, or by proposing a separate immunogenicity study.

If ADAs are known to affect the PK of the RP then ADA rate and kinetics assessments could be performed along with assessment of their impact on PK through pre-specified subgroup analysis of ADA-negative and -positive subjects.

The observation period required for immunogenicity testing will depend on the expected time of antibody development and should be justified by the manufacturer. Sampling during immunogenicity testing should include baseline sampling (prior to treatment) for pre-existing antibodies, sampling during treatment and in some cases post-treatment, particularly if ADAs persist or are undetectable at earlier time points (due to immunosuppressive properties of the product or technical problems such as drug interference). The sampling schedule should be synchronized for evaluation of PK as well as for assessment of safety and efficacy to provide an understanding of the impact of antibodies on clinical outcome. Generally, for chronic administration, 6-month data are acceptable to exclude excessive immunogenicity, but in some

cases a longer evaluation period may be appropriate pre-licensing to assess antibody incidence and possible clinical effects.

Furthermore, notable differences in immunogenicity between the biosimilar and RP would require further investigation of the underlying cause, and data and justification provided to support any claim that the difference noted was not clinically relevant. An analysis of the clinical impact of ADAs in both arms on PK, efficacy and/or safety should be performed through stratified analysis of ADA-negative and -positive subjects.

Any potential for the production of neutralizing antibodies against critical endogenous factors (for example, following epoetin administration) will necessitate clinical studies in patients.

As is the case with the RP, the biosimilar should also undergo robust post-marketing surveillance that includes assessment of any serious adverse events related to immunogenicity.

9.7 Authorization of indications

The decision to authorize the requested indications will be dependent upon the demonstration of similarity between the biosimilar and RP. The extension of indications from the RP to the biosimilar is only possible if the following requirements are fulfilled:

- similarity in analytical characteristics and functional properties has been confirmed in sensitive orthogonal assays which provide information on the clinically relevant mechanism of action and/or involved receptor(s) as part of the comparability exercise; and
- this is supported by clinical data (comparative PK and/or PD study – see sections 9.1–9.3 above) plus a comparative clinical trial performed in a patient population that allows sensitive measurement of the intended clinical parameters, if necessary (see sections 9.4–9.6 above).

For example, authorization of all indications may be obtained based on highly comparable functional data – for example, for biosimilars of mAbs such as infliximab and adalimumab if they show fully comparable activity (including ADCC, CDC, reverse signalling and apoptosis) both in terms of binding to soluble TNF and membranous TNF.

10. Pharmacovigilance

Following approval, many NRAs consider a biosimilar to have its own life-cycle and there is no formal requirement to re-establish similarity to the RP when comparability exercises are conducted following manufacturing changes *(9, 17).* Both RP and biosimilar manufacturers are responsible for ensuring that their products remain safe and efficacious throughout their life-cycle by preventing significant changes to individual products. In this context, it is important to emphasize that the required data can be obtained only by having robust pharmacovigilance systems in place that allow for the collection of product-specific data.

As with all medicinal products, further close monitoring of the safety and efficacy of a biosimilar in all approved indications, along with continued benefit–risk assessment, are necessary in the post-marketing phase. Any specific safety monitoring or risk-minimization measures imposed on the RP or product class should be incorporated into the pharmacovigilance plan for the relevant biosimilar unless a compelling justification can be provided to show that this is not necessary. Furthermore, participation in existing disease registries should be encouraged and is mandatory if also mandatory for the RP. Post-marketing safety reports should include all information on product safety received by the marketing

authorization holder. The safety information must be evaluated in a scientific manner and this should include evaluation of the frequency and cause of adverse events.

The manufacturer should submit a pharmacovigilance plan describing a safety specification, pharmacovigilance activities and risk-minimization activities at the time of submission of the marketing authorization application or whenever a safety concern arises post-marketing. The principles of pharmacovigilance planning can be found in relevant guidelines such as ICH E2E *(37)*. The safety specification should describe important identified or potential safety issues for the RP and for the substance class as well as any that are specific to the biosimilar. If there are any remaining uncertainties regarding the biosimilar – due for example to the use of a novel excipient or device – then these should be included in the pharmacovigilance plan and followed up post-marketing.

Manufacturers should ensure that at the time of the marketing authorization they have in place an appropriate pharmacovigilance system, including the services of a qualified person responsible for monitoring pharmacovigilance activities and the necessary means for notification of adverse reactions that occur in any of the countries in which the product is marketed.

After the marketing authorization has been granted, it is the responsibility of the NRA to monitor closely the compliance of manufacturers with their marketing commitments, particularly with regard to their pharmacovigilance obligations as described here.

In addition, as with all biological products, an adequate system for ensuring the specific identification of the biosimilar (that is, traceability) is essential. The NRA shall provide a legal framework for proper pharmacovigilance surveillance and ensure the ability to identify any biological marketed in its area of jurisdiction that is the subject of an adverse reaction report. In addition to the international nonproprietary name (INN) *(38)* an adverse reaction report for any biological should also include all other important indicators, including the proprietary (brand) name, manufacturer's name and lot number. The country of origin is not strictly required.

11. Labelling and prescribing information

The biosimilar should be clearly identifiable by a unique trade name together with the INN. From the perspective of WHO there is no specific INN nomenclature for biosimilars – that is, there is no part of an INN which indicates that a product is a biosimilar. Biosimilars are assigned INNs using the process and rules used for all biological products. In many cases, the INN for a biosimilar is the same as that for its RP – for example, for G-CSF biosimilars that have used Neupogen as the RP, both the biosimilar and the RP have the INN "filgrastim" *(39, 40)*. Provision of the lot number is essential as it is an important part of production information and is critical for traceability whenever problems with a product are encountered.

The prescribing information for a biosimilar should be as similar as possible to that of the RP except for product-specific aspects such as use of different excipient(s) and/or presentations. This similarity is particularly important for posology and for safety-related information, including contraindications, warnings and known adverse events. However, if there are fewer indications for the biosimilar than for the RP, the related text in various sections may be omitted unless it is considered important in informing doctors and patients of certain risks – for example, as a result of potential off-label use. In such cases it should be clearly stated in the prescribing information that the biosimilar is not intended for use in the specific indication(s) and the reasons why.

12. Roles and responsibilities of NRAs

One of the responsibilities of an NRA is to set up appropriate regulatory oversight for the licensing and post-marketing surveillance of biosimilars that are developed and/or authorized for use in its area of jurisdiction. The experience and expertise of the NRA in evaluating biological products is a key prerequisite for appropriate regulatory oversight of these products. The NRA is responsible for clearly defining a suitable regulatory framework for licensing biological products, including biosimilars *(41)*.

As the development of biological products is a rapidly evolving area, NRAs may need to conduct regular reviews of their licensing, the adequacy of their regulations for providing oversight, and the processes and policies that constitute the regulatory framework. Such a process of review is an essential component of well-functioning and up-to-date regulatory oversight of biological products *(42)*. Some countries have licensed products called "biosimilars" that were approved prior to the establishment of a regulatory framework for biosimilar approval. WHO recommends avoiding use of this term (or other equivalent term) for products that have not been evaluated in line with the principles set out in these Guidelines. NRAs should develop a specific, appropriate, regulatory framework for approving biosimilars that is distinct from the regulatory procedures previously applied to products with a version of the same active ingredient intended for the same use but for which regulatory evaluation was not well defined *(41, 43)*. In addition, the terminology used for such products should not be confused by calling them "biosimilars".

NRAs could improve access to biosimilars of assured quality, safety and efficacy by improving the efficiency of their regulatory evaluation – for example, by making efforts to reduce the time taken for evaluation without compromising the quality of the review process *(41, 43)*. In addition, efforts should be made to avoid the unnecessary duplication of studies *(44)*.

Most countries either use or amend their existing legislation and applicable regulations or develop entirely novel regulatory frameworks for the authorization of biosimilars. In some jurisdictions, regulations for licensing subsequent entry versions of biotherapeutic products are intricately linked with policies for innovation. Hence an NRA may need to coordinate and communicate with other stakeholders to ensure consistency *(45)*.

Authors and acknowledgements

The various drafts of these WHO Guidelines were prepared by a WHO drafting group comprising: Dr P. Aprea, Administración Nacional de Medicamentos, Alimentos y Tecnología Medica, Argentina; Dr S. Barry, Health Products Regulatory Authority, Ireland; Dr M-C Bielsky, Medicines and Healthcare products Regulatory Agency, the United Kingdom; Dr N. Ekman (*lead author for the quality section*), Finnish Medicines Agency, Finland; Dr H-K Heim (*lead author for the nonclinical section*), Federal Institute for Drugs and Medical Devices, Germany; Dr J. Joung, Ministry of Food and Drug Safety, Republic of Korea; Dr P. Kurki, University of Helsinki, Finland; Dr E. Lacana, United States Food and Drug Administration, the USA; Dr C. Njue, Health Canada, Canada; Dr E. Nkansah, Food and Drug Authority, Ghana; Dr M. Savkina, Federal State Budgetary Institution Scientific Centre for Expert Evaluation of Medicinal Products, Russian Federation; Dr R. Thorpe, Consultant, the United Kingdom; Dr T. Yamaguchi, Pharmaceuticals and Medical Devices Agency, Japan; Dr M. Wadhwa, National Institute for Biological Standards and Control, the United Kingdom; Dr J. Wang, Health Canada, Canada; Dr J. Wang, National Institutes for Food and Drug Control, China; Dr J. Welch, United States Food and Drug Administration, the USA; Dr M. Weise, Federal Institute for Drugs and Medical Devices, Germany; Dr E. Wolff-Holz (*lead author for the clinical section*), Paul-Ehrlich-Institut, Germany; and Dr H-N Kang, World Health Organization, Switzerland.

The first draft document was then posted on the WHO Biologicals website from 27 April to 24 May 2021 for a first round of public consultation. Comments were received from: M. Baldrighi (consolidated comments), Medicines for Europe, Belgium; Biological Products Office (consolidated comments), Agência Nacional de Vigilância Sanitária, Brazil; M.A.A. Boller (consolidated comments), National Institute for Quality Control in Health, Brazil; R.B. Arcuri, Grupo FarmaBrasil, Brazil; Dr A. Cook, Medicines and Healthcare products Regulatory Agency, the United Kingdom; Dr M. Gencoglu (consolidated comments), International Federation of Pharmaceutical Manufacturers & Associations (IFPMA), Switzerland; Dr C.P.V. González, Universidad Nacional de Colombia, Colombia; D. Goryachev, Federal State Budgetary Institution Scientific Center for Examination of Medical Devices, Russian Federation; Dr S.S. Jadhav, Serum Institute India, India; Z. Kusynová (consolidated comments), International Pharmaceutical Federation; Dr G.E. Medgyesi, National Institute of Pharmacy and Nutrition, Hungary; C. Moreno, Comité de veeduría y cooperación en Salud, Colombia; Dr R. Naruse, Pharmaceuticals and Medical Devices Agency, Japan; B.A. Nhaquila, Ministério da Saúde, Mozambique; Dr T.P. Petkovic (consolidated comments), Swissmedic, Switzerland; C. Rao, Third World Network; Dr E. Satterwhite (consolidated comments), International Generic and Biosimilar Medicines Association (IGBA), the USA; Dr M. Savkina, Federal State Budgetary Institution Scientific Centre for Expert Evaluation of Medicinal Products, Russian Federation; K. Sehmi, International Alliance of Patients' Organizations, Switzerland; A.P. Seisdedos, Public Health Institute of Chile, Chile; Dr W. Wei, National Medical Products Administration, China; Dr S. Yim, United States Food and Drug Administration, the USA; and A.M.A. Zuñiga (consolidated comments), Instituto Nacional de Vigilancia de Medicamentos y Alimentos, Colombia.

The main issues addressed during the public consultation were reviewed at a WHO informal consultation to revise the WHO Guidelines on evaluation of similar biotherapeutic products, held virtually from 30 June to 2 July 2021 and attended by: *Drafting group members*: Dr P. Aprea, Administración Nacional de Medicamentos, Alimentos y Tecnología Medica, Argentina; Dr M-C Bielsky, Medicines and Healthcare products Regulatory Agency, the United Kingdom; Dr N. Ekman, Finnish Medicines Agency, Finland; Dr H-K Heim, Federal Institute for Drugs and Medical Devices, Germany; Dr J. Joung, Ministry of Food and Drug Safety, Republic of Korea; Dr P. Kurki, University of Helsinki, Finland; Dr E. Lacana, United States Food and Drug Administration, the USA; Dr C. Njue, Health Canada, Canada; Dr E. Nkansah, Food and Drug Authority, Ghana; Dr M. Savkina, Federal State Budgetary Institution Scientific Centre for Expert Evaluation of Medicinal Products, Russian Federation; Dr R. Thorpe, Consultant, the United Kingdom; Dr T. Yamaguchi, Pharmaceuticals and Medical Devices Agency, Japan; Dr M. Wadhwa, National Institute for Biological Standards and Control, the United Kingdom; Dr J. Wang, Health Canada, Canada; Dr J. Wang, National Institutes for Food and Drug Control, China; Dr M. Weise, Federal Institute for Drugs and Medical Devices, Germany; and Dr E. Wolff-Holz, Paul-Ehrlich-Institut, Germany. *Other participants*: Dr M. Allam, Dr H. Bahaa and Dr M. Sayed, Egyptian Drug Authority, Egypt; Dr A. Al-Oballi, Jordan Food and Drug Administration, Jordan; Mr A. Alshahrani, Saudi Food & Drug Authority, Saudi Arabia; Dr D. Baek and Ms J. Kim, Ministry of Food and Drug Safety, Republic of Korea; Ms H.M. Chua, Malaysia National Pharmaceutical Regulatory Agency, Malaysia; Mr J. Gangakhedkar and Mr P. Jagtap, Central Drugs Standard Control Organisation, India; T. Lyaskovsky, Ministry of Health of Ukraine, Ukraine; Dr S. Okudaira, Pharmaceuticals and Medical Devices Agency, Japan; Ms W. Ondee, Ministry of Public Health, Thailand; Dr P.S. Sotomayor and Dr J.I. Solis Ricra, Ministry of Health, Peru; and Dr J. Uviase, National Agency for Food and Drug Administration and Control, Nigeria. *Representatives of the Developing Countries Vaccine Manufacturers Network*: Dr F. Ahmed, Incepta Pharmaceuticals Ltd, Bangladesh; Dr Y. Rajendran, Zydus Cadila Healthcare, India; and Dr

H.G. Tonioli Defendi, R&D Biomanguinhos, Brazil. *Representatives of the Emerging Biopharmaceutical Manufacturers Network*: Dr S. Yi O Cho, Instituto Butantan, Brazil; and Dr A. Qu, China National Biotec Group Company Ltd, China. *Representatives of the IFPMA*: Dr V. Acha, MSD, Switzerland; Dr M. Gencoglu, IFPMA, Switzerland; and Dr K. Ho, F. Hoffmann-La Roche Ltd, Switzerland. *Representatives of IGBA*: Dr M. Baldrighi, Medicines for Europe, Belgium; Dr M. Schiestl, Sandoz Biopharmaceuticals, Austria; and Dr K. Watson, Celltirion, Republic of Korea. *Representative of the Latin American Association of Pharmaceutical Industries*: Dr E. Spitzer, Latin American Association of Pharmaceutical Industries, Argentina. *Representative of the Singapore Association of Pharmaceutical Industries*: Dr S. Chong, Singapore Association of Pharmaceutical Industries, Singapore. *WHO Secretariat*: Dr A. Fukushima, Dr H-N Kang, Dr I. Knezevic, Dr G. Pante and Dr M. Simão, Access to Medicines and Health Products, World Health Organization, Switzerland.

Based on the outcomes of the above informal consultation, the document WHO/BS/2022.2413 was prepared by the original drafting group and posted on the WHO Biologicals website from 8 November 2021 to 7 January 2022 for a second round of public consultation. Comments were received from: Dr A.E.C.C. Almeida (consolidated comments), National Institute for Quality Control in Health, Brazil; Mr A. Alshahrani (consolidated comments), Saudi Food & Drug Authority, Saudi Arabia; M-C. Annequin (consolidated comments), Agence nationale de sécurité du médicament et des produits de santé, France; R.B. Arcuri, Grupo FarmaBrasil, Brazil; A.M. Awamlehi, Jordan Food and Drug Administration, Jordan; M. Baldrighi (consolidated comments), Medicines for Europe, Belgium; Ms K. Choudhury, Guru Govind Singh Indraprastha University, India; Ms H.M. Chua, Malaysia National Pharmaceutical Regulatory Agency, Malaysia; Dr M. Gencoglu (consolidated comments), IFPMA, Switzerland; Dr C.P.V. González, Universidad Nacional de Colombia, Colombia; S.M. Hassan, Malaysia National Pharmaceutical Regulatory Agency, Malaysia; Dr P. Huleatt, Australian High Commission, Singapore; U. Katneni, United States Food and Drug Administration, the USA; Dr M. Kucuku, National Agency for Medicines and Medical Devices, Albania; Dr G.E. Medgyesi, National Institute of Pharmacy and Nutrition, Hungary; Dr S.N. Niazi, University of Illinois Chicago, the USA; Dr S. Roosendaal, Quality RA B.V., Netherlands; Dr E. Satterwhite (consolidated comments), IGBA, the USA; Dr J. Southern, South Africa National Regulatory Authority, South Africa; Dr H.G. Tonioli Defendi, R&D Biomanguinhos, Brazil; Dr S. Wendel, Hospital Sirio Libanes, Brazil; and Dr G. Zenhäusern (consolidated comments), Swissmedic, Switzerland.

Further changes were made to document WHO/BS/2022.2413 by the Expert Committee on Biological Standardization.

References

1. Kang H-N, Thorpe R, Knezevic I et al. The regulatory landscape of biosimilars: WHO efforts and progress made from 2009 to 2019. Biologicals. 2020;65:1–9 (https://www.sciencedirect.com/science/article/pii/S1045105620300269?via%3Dihub , accessed 26 September 2021).

2. Joung J, Robertson JS, Griffiths E, Knezevic I, on behalf of the WHO Informal Consultation Group. WHO informal consultation on regulatory evaluation of therapeutic biological medicinal products held at WHO Headquarters, Geneva, 19–20 April 2007. Biologicals. 2008;36(4):269–76 (https://www.sciencedirect.com/science/article/abs/pii/S1045105607001169, accessed 26 September 2021).

3. Guidelines on evaluation of similar biotherapeutic products (SBPs). In: WHO Expert Committee on Biological Standardization: sixtieth report. Geneva: World Health Organization; 2013: Annex 2 (WHO Technical Report Series, No. 977; http://who.int/biologicals/publications/trs/areas/biological_therapeutics/TRS_977_Annex_2.pdf, accessed 26 September 2021).

4. Resolution WHA67.21. Access to biotherapeutic products including similar biotherapeutic products and ensuring their quality, safety and efficacy. Sixty-seventh World Health Assembly, Geneva, 18–26 May 2014. Geneva: World Health Organization; 2014 (http://apps.who.int/gb/ebwha/pdf_files/WHA67/A67_R21-en.pdf, accessed 26 September 2021).

5. Revision of the WHO Guidelines on evaluation of similar biotherapeutic products (SBPs). In: WHO Expert Committee on Biological Standardization: Report of the seventy-second and seventy-third meetings. Geneva: World Health Organization; 2021 (WHO Technical Report Series, No. 1030, section 3.1.4, pp. 30–31; https://www.who.int/publications/i/item/9789240024373, accessed 28 September 2021).

6. Kurki P, Kang H-N, Ekman N, Knezevic I, Weise M, Wolff-Holz E. Regulatory evaluation of biosimilars: refinement of principles based on the scientific evidence and clinical experience. BioDrugs (in press).

7. Wadhwa M, Kang H-N, Thorpe R, Knezevic I. WHO Informal consultation on revision of guidelines on evaluation of similar biotherapeutic products. Virtual meeting held 30 June to 2 July 2021. Biologicals (in press).

8. Guidelines on the quality, safety and efficacy of biotherapeutic protein products prepared by recombinant DNA technology. In: WHO Expert Committee on Biological Standardization: sixty-fourth report. Geneva: World Health Organization; 2014: Annex 4 (WHO Technical Report Series, No. 987; http://www.who.int/biologicals/biotherapeutics/TRS_987_Annex4.pdf?ua=1, accessed 26 September 2021).

9. Comparability of biotechnological/biological products subject to changes in their manufacturing process. ICH Harmonised Tripartite Guideline Q5E (Step 4). Geneva: International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use; 2004 (https://database.ich.org/sites/default/files/Q5E%20Guideline.pdf, accessed 30 October 2021).

10. Guidelines on evaluation of monoclonal antibodies as similar biotherapeutic products (SBPs). In: WHO Expert Committee on Biological Standardization: sixty-seventh report. Geneva: World Health Organization; 2017: Annex 2 (WHO Technical Report Series, No. 1004; http://who.int/biologicals/biotherapeutics/WHO_TRS_1004_web_Annex_2.pdf?ua=1, accessed 26 September 2021).

11. Recommendations for the preparation, characterization and establishment of international and other biological reference standards (revised 2004). In WHO Expert

Committee on Biological Standardization: fifty-fifth report. Geneva: World Health Organization; 2005: Annex 2 (WHO Technical Report Series, No. 932; https://www.who.int/bloodproducts/publications/TRS932Annex2_Inter_biolefstandardsrev2004.pdf?ua=1, accessed 28 September 2021).

12. Thorpe R, Wadhwa M. Intended use of reference products & WHO international standards/reference reagents in the development of similar biological products (biosimilars). Biologicals. 2011;39(5):262–5 (https://www.sciencedirect.com/science/article/pii/S1045105611000704?via%3Dihub, accessed 28 September 2021).

13. Prior S, Metcalfe C, Hufton SE, Wadhwa M, Schneider CK, Burns C. Maintaining "standards" for biosimilar monoclonal antibodies. Nat Biotechnol. 2021;39(3):276–80 (available at: https://www.nature.com/articles/s41587-021-00848-0, accessed 28 September 2021).

14. Report of a WHO informal consultation on international standards for biotherapeutic products. In: WHO Expert Committee on Biological Standardization: sixty-sixth report. Geneva: World Health Organization; 2016 (WHO Technical Report Series, No. 999, section 2.2.4, pp. 13–15; https://www.who.int/publications/i/item/9789241209991, accessed 28 September 2021).

15. WHO good manufacturing practices for pharmaceutical products: main principles. In: WHO Expert Committee on Specifications for Pharmaceutical Preparations: forty-eighth report. Geneva: World Health Organization; 2014: Annex 2 (WHO Technical Report Series, No. 986; http://www.who.int/medicines/areas/quality_safety/quality_assurance/TRS986annex2.pdf?ua=1, accessed 26 September 2021).

16. WHO good manufacturing practices for biological products. In: WHO Expert Committee on Biological Standardization: sixty-sixth report. Geneva: World Health Organization; 2016: Annex 2 (WHO Technical Report Series, No. 999; http://www.who.int/biologicals/areas/vaccines/Annex_2_WHO_Good_manufacturing_practices_for_biological_products.pdf?ua=1, accessed 26 September 2021).

17. Guidelines on procedures and data requirements for changes to approved biotherapeutic products. In: WHO Expert Committee on Biological Standardization: sixty-eighth report. Geneva: World Health Organization; 2018: Annex 3 (WHO Technical Report Series, No. 1011; https://www.who.int/publications/m/item/approved-biotherapeutics-a3-trs-no-1011, accessed 26 September 2021).

18. Quality risk management – Q9. ICH Harmonised Tripartite Guideline. Current Step 4 version dated 9 November 2005. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use; 2005 (https://database.ich.org/sites/default/files/Q9%20Guideline.pdf, accessed 2 October 2021).

19. Preclinical safety evaluation of biotechnology-derived pharmaceuticals – S6(R1). ICH Harmonised Tripartite Guideline. Current Step 4 version: June 2011. International Conference on Harmonisation of Technical Requirements for Registration of

Pharmaceuticals for Human Use; 2011 (https://database.ich.org/sites/default/files/S6_R1_Guideline_0.pdf, accessed 2 October 2021).

20. Blum WF, Alherbish A, Alsagheir A, El Awwa A, Kaplan W, Koledova E et al. The growth hormone-insulin-like growth factor-I axis in the diagnosis and treatment of growth disorders. Endocr Connect. 2018;7(6):R212–22 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5987361/, accessed 15 April 2022).

21. Romer T, Zabransky M, Walczak M, Szalecki M, Balser S. Effect of switching recombinant human growth hormone: comparative analysis of phase 3 clinical data. Biol Therapy. 2011;1(1):Article number 5 (https://link.springer.com/article/10.1007/s13554-011-0004-8, accessed 15 April 2022).

22. Bielsky M-C, Cook A, Wallington A, Exley A, Kauser S, Hay L et al. Streamlined approval of biosimilars: moving on from the confirmatory efficacy trial. Drug Discov Today. 2020;25(11):1910–8 (https://www.sciencedirect.com/science/article/pii/S1359644620303433?via%3Dihub , accessed 2 October 2021).

23. Guideline on the investigation of bioequivalence. Committee for Medicinal Products for Human Use. London: European Medicines Agency; 2010 (Document CPMP/EWP/QWP/1401/98 Rev. 1/Corr; (https://www.ema.europa.eu/documents/scientific-guideline/guideline-investigation-bioequivalence-rev1_en.pdf, accessed 2 October 2021).

24. Li J, Florian J, Campbell E, Schrieber SJ, Bai JPF, Weaver JL et al. Advancing biosimilar development using pharmacodynamic biomarkers in clinical pharmacology studies. Clin Pharmacol Ther. 2020;107(1):40–2 (https://ascpt.onlinelibrary.wiley.com/doi/epdf/10.1002/cpt.1653, accessed 2 October 2021).

25. Civoli F, Kasinath A, Cai X-Y, Wadhwa M, Exley A, Oldfield P et al. Recommendations for the development and validation of immunogenicity assays in support of biosimilar programs. The AAPS J. 2020;22(7) (abstract: https://link.springer.com/article/10.1208%2Fs12248-019-0386-y, accessed 2 October 2021).

26. Choice of control group and related issues in clinical trials – E10. ICH Harmonised Tripartite Guideline. Current Step 4 version: 20 July 2000. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use; 2000 (https://database.ich.org/sites/default/files/E10_Guideline.pdf, accessed 2 October 2021).

27. ICH Topic E 9. Statistical principles for clinical trials. Step 5. Note for guidance on statistical principles for clinical trials. London: European Medicines Agency; 2006 (Document CPMP/ICH/363/96; https://www.ema.europa.eu/en/documents/scientific-guideline/ich-e-9-statistical-principles-clinical-trials-step-5_en.pdf, accessed 2 October 2021); and Addendum on estimands and sensitivity analysis in clinical trials

to the guideline on Statistical principles for clinical trials – E9(R1). ICH Harmonised Guideline. International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use; 2019 (https://database.ich.org/sites/default/files/E9-R1_Step4_Guideline_2019_1203.pdf, accessed 31 October 2021).

28. ICH guideline E17 on general principles for planning and design of multi-regional clinical trials. Step 5. London: European Medicines Agency; 2017 (Document EMA/CHMP/ICH/453276/2016 Rev.1; https://www.ema.europa.eu/en/documents/scientific-guideline/ich-guideline-e17-general-principles-planning-design-multi-regional-clinical-trials-step-5-first_en.pdf, accessed 2 October 2021).

29. Scientific considerations in demonstrating biosimilarity to a reference product. Guidance for industry. U.S. Department of Health and Human Services, Food and Drug Administration; 2015 (https://www.fda.gov/media/82647/download, accessed 2 October 2021).

30. Shankar G, Arkin S, Cocea L, Devanarayan V, Kirshner S, Kromminga A et al. Assessment and reporting of the clinical immunogenicity of therapeutic proteins and peptides – harmonized terminology and tactical recommendations. AAPS J. 2014;16(4):658–73 (abstract: https://pubmed.ncbi.nlm.nih.gov/24764037/, accessed 2 October 2021).

31. Guideline on Immunogenicity assessment of therapeutic proteins. Committee for Medicinal Products for Human Use (CHMP). London: European Medicines Agency; 2017 (Document EMEA/CHMP/BMWP/14327/2006 Rev 1; https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-immunogenicity-assessment-therapeutic-proteins-revision-1_en.pdf, accessed 31 October 2021).

32. Immunogenicity testing of therapeutic protein products – developing and validating assays for anti-drug antibody detection. U.S. Department of Health and Human Services, Food and Drug Administration; 2019 (https://www.fda.gov/media/119788/download, accessed 2 October 2021).

33. Haag-Weber M, Eckardt K-U, Hörl WH, Roger SD, Vetter A, Roth K. Safety, immunogenicity and efficacy of subcutaneous biosimilar epoetin-α (HX575) in non-dialysis patients with renal anemia: a multi-center, randomized, double-blind study. Clin Nephrol. 2012;77(1):8–17 (https://www.researchgate.net/publication/51879081_Safety_immunogenicity_and_efficacy_of_subcutaneous_biosimilar_epoetin-A_HX575_in_non-dialysis_patients_with_renal_anemia_a_multi-center_randomized_double-blind_study, accessed 2 October 2021).

34. Wadhwa M, Knezevic I, Kang H-N, Thorpe R. Immunogenicity assessment of biotherapeutic products: an overview of assays and their utility. Biologicals. 2015;43(5):298–306 (https://www.sciencedirect.com/science/article/pii/S1045105615000627?via%3Dihub, accessed 2 October 2021).

35. Shankar G, Devanarayan V, Amaravadi L, Barrett YC, Bowsher R, Finco-Kent D et al. Recommendations for the validation of immunoassays used for detection of host antibodies against biotechnology products. J Pharm Biomed Anal. 2008;48(5):1267–81 (abstract: https://www.sciencedirect.com/science/article/abs/pii/S0731708508005001 accessed 2 October 2021).

36. Gupta S, Devanarayan V, Finco D, Gunn GR 3rd, Kirshner S, Richards S et al. Recommendations for the validation of cell-based assays used for the detection of neutralizing antibody immune responses elicited against biological therapeutics. J Pharm Biomed Anal. 2011;55(5):878–88 (abstract: https://pubmed.ncbi.nlm.nih.gov/21531522/, accessed 2 October 2021).

37. ICH Topic E 2 E. Pharmacovigilance planning (Pvp). Step 5. Note for guidance on planning pharmacovigilance activities. London: European Medicines Agency; 2005 (Document CPMP/ICH/5716/03; https://www.ema.europa.eu/en/documents/scientific-guideline/international-conference-harmonisation-technical-requirements-registration-pharmaceuticals-human-use_en-25.pdf, accessed 2 October 2021).

38. International Nonproprietary Names (INN) for biological and biotechnological substances (a review). Geneva: World Health Organization; 2019 (Document WHO/EMP/RHT/TSN/2019.1; https://www.who.int/medicines/services/inn/BioReview2019.pdf?ua=1, accessed 2 October 2021).

39. Guidance on the use of international nonproprietary names (INNs) for pharmaceutical substances. Geneva: World Health Organization; 2017 (https://www.who.int/medicines/services/inn/FINAL_WHO_PHARM_S_NOM_1570_web.pdf?ua=1).accessed 2 October 2021).

40. Robertson JS, Chui W-K, Genazzani AA, Malan SF, de la Rica Manjavacas AL, Mignot G et al. The INN global nomenclature of biological medicines: a continuous challenge. Biologicals. 2019;60:15–23 (https://www.sciencedirect.com/science/article/pii/S104510561930048X, accessed 2 October 2021).

41. Regulatory harmonization. 16th International conference of drug regulatory authorities (ICDRA). WHO Drug Information. 2014;28(3):297–306 (https://apps.who.int/iris/bitstream/handle/10665/331105/DI283-297-306-eng.pdf?sequence=1&isAllowed=y, accessed 2 October 2021).

42. Regulatory assessment of approved rDNA-derived biotherapeutics. In: WHO Expert Committee on Biological Standardization: sixty-sixth report. Geneva: World Health Organization; 2016: Annex 3 (WHO Technical Report Series, No. 999; http://www.who.int/biologicals/areas/biological_therapeutics/Annex_3_Regulatory_assessment_of_approved_rDNA-derived_biotherapeutics.pdf?ua=1, accessed 26 September 2021).

43. Kang H-K, Knezevic I. Regulatory evaluation of biosimilars throughout their product life-cycle. Bull World Health Organ. 2018;96:281–5 (https://www.who.int/bulletin/volumes/96/4/17-206284.pdf, accessed 2 October 2021).

44. Good reliance practices in the regulation of medical products: high level principles and considerations. In: WHO Expert Committee on Specifications for Pharmaceutical Preparations: fifty-fifth report. Geneva: World Health Organization; 2021: Annex 10 (WHO Technical Report Series, No. 1033; https://apps.who.int/iris/bitstream/handle/10665/340323/9789240020900-eng.pdf, accessed 5 April 2022).

45. Kang H-K, Thorpe R, Knezevic I, Levano MC, Chilufya MB, Chirachanakul P et al. Regulatory challenges with biosimilars: an update from 20 countries. Ann N Y Acad Sci. 2021;1491(1):42–59 (https://nyaspubs.onlinelibrary.wiley.com/doi/10.1111/nyas.14522, accessed 2 October 2021).