16 March 2026
EMA/CHMP/BMWP/60916/2025
Committee for Medicinal Products for Human Use (CHMP)
Biosimilar Medicinal Products Working Party (BMWP)

Reflection paper on a tailored clinical approach in biosimilar development

Draft for internal consultation agreed by Biosimilar Medicines Working Party	21 October 2024
Consultation with MWP, BWP and SAWP	17 January 2025
Draft agreed by Biosimilar Medicinal Products Working Party	12 February 2025
Adopted by CHMP for release for consultation	17 March 2025
Start of public consultation	27 March 2025
End of consultation (deadline for comments)	30 September 2025
Agreed by Biosimilar Medicinal Products Working Party	02 March 2026
Adopted by CHMP	16 March 2026

Keywords	**Reflection Paper, Biosimilar, Comparative Efficacy Study, Tailored clinical approach**

Official address Domenico Scarlattilaan 6 ● 1083 HS Amsterdam ● The Netherlands
Address for visits and deliveries Refer to www.ema.europa.eu/how-to-find-us
Send us a question Go to www.ema.europa.eu/contact **Telephone** +31 (0)88 781 6000 An agency of the European Union

Reflection paper on a tailored clinical approach in biosimilar development

Table of contents

1. Introduction

In the field of pharmacology, it is a well-established scientific principle that the biological activity (resulting in efficacy and to a large degree, safety) of any pharmacologically active substance, whether it is a small molecule like paracetamol or a large protein like a monoclonal antibody, stems from its interaction with its receptor(s) (including membrane receptors, ligands, substrates, and other targets).

Such interactions are usually highly specific for the pharmacologically active substance in question. It is also scientifically well understood that these receptor interactions are determined by the structure of the pharmacologically active substance. In other words, structure determines function, and as an immediate corollary, the same structure results in the same biological activity.

This scientific principle also extends to recombinant proteins and other biological products; the biological activity, and ultimately the therapeutic effects, are dictated by the structure. Consequently, if the structure of two proteins is the same, or at least highly similar, then these two proteins will bind to the same receptors in the same quantitative manner, and will therefore have the same pharmacological properties, and ultimately the same clinical efficacy.

This scientific principle has been widely accepted and used to support changes in the manufacturing processes of biological products with well-defined structural attributes. Significant changes in the manufacturing processes of biological medicines like monoclonal antibodies have been approved by confirmation of structural and functional comparability through a comprehensive comparative analytical testing without the need for new clinical data. This experience, together with technical advances in analytical characterisation, supports the notion that, under specific conditions, analytical comparability exercises and pharmacokinetic (PK) data can be sufficient for demonstrating biosimilarity.

This reflection paper will examine settings where similar clinical efficacy and safety of a biosimilar candidate and its reference medicinal product (RMP) can be inferred from a conclusion of analytical, i.e. physicochemical and functional similarity and comparable pharmacokinetics (PK). A further driver for this Reflection paper is the regulatory experience indicating that, in the past, the results of the Comparative Efficacy Studies (CES) generally did not add pivotal information for the assessment of biosimilarity (Guillen *et al*., Kirsch-Stefan *et al*., Bielsky MC *et al*., IPRP workshop report 2024).

2. Scope

This Reflection Paper will discuss the necessity of CES for demonstration of biosimilarity. In order to place those reflections into context, the Reflection Paper will first consider the current practice with respect to analytical comparability exercises and consider their predictive value. Subsequently, some reflections will be provided regarding the contribution of CES and other human *in vivo* studies, especially PK/PD studies, to the biosimilarity exercise, as well as to the assessment of immunogenicity.

This Reflection Paper is not intended to replace current guidance or current practice with regard to analytical comparability exercises.

All general EMA and ICH quality guidelines for the standard development of a biological medicinal product remain fully applicable to biosimilar development, irrespective of whether a CES is conducted.

3. Discussion

3.1. Quality

3.1.1. General basis and background

Assessing the similarity of biological active substances is challenging because these active substances usually comprise of complex and heterogeneous mixtures. The comparability paradigm emerged approximately 30 years ago as a concept, triggered by the special challenges that biologicals posed. ICH Q5E guideline defines '*comparable*' as '*a conclusion that products have highly similar quality attributes before and after manufacturing process changes and that no adverse impact on the safety or efficacy, including immunogenicity, of the finished product occurred.*' The body text of the guideline further states that '*The demonstration of comparability does not necessarily mean that the quality attributes of the pre-change and post-change product are identical, but that they are highly similar and that the existing knowledge is sufficiently predictive to ensure that any differences in quality attributes have no adverse impact upon safety or efficacy of the finished product.*' The ICH Q5E emphasises the importance of sensitive analytical technologies to determine whether physicochemical differences are present.

The concept of comparability has been used effectively for many years. It recognises that biologicals are inherently variable and that minor differences in quality attributes (QAs) are often clinically irrelevant. The concept of comparability allows for quality differences as long as they do not translate into relevant clinical differences, in other words, comparability does not require that products should be identical. This concept has been used to support the implementation of manufacturing process changes for biological products, without imposing that products should be identical in a physicochemical sense, which may not be achievable, and prevents the need of conducting unnecessary comparative clinical studies. Since the 1990s, major manufacturing changes have been substantiated and implemented based on comparability exercises not including CES. This applies to scenarios such as replacing a product's Master Cell Bank, or a change in the expression system, situations that are from a scientific viewpoint comparable to the development of a biosimilar product (Van der Plas *et al.* 2020).

3.1.2. Prerequisites for similarity assessment

The Guideline on Similar Biological Medicinal Products (CHMP/437/04) states that '*The scientific principles of (..) a biosimilar comparability exercise is based on those applied for evaluation of the impact of changes in the manufacturing process of a biological medicinal product (as outlined in ICH Q5E)*.' The CHMP guideline also underscores that 'comparable safety and efficacy of a biosimilar to its reference medicinal product has to be demonstrated or otherwise justified'.

In line with the concept of comparability, the general requirement for biosimilars is that their QAs are highly similar to those of the reference medicinal product, but that they do not need to be identical. Minor differences are allowed, provided they have no impact on clinical safety or efficacy (see section 3.1.7). Conversely, potentially clinically relevant differences in QAs are not compatible with the biosimilar approach, and in such situations clinical data cannot be used to justify substantial differences in quality attributes.

A comprehensive set of relevant QAs providing detailed information regarding the physicochemical and functional properties of the biological molecule is essential for the demonstration of similarity between a biosimilar candidate and its RMP.

Following identification of the QAs, a risk assessment using prior knowledge in combination with scientifically sound justification should be performed (risk assessment is further discussed in e.g., the relevant ICH guidelines (ICH Q8, Q9 and Q10)). Prior knowledge provides understanding of the critical QAs (CQAs) impacting the interaction with receptor(s) (including membrane receptors, ligands, substrates, and other targets). Whilst it is acknowledged that a quantitative correlation between evaluated CQAs and clinical performance may not always be feasible, available prior knowledge should be such that a robust risk assessment of QA criticality can be conducted. Selection of QAs and an initial criticality assessment and ranking should be completed prior to product development. However, as development proceeds, the knowledge accumulated from the characterisation studies provides increased insight into the QAs, which need to be properly reflected in the design of the analytical similarity exercise and data evaluation approaches to be provided in support of the Marketing Authorisation Application (MAA). Rigorous evaluation of QAs (including functional aspects) during risk assessment in terms of potential impact on PK/PD, efficacy and safety, including immunogenicity, is pivotal for tailoring of clinical data requirements and will therefore have to be thoroughly justified using an interdisciplinary approach.

A commercial manufacturing process with appropriate manufacturing process controls and demonstrated batch-to-batch consistency should be developed. Any batches not fulfilling the pre-determined specifications will be rejected and do not reach patients. This will ensure consistent quality profile across commercial production batches and high similarity between the QAs of the biosimilar batches and the RMP.

In summary, for waiving the need of a CES, the following prerequisites need to be fulfilled:

- comprehensive knowledge regarding the molecule's Mechanism of Action (MoA) is available;
- detailed characterisation of the structure and functionally relevant QAs, including stability indicating QAs, is possible using orthogonal and state-of-the-art analytical methods;
- functional assays (biological tests such as potency tests, receptor binding assays, etc.) are available, both to assess comparability of functional properties directly, and indirectly as surrogates for higher-order structure of the molecule. These functional assays are sometimes referred to as *in vitro* pharmacology; they are expected to confirm that functional (pharmacological) aspects are indeed comparable, as inferred from the comparability of other QAs;
- a validated manufacturing process and control strategy (including but not limited to specification/release testing) to assure future consistency of the biosimilar product;
- a pre-established similarity assessment plan (see section 3.1.3).

These prerequisites will support that an analytical comparability exercise, and data from human PK studies, as appropriate, is able to assure similarity of the biosimilar to its RMP.

On the other hand, waiving a CES is not acceptable in situations where there is a lack of sufficiently sensitive analytical methods or where the MoA is not understood. Furthermore, waiving a CES may not be acceptable when the MoA and structure function relationship(s) are incompletely understood, and consequently, CQAs are difficult to identify and risk-assess. For example, many cell-based medicinal products would today fall under this category.

3.1.3. Similarity assessment plan

As described in section 3.1.2 biosimilar product development is an iterative process.

In order to increase the overall robustness of any biosimilar development programme, it is considered essential that a similarity assessment plan is developed and documented. The plan should capture all critical parts of the physicochemical and functional similarity assessment, such as:

- the number of RMP batches to be included and the sampling plan (see section 3.1.4);
- the number and nature of biosimilar batches (primarily batches manufactured using the commercial manufacturing process and scale) (see section 3.1.4);
- justifications for the list of QAs, including criticality and known link to clinical parameters (PK, efficacy, safety, immunogenicity), that will be considered in the similarity assessment;
- the analytical methods and assays that will be used and the degree of method validation/qualification required; these assays should also include a justified list of functional/biological assays (e.g., receptor binding assays and cell-based potency assays) (see section 3.1.5);
- justification of overall approach planned for the similarity assessment and the planned statistical analysis (see section 3.1.6).

For highly critical QAs (e.g. potency, protein content) it is expected that the results from all biosimilar batches are within the established similarity range. A good biosimilarity development plan should, however, pre-specify what actions will be taken in case a critical QA falls outside the established similarity range. This will reduce the risk that post-hoc justifications would lead to rejection of the application (see section 3.1.7).

Applicants are recommended to make use of the EMA scientific advice procedure to present to and to reach agreement from CHMP on their similarity assessment plan.

3.1.4. Batches to be included in the similarity assessment

The conclusion on similarity should primarily be based on comparative characterisation studies conducted on batches manufactured using the commercial manufacturing process and scale for the biosimilar product. In addition, development batches could be included if comparability to commercial scale batches has been unquestionably demonstrated in line with the principles of ICHQ5E.

A sufficient number of biosimilar batches needs to be tested, usually 6 or more batches originating from independent drug substance batches. All commercial-scale biosimilar batches produced, including process performance qualification batches and batches applied in the clinical trial(s), should be included in the similarity assessment. Any exceptions to this should be described and justified in the similarity assessment plan.

Although it is impossible to specify the exact number of RMP batches needed for every product and scenario, in general at least 10 RMP batches should be included in the similarity assessment. Experience has shown that, in many cases, 15-30 batches of the RMP is optimal, depending on factors like batch independency, criticality and variability of the QAs, the analytical procedures used to investigate them, and the approach applied to assess similarity (see section 3.1.6.). If less than 10 RMP batches are used, this could lead to difficulties in estimating the real batch-to-batch variability present in the RMP with consequences both on the similarity assessment and on the design of the overall control strategy, including release testing for the biosimilar. The final number of RMP batches should be justified by the Applicant in the similarity assessment plan. Continued sampling over time is meaningful to take into account potential shifts or drifts in the RMP, irrespective of the number of batches already sampled.

3.1.5. Analytical considerations

The analytical methods should be state-of-the-art, and ideally orthogonal methods should be used. The previously applied requirements to perform side-by-side analysis have largely become obsolete because most state-of-the-art methods have good analytical precision with little between run/day-to-day variability (or, at least, this variability is similar to within day variability/precision). However, side-by-side analysis might remain meaningful in a situation with strong between run variability, for example, Surface Plasmon Resonance analysis.

In addition to physicochemical QAs, it is expected that relevant and discriminatory functional assays (e.g., receptor binding studies, cell-based potency assays) are available, both to support the evaluation of the criticality of physicochemical QAs, and to provide comparative data between biosimilar and its RMP. Such comparative functional assay data provide evidence that the biological activity, and therefore the clinical activity is the same. Where relevant, such comparative data may not only include potency, concentration-response relationships and binding to targets but also binding to other receptors which may be related to pharmacokinetics, e.g., the FcRn binding for monoclonal antibodies.

In order to preserve RMP batches, freezing has occasionally been proposed and accepted. However, adequate data needs to be provided to show that the freezing/thawing process and storage under frozen conditions does not affect the relevant QAs of the RMP batches.

It is acknowledged that, during the period of development of the biosimilar medicinal product, analytical methods can change. The adequacy of the results from the former methods needs to be confirmed in the MAA dossier and, if needed, re-analyses of batches with the new method provided.

It should be emphasised that the accuracy and precision of the analytical methods need to be high enough so that the differences seen during the characterisation studies mainly reflect real batch-to-batch variability as opposed to variability of the analytical method itself.

3.1.6. Assessment of physicochemical and functional similarity

In order to generate robust evidence for similarity, the Applicant is recommended to follow the general principles outlined in the Guideline on similar biological medicinal products containing biotechnology-derived proteins as active substance: quality issues (EMA/CHMP/BWP/247713/2012). The overall statistical principles relevant for the comparative assessment of quality attributes are described in detail in the Reflection paper on statistical methodology for the comparative assessment of quality attributes in drug development (EMA/CHMP/138502/2017).

The most widely used approach to demonstrate physicochemical and functional similarity for quantitative QAs of high and moderate criticality relies on demonstrating that the QAs of the biosimilar lie within pre-established similarity ranges (also called quality ranges). All biosimilar batches that fall within the batch-to-batch variability of the RMP are expected to have the same clinical performance. Thus, differences within the similarity ranges (e.g. a difference in data distribution) are assumed not to have a relevant impact on safety or efficacy.

The approach based on similarity ranges is derived from the ICH Q5E guideline and is described in the EMA 2014 guideline on "Similar biological medicinal products containing biotechnology-derived proteins as active substance: quality issues". It is noted that the same general approach is referenced for example in the FDA 2025 guideline on "Development of Therapeutic Protein Biosimilars: Comparative Analytical Assessment and Other Quality-Related Considerations Guidance for Industry" and in the WHO 2022 "Guidelines on evaluation of biosimilars".

Similarity ranges are determined based on results from characterization studies of multiple RMP batches. Different approaches can be used for establishing these ranges; however, the X-sigma interval is the most commonly used. The X-sigma interval is defined as Mean ± x*SD, where the multiplier x must be scientifically justified and proportionate to the criticality of the quality attribute being assessed. A smaller multiplier is generally applied for high criticality quality attributes.

The similarity ranges should reflect the batch-to-batch variability present in the RMP, unless it can be determined which differences would be acceptable. However, for certain QAs, such as product-related impurities, it can be sufficient to rule out an increase in the impurity levels. For other QAs, there can be pre-determined general expectations that need to be fulfilled; protein content and most process-related impurities are examples of these. Furthermore, for QAs with the lowest risk ranking or for those quality attributes that cannot be quantitatively measured raw data/graphical comparisons can be used.

For the MAA, the characterisation data should be tabled and presented separately for each batch. The results should also be visualised graphically using suitable methods, such as dotplots and/or boxplots with similarity ranges and information on data spread included.

Other statistical approaches for demonstrating physicochemical and functional similarity may also be used. It is recommended to discuss such approaches in advance with the Agency.

3.1.7. Considerations for specific quality attributes

The analytical similarity package needs to provide convincing evidence that any differences between the biosimilar and reference medicinal product would have no meaningful impact on safety or efficacy. As discussed below, substantial differences which directly impact the MoA, or which could lead to an inferior safety profile, are not compatible with the biosimilarity concept.

This reflection paper uses different qualifiers to specify analytical differences, such as "minor differences" or "major differences" throughout the text. For the sake of clarity, any difference—whether minor or major—in terms of measurable values and clinical impact must be justified and discussed by the Applicant.

A finding that some data points fail to meet similarity (e.g., fall outside the biosimilarity range) for some QAs does not *a priori* preclude biosimilarity or a tailored clinical development programme without a CES. Nonetheless, the presence of (minor) differences may increase the overall uncertainty, which needs to be considered in the conclusion on biosimilarity. If the similarity ranges are not met for some QAs, and the supporting data package and justifications are insufficient to rule out a possible impact on efficacy or safety, developers should consider adapting the manufacturing process of the biosimilar to better align with the quality profile of the RMP. Any remaining uncertainties on the quality or PK level can lead to additional analytical and/or functional data requirements by the CHMP.

For attributes that fall outside the similarity ranges, the level of supporting data required to justify an approval depends on the criticality of the QA in question. Therefore, it is expected that any differences are supported by an appropriate risk assessment which considers the criticality of the QA. It is expected that the applicant has a sufficient understanding of the MoA of the product and has a clear understanding of whether the QA could have a direct impact on the efficacy or safety of the product. Where any quality differences are observed, however minor, the applicant will be expected to present a detailed discussion on the potential impact on safety and efficacy. This discussion can include peer-reviewed literature references, and/or supportive analytical and functional/biological data, where relevant. Confirmed differences in the most critical QAs can generally not be justified by supportive data.

As noted above, the accuracy and precision of the analytical methods need to be sufficiently high. The issue whether any observed difference between the biosimilar and the RMP is due to analytical variability or is a true difference should be carefully considered in any discussion of analytical data. Where applicants consider that analytical variability is the underlying reason for anomalous results, firstly, every effort should be made to improve the overall performace of the method. In addition, multiple repeat testing of the same batch/sample can improve the precision of results obtained, and the use of side-by-side testing could be used to support any inter-assay variability. Claims about analytical variability would not be acceptable without supportive data. It is also important to recognise that differences detected using a sensitive assay typically cannot be overcome by providing additional data from a less sensitive assay. In certain cases, differences observed with a highly sensitive functional assay might be addressed by demonstrating similarity using assays that are more physiologically relevant. However, such approaches must be carefully justified and scientifically supported. Where use of more variable assays is unavoidable (e.g., certain cell-based assays, supported with data demonstrating analytical variability), experience has shown that alternative experimental designs, e.g., analysis of the batches at different time points or the use of method trending data, can provide valuable insights, as it can point to the variance contribution of the assay over time and improve the interpretation of the data.

For QAs that fall outside the similarity ranges, characterisation data using orthogonal assays can provide supportive evidence. Consideration should also be given to increasing the number of batches tested to provide a greater understanding of the true range of variability of that QA in the biosimilar and RMP. There are several approaches which could be included in the similarity assessment plan to address the situation where unanticipated differences in QAs are found. This may help to avoid rejection of the application. Based on experience with biosimilar applications in the EU, some examples are discussed below for particular QAs. However, this is not an exhaustive list, and it is up to the applicant to justify that the additional supportive data package is sufficient to address any uncertainties.

3.1.7.1. Primary and higher order structure

Demonstration of comparable molecular structure of a biosimilar and the RMP is essential to achieve comparable binding affinity to the target. Secondary and tertiary structures determine how a protein folds and maintains its stability, hence any variations in these structures can lead to differences in the protein's functional form, affecting its efficacy and/or safety. Even minor differences in higher order structure can have significant implications for the biosimilarity claim. Therefore highly sensitive and orthogonal analytical methods for higher order structure should be used to thoroughly characterise any potential differences in conformation.

Differences in the primary amino acid structure contradict the biosimilarity concept. While it is noted that low-level sequence variants may occur, these are not considered to be a difference in the primary amino acid sequence; instead, they are product-related substances that can be acceptable if properly described, justified, and controlled. Differences in post-translational modifications are frequently seen, including differences in N/C terminal variants, oxidation, deamidation, etc. An appropriate panel of orthogonal testing is expected to ensure that any apparent differences in post-translational modifications are not clinically relevant. For example, for mAbs, additional computational modelling showing that the deamidation, oxidation and isomerisation sites are not located in an epitope binding region or Fc region or that any differences observed have no impact on binding and/or activity may be relevant. In some cases, additional structure-function data could be provided to show the relationship between the particular post-translational modification and biological activity. Such data could be useful in providing assurance that any differences are unlikely to have an effect on efficacy or safety *in vivo*.

3.1.7.2. Protein content

The batch or batches of the biosimilar candidate used in the comparative clinical PK study should be carefully selected to sufficiently match the protein concentration of the RMP. The actual protein content of each batch used in the PK study should be determined in order to align between biosimilar and RMP, as close as possible (see EMA Clinical pharmacology and pharmacokinetics: questions and answers for further details).

CHMP has encountered several examples where similar PK profiles could not be achieved due to differences in protein content that were only discovered later. Applicants should not rely on the label claim of the RMP. Instead, an accurate determination of the protein concentration of multiple RMP batches should be performed. If such determinations are based on UV-visible spectroscopy, then the used extinction coefficient should be substantiated.

3.1.7.3. Biological activity

Demonstration of comparable bioactivity is of critical importance. If batches of the biosimilar candidate fail to meet the similarity ranges for biological activity, conclusion of biosimilarity is unlikely.

3.1.7.4. Charge variant analysis

Differences in the charge profile between a biosimilar and its RMP are not uncommon due to the many factors that can influence the overall charge profile of a biological medicinal product. Differences in charge profile could be acceptable where the applicant has conducted thorough analyses to determine the causes of these variations. Examples could include peak fractionation studies, where the acidic and basic fractions are purified and further analysed using an appropriate panel of physicochemical assays and biological assays. Any such supportive data should identify the relevant variants and provide convincing evidence that the identified differences will not have any clinically meaningful impact. For instance, where differences in charge are due to differences in C-terminal lysine clipping (a modification known from scientific literature not to impact safety and efficacy), applicants may provide data from samples treated with enzymes such as carboxypeptidase to provide supportive experimental evidence.

3.1.7.5. Glycosylation

Minor differences in the glycosylation pattern are commonly seen. Based on experience to date, certain differences between the biosimilar candidate and reference medicinal product can be challenging to justify, as such differences could lead to clinically relevant changes, especially for certain hormones, enzymes, and cytokines, and also for mAbs with Fc-effector functions. For example, a different level of afucosylation may impact effector function of a mAb, leading to a change in biological activity. Changes in high mannose species and sialylation might impact clearance and PK, and differences in non-human glycan epitopes such as α-galactose and N-glycolylneuraminic acid could impact immunogenicity. Applicants are strongly encouraged to evalute the glycosylation profile of the RMP during the early development of their biosimilar candidate. If glycosylation has the potential to affect efficacy, safety/immunogenicity, or PK, applicants should make every effort to closely align the glycosylation profile of the biosimilar with that of the RMP to minimise the risk of rejection of the claim of biosimilarity.

Where differences in glycosylation profile are unavoidable, a robust data package is expected to justify that this will not have an impact on efficacy or safety, including immunogenicity.

For biosimilar monoclonal antibodies with differing glycoprofiles where effector function is part of the MoA, a comprehensive panel of tests should be provided to show that differences in glycosylation do

not impact on efficacy or safety. In particular, differences in afucosylation and high mannose are considered of critical importance due to the potential impact on FcγRIII binding and ADCC activity of mAbs. The supportive data package should include at minimum the following:

- a comprehensive panel of Fc receptor binding assays, including relevant genotypic variants of FcγRIIa and FcγRIIIa;
- extensive data from ADCC assays - this usually requires more than one assay format to provide sufficiently convincing evidence e.g., data using different sources of effector cells such as PBMCs and NK cells, and/or using assays which more closely reflect the physiological situation e.g., using patient cells, inclusion of patient serum in the assay, or other relevant approaches;
- data on correlation between afucosylation, high mannose and ADCC to establish a correlation between the afucosylation/high mannose and ADCC is encouraged, where appropriate. In such cases, applicants should consider using experimentally generated samples which cover a wide range of afucosylation or high mannose. Such data may allow for a predictive approach where the release specifications for afucosylation/high mannose could be set to ensure that all commercial batches of the biosimilar would have comparable ADCC to the reference medicinal product. Such experimental approaches could be useful in addressing the residual uncertainty due to differences in the glycoprofile.

Ultimately, for mAbs where effector function is part of the MoA, if there are clear differences in ADCC or any other relevant Fc-functions (e.g. ADCP, CDC) between the biosimilar and the RMP, approval as a biosimilar may not be possible. In such cases, adapting the manufacturing process to achieve a similar glycoprofile should be pursued.

Certain products have complex glycosylation profiles and multiple N-linked and O-linked sites of glycosylation, for example some recombinant hormones and enzymes. For such products, differences in glycoprofile may preclude approval in the absence of a CES.

3.1.7.6. Impurities

Product-related impurities are inherent to biological medicines. For example, differences in aggregates may increase the likelihood of product immunogenicity. Where differences have been observed in impurity levels between a biosimilar and its RMP, experience has shown that further characterisation data may alleviate potential clinical safety and efficacy concerns. Such studies have included MoA studies performed with the individual impurities at levels beyond those observed during the analytical similarity study. Inclusion of batches of the RMP in such fractionation studies will strengthen the overall understanding of the structural properties of the molecule and thus help support the suitability of the data provided to substantiate the claim of biosimilarity. Complementary studies should be adequately designed to support any conclusions that the differences observed in the impurity profiles have no clinically meaningful impact. Prior knowledge based on data of other products or additional non-clinical studies can also be helpful in determining whether a particular impurity is a relevant safety concern. However, reducing impurities, in particular one(s) that are not present in the RMP to levels as low as technically reasonable, is always preferred over immunological characterisation because the latter is subject to high uncertainties in respect to predictability for the clinical situation. Comparative accelerated and/or stress stability studies can also be helpful in demonstrating comparable degradation profiles and kinetics.

3.2. Clinical

3.2.1. Utility and Limitations of Comparative Clinical Efficacy/Safety Trials

In the European Union Regulatory Framework for biosimilars, CES that also include supportive safety and immunogenicity data have historically played an important role. CES are intended to address remaining uncertainties regarding biosimilarity following the analytical comparison of a biosimilar candidate and its RMP, and to confirm equivalent clinical performance. They have typically been required in biosimilar developments, except for certain biologic molecules with low structural and functional complexity.

In general, state of the art analytical tools are sufficiently sensitive to detect differences between a biosimilar and its RMP, and CES may not add scientific knowledge that is essential for the conclusion of biosimilarity (Guillen *et al*., Kirsch-Stefan *et al*., Bielsky MC *et al*., IPRP workshop report 2024). Therefore, it is expected that CES will generally no longer be needed for biologicals whose physicochemical and functional properties can be well characterised and the results properly interpreted. CES, however, may still be important in cases where a biological is not well-characterisable and/or has an unknown or poorly understood MoA or structure-function relationship (see section 3.1.2). In such cases, the remaining uncertainties would need to be addressed by clinical data to ensure similar clinical performance.

In addition, a CES may still be required in scenarios that do not allow for a meaningful characterisation of PK, e.g., locally applied products with negligible systemic absorption.

3.2.2. The relevance of pharmacokinetic (PK) studies in biosimilar development

Comparative PK studies are still an essential part of a biosimilar development programme.

Generally, requirements for comparative PK studies outlined in guidelines (Guideline on similar biological medicinal products containing biotechnology-derived proteins as active substance: non-clinical and clinical issues, EMEA/CHMP/BMWP/42832/2005 Rev1) and (Guideline on similar biological medicinal products containing monoclonal antibodies – non-clinical and clinical issues, EMA/CHMP/BMWP/403543/2010) are expected to apply.

If evaluation of PK is not feasible (e.g. for heparins), relevant PD/biomarker endpoints should be measured instead (see 3.2.3 Pharmacodynamics).

Traditionally, PK studies have not been pivotal in answering questions related to efficacy, safety and immunogenicity in biosimilar development programmes. The objective has rather been to provide evidence for equivalent exposure between the biosimilar and the reference product as part of the totality of evidence supporting biosimilarity. For a PK trial in a tailored approach without a CES, the primary objective of the trial is still to provide evidence for equivalent exposure. Conclusions about the similar efficacy, safety and immunogenicity between the biosimilar and the reference product needs to be primarily inferred from the analytical comparability exercise. In addition, the PK trial can provide supportive comparative safety and immunogenicity data. (see 3.2.4 Safety and Immunogenicity)

3.2.3. Pharmacodynamics (PD)

In previous biosimilar developments, the availability of an accepted PD surrogate endpoint that could be investigated as part of the PK study was considered a prerequisite for waiving a CES. PD assessments offer insights into the biological effects of the biosimilar and may confirm its MoA and therapeutic potential. However, for a biosimilar molecule whose functional properties are well-

characteriseable with state of the art methods (see section 3.2.1), PD data in humans are not expected to add relevant scientific information. If PD markers are nevertheless measured, formal statistical equivalence testing is generally not needed, however results should reasonably support a conclusion of biosimilarity. In exceptional cases, where PK measurements are not feasible and meaningful PD markers are available, comparative evaluation of PD with predefined equivalence margins and appropriate sample size calculation would be required (see also 3.2.2). It should be considered that PD endpoints may not be meaningfully interpretable in healthy volunteers.

3.2.4. Safety and Immunogenicity

For biologicals whose physicochemical and functional properties can be well-characterised, the conclusion on similar clinical performance largely rests on the analytical and pharmacokinetic comparison. However, supportive safety and immunogenicity data should generally be collected as part of the comparative PK study/ies and analysed descriptively. Regulatory experience suggests that the ADA development in comparative PK studies corresponds well with that observed in CES (Ping *et al.*, 2024; Kurki *et al.,* 2021).

In some cases, immunogenicity data from a single-dose PK study may not be sufficient, especially if ADAs are known to exert relevant effects on efficacy (e.g., due to neutralising antibodies) or safety (e.g., serious infusion reactions) but develop later in the treatment course. In such cases, PK studies and concurrent ADA measurements may have to be extended to more than one administration in an appropriate healthy volunteer or patient population. The applicant should, as far as possible, assess the immunogenic risk of the RMP and the timeframe of ADA development to design a comparative PK study of adequate duration.

3.3. Conclusion

Based on the advancements in analytical technology and the regulatory experience gained, a tailored approach for clinical development of biosimilar candidates is possible. CES are no longer expected to be required for approval of biosimilars that can be thoroughly characterised using state-of-the-art analytical methods and have demonstrated similarity in physicochemical and functional properties. Comparative clinical PK studies are still essential elements in the biosimilar development and can provide supportive safety and immunogenicity data. This tailored clinical approach is expected to be applicable for the majority of biosimilar candidates. A regulatory option that, under certain conditions, allows authorisation of biosimilars based on demonstrated comparability at the analytical level with a limited clinical data package streamlines the development process without compromising efficacy and safety.

4. References

Relevant EU and International guidelines on biosimilars development.

Overarching biosimilar guidelines

Guideline on similar biological medicinal products, CHMP/437/04 Rev 1, 23 October 2014.

Guideline on similar biological medicinal products containing biotechnology-derived proteins as active substance: non-clinical and clinical issues, EMEA/CHMP/BMWP/42832/2005 Rev1, 18 December 2014.

Guideline on similar biological medicinal products containing biotechnology-derived proteins as active substance: quality issues (revision 1), EMA/CHMP/BWP/247713/2012, 22 May 2014.

Product-specific biosimilar guidelines

Guideline on similar biological medicinal products containing monoclonal antibodies – non-clinical and clinical issues, EMA/CHMP/BMWP/403543/2010, 30 May 2012.

Other guidelines relevant for biosimilars

Guideline on comparability of biotechnology-derived medicinal products after a change in the manufacturing process - non-clinical and clinical issues, EMEA/CHMP/BMWP/101695/2006, 19 July 2007.

Guideline on Immunogenicity assessment of biotechnology-derived therapeutic proteins, EMEA/CHMP/BMWP/14327/2006 Rev 1, 18 May 2017.

Guideline on immunogenicity assessment of monoclonal antibodies intended for *in vivo* clinical use, EMA/CHMP/BMWP/86289/2010, 24 May 2012.

Reflection paper on statistical methodology for the comparative assessment of quality attributes in drug development (EMA/CHMP/138502/2017).

Commission Regulation (EC) No 1234/2008 of 24 November 2008 concerning the examination of variations to the terms of marketing authorisations for medicinal products for human use and veterinary medicinal products.

ICH guideline Q5E on Comparability of Biotechnological/Biological Products (CPMP/ICH/5721/03).

ICH guideline Q8 on Pharmaceutical Development (EMA/CHMP/ICH/167068/2004)

ICH guideline Q9 on Quality Risk Management (EMA/CHMP/ICH/24235/2006).

ICH guideline Q10 on Pharmaceutical Quality System (EMA/CHMP/ICH/214732/2007)

Clinical pharmacology and pharmacokinetics: questions and answers | European Medicines Agency (EMA)

Relevant Literature References

R Martijn van der Plas, Marcel H N Hoefnagel, Hans L Hillege, Kit C B Roes; Pragmatic rules for comparability of biological medicinal products. Biologicals 2020 Jan:63:97-100. https://doi.org/10.1016/j.biologicals.2019.11.002

Pekka Kurki, Sean Barry, Ingrid Bourges *et al*. Safety, Immunogenicity and Interchangeability of Biosimilar Monoclonal Antibodies and Fusion Proteins: A Regulatory Perspective. Drugs. 2021; 81(16): 1881–1896.

Guillen, E., Ekman, N., Barry, S., Weise, M. and Wolff-Holz, E. (2023), A Data Driven Approach to Support Tailored Clinical Programs for Biosimilar Monoclonal Antibodies. Clin Pharmacol Ther, 113: 108-123. https://doi.org/10.1002/cpt.2785.

Kirsch-Stefan, N., Guillen, E., Ekman, N. *et al*. Do the Outcomes of Clinical Efficacy Trials Matter in Regulatory Decision-Making for Biosimilars?. BioDrugs 37, 855–871 (2023). https://doi.org/10.1007/s40259-023-00631-4

Marie-Christine Bielsky, Anne Cook, Andrea Wallington, Andrew Exley, Shahin Kauser, Justin L. Hay, Leonard Both, David Brown, Streamlined approval of biosimilars: moving on from the confirmatory efficacy trial. Drug Discovery Today. Volume 25, Issue 11, 2020, Pages 1910-1918, ISSN 1359-6446, https://doi.org/10.1016/j.drudis.2020.09.006.

IPRP workshop report 2024: https://admin.iprp.global/sites/default/files/2024-07/IPRP_BWG_Final%20IPRP%20Scientific%20Workshop%20Summary%20Report_2024_0506.pdf

FDA 2025 "Development of Therapeutic Protein Biosimilars: Comparative Analytical Assessment and Other Quality-Related Considerations Guidance for Industry" FDA-2019-D-2102.

WHO 2022 "Guidelines on evaluation of biosimilars". https://www.who.int/publications/m/item/guidelines-on-evaluation-of-biosimilars

5. List of Abbreviations

Abbreviation	Definition
ADCC	Antibody-dependent cellular cytotoxicity
BMWP	(EMA CHMP) Biosimilar Medicinal Products Working Party
BWP	(EMA CHMP) Biologics Working Party
CDR	Complementarity Determining Region
CES	Comparative Efficacy Studies
CHMP	(EMA) Committee for Medicinal Products for Human Use
CQA	Critical Quality Attribute
MAA	Marketing Authorisation Application
MoA	Mechanism of Action
PK	Pharmacokinetics
PD	Pharmacodynamics
RMP	Reference Medicinal Product
WHO	World Health Organization
WP	Working Party