17 March 2025
EMA/CHMP/BMWP/60916/2025
Committee for Medicinal Products for Human Use (CHMP)

Reflection paper on a tailored clinical approach in biosimilar development

Draft

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	1 April 2025
End of consultation (deadline for comments)	30 September 2025

Comments should be provided using this EUSurvey form. For any technical issues, please contact the EUSurvey Support.

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 (finally resulting in efficacy and 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 prerequisites, analytical comparability exercises and pharmacokinetic (PK) data could be sufficient for demonstrating biosimilarity.

This reflection paper will examine settings for biosimilars where similar clinical efficacy and safety can be inferred from a conclusion of physicochemical and biological similarity and comparable pharmacokinetics. Currently, Comparative Efficacy Studies (CES) (in which safety and immunogenicity data are also routinely captured) can already be waived in case an accepted pharmacodynamic (PD) surrogate endpoint exists, but even this prerequisite might not be needed.

A further driver for this Reflection paper is the regulatory experience indicating that the results from the CES in the past generally did not add relevant additional information to the biosimilarity exercise (Guillen *et al*., Kirsch-Stefan *et al*., Bielsky MC *et al*., IPRP workshop report 2024).

In addition, trends are observed regarding the types of biological medicinal products losing market exclusivity, where feasibility of performing comparative efficacy trials appears limited. This is firstly due to originator products having narrow indications with small number of patients as well as originator products being used in increasingly complex add-on therapy settings.

Taken together, a regulatory option that, under certain prerequisites, allows authorisation based on demonstrated comparability at the quality level with a limited (tailored) clinical data package (based on a comparative PK trial) would provide a viable path forward for approving biosimilars with less clinical data.

Based on the points outlined above, a tailored approach for clinical development of biosimilar candidates can be envisioned. In certain cases, CES may no longer be required for approval of biosimilars that can be thoroughly characterised and have shown high similarity on an analytical and *in vitro* pharmacology level. Comparative clinical pharmacokinetic studies are still essential elements in biosimilar development but some adjustments to the data requirements, such as inclusion of immunogenicity parameters and/or modifying the study design (e.g., one-dose vs multiple-dose), could be considered.

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, including *in vitro* pharmacology, and consider their predictive value. Subsequently, some reflections will be provided with regard to the contribution of CES, and other human *in vivo* studies, especially PK/PD studies, and to the assessment of immunogenicity.

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

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 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 proven to be useful and successful over 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 to take into consideration quality differences (in other words, it does not impose that products should be identical) as long as they do not translate into significant clinical differences. This concept has been used for instance to support the implementation of necessary manufacturing process changes for biological products, without imposing that products should be identical in a physicochemical sense, which may not be achievable or requiring the conduct of unnecessary comparative clinical studies. Since the 1990s, major manufacturing changes have been substantiated and implemented based on a comparability exercise, and without comparative efficacy trials. This includes situations such as replacing a product's Master Cell Bank, a situation that is from a scientific viewpoint comparable to the development of a biosimilar product.

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 (RMP) 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 these slight differences have no impact on clinical safety or efficacy. However, large differences in QAs are not compatible with the biosimilar approach, and such a situation cannot be remedied by clinical data. Consequently, a CES is not intended to justify the presence of large differences, but to address residual uncertainty after the evaluation of Quality and Non-clinical data, should it exist. As explicitly stated in the Guideline on Similar Biological Medicinal Products (CHMP/437/04): '*The aim of clinical data is to address slight differences shown at previous steps and to confirm comparable clinical performance of the biosimilar and the reference medicinal product. Clinical data cannot be used to justify substantial differences in quality attributes. However, in case the mechanism of action (MoA) and structure-function relationship is not sufficiently understood, a CES might still be needed*.'

A comprehensive set of relevant QAs providing detailed information regarding the structural 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 ICH Q9 guideline (EMA/CHMP/ICH/24235/2006)). Prior knowledge provides understanding of the critical QAs (CQAs) impacting the interaction with receptor(s) (including membrane receptors, ligands, substrates, and other targets). These interactions form the basis of subsequent biological effects, i.e., pharmacology, toxicology, PK/PD, etc. 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 and *in vitro* pharmacology during risk assessment in terms of potential impact on PK/PD, efficacy and safety, including immunogenicity, becomes 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 should be developed to ensure that a biosimilar product which is highly similar to the RMP can be consistently produced. A robust manufacturing control system and demonstrated batch-to-batch consistency of the biosimilar are prerequisites for a successful similarity assessment and ensure that batches that do not fulfil pre-determined specifications are rejected and do not reach the patient. The overall manufacturing control system will therefore ensure consistency of the quality profile of commercially manufactured batches and high similarity between the QAs of the commercial biosimilar and the RMP.

In summary, the following prerequisites allow for a successful comparability exercise, which is fundamental for the approach outlined in this reflection paper:

- comprehensive knowledge regarding the molecule's MoA is available;
- detailed characterisation of the structure and functionally relevant QAs is possible using orthogonal and state-of-the-art analytical methods;
- functional assays (*in vitro* pharmacology 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.

- 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 protocol (see section 3.1.3)

These prerequisites will support that an analytical comparability exercise, expanded with *in vitro* pharmacology data, and data from human PK studies, as appropriate, is able to assure similarity of the biosimilar to its RMP. This similarity implies that there are no meaningful differences in structure and other QAs, that interactions with relevant receptors/targets are comparable, and therefore that comparable efficacy and safety can be inferred.

3.1.3. Similarity assessment protocol

In general, product development is an iterative process. Development starts based on 'prior knowledge'; information which is collected during the development process is used to amend, focus, and fine-tune that development process, and to define more precisely which specific studies are needed, and how these specific studies should be conducted. A clear plan for the development activities should be available.

In order to increase the overall robustness of any biosimilar development programme, it is essential that a similarity assessment protocol is developed and documented prior to the initiation of the pivotal similarity studies (i.e., the final assessment of similarity). Adequate consideration should be given to section 6 of the Reflection Paper on statistical methodology for the comparative assessment of quality attributes in drug development (EMA/CHMP/138502/2017). The protocol should capture all critical parts of the analytical and functional similarity assessment, such as:

- the number of RMP batches to be included and the sampling plan;
- the number and nature of biosimilar batches (primarily batches manufactured using the commercial manufacturing process and scale);
- 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;
- justification of the similarity condition and acceptance criteria/ranges to be applied, as well as the overall approach planned for 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 *in vitro* pharmacology/biological assays (e.g., receptor binding assays and cell-based potency assays);
- a sufficiently detailed plan on the handling and consequences of potential differences (e.g., biosimilar batches which fail to meet the established similarity criteria);
- a discussion on why a tailored clinical development approach is considered applicable for the biosimilar under development; whether it can be assumed based on scientific knowledge that similarity demonstrated for critical QAs will ensure the desired clinical performance.

Applicants are strongly recommended to make use of the EMA scientific advice procedure to present and to reach agreement on their similarity assessment protocol before starting the pivotal similarity assessment.

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. However, developers should take into account that the use of development batches can introduce uncertainties in the evidence for analytical and functional similarity.

A sufficient number of biosimilar batches needs to be tested. Usually, 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 protocol.

Although it is impossible to specify the exact number of RMP batches needed for every product and scenario, experience has shown that 15-30 batches of RMP are generally appropriate, 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.). The RMP batches are expected to be stored under recommended (label) conditions and tested within their shelf life. Any exception to this has to be fully substantiated with experimental data. Age at the time of testing (relative to expiry date) should be considered. 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 *in vitro* pharmacology (e.g., receptor binding studies, cell-based potency assays) are available, both to support the identification of physicochemical QAs, and to provide comparative data between biosimilar and its reference medicinal product. Such comparative *in vitro* pharmacology 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) and the Reflection paper on statistical methodology for the comparative assessment of quality attributes in drug development (EMA/CHMP/138502/2017).

Currently, the most widely used approach for demonstrating physicochemical and functional similarity is to show that the biosimilar developer is able to manufacture a biosimilar candidate having all relevant QAs within the batch-to-batch variability of the RMP. The manufacturing control system, including batch release testing for the most critical QAs, ensures that the quality profile of future biosimilar batches remains similar to the batches tested for similarity, as well as to the RMP. Any biosimilar batches released within the batch-to-batch variability of the RMP are expected to have the same clinical performance, and differences within the ranges are assumed not to have a relevant impact on safety or efficacy.

This approach 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 similar approaches are referenced in the FDA 2019 guideline on "Development of Therapeutic Protein Biosimilars: Comparative Analytical Assessment and Other Quality-Related Considerations Guidance for Industry" and the WHO 2022 "Guidelines on evaluation of biosimilars".

A number of novel statistical approaches for demonstrating physicochemical and functional similarity have been proposed in the literature. Applicants are encouraged to explore the possibilities of making use of these and other statistical approaches where relevant.

3.1.6.1. Similarity condition

A similarity condition is a concise description for when two data distributions allow a conclusion of similarity (EMA/CHMP/138502/2017). For most QAs, it is feasible to establish their criticality based on prior knowledge of their structure-function relationship. However, defining similarity conditions based on the maximum allowed difference between the two underlying data distributions for the specific QA purely based on clinical performance is difficult, as a clear correlation between the quantitative level of individual QAs and the clinical performance is usually lacking. The final clinical performance of a molecule is a result of several QAs; therefore, similarity between the biosimilar and the RMP needs to be considered holistically, using a set of orthogonal methods.

For QAs with a continuous scale of measurement, a "population in population" approach will, to a large extent, overcome the difficulties in determining and justifying the allowed differences in the underlying distributions. For the population in population approach, the similarity condition is defined as a pre-determined portion of the biosimilar population that should be within a prespecified population portion of the reference medicinal product.

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. Finally, comparisons of QAs with a nominal scale of measurement or comparisons against an expectation (primary amino acid sequence), as well as visual comparisons of e.g., chromatograms, are not compatible with the population in population approach. It is noted that for such QAs, a similarity condition has not always been well-defined. To avoid this, the applicant is recommended to describe in the similarity assessment protocol the conditions for similarity of all types of QAs, not only for those QAs having a continuous scale of measurement.

3.1.6.2. Similarity criteria

For making decisions on whether the similarity condition (such as "population within population") is fulfilled, a similarity criterion is needed. Ideally, the choice of this similarity criterion should be based on its operating characteristics, i.e., the probability of false positive decisions (which is of main interest from a regulatory point of view) and the probability of true positive decisions. The criticality of the QA could also be considered when selecting the similarity criterion. For more details, refer to Reflection paper on statistical methodology for the comparative assessment of quality attributes in drug development (EMA/CHMP/138502/2017).

3.1.7. Uncertainties in the similarity assessment

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, differences which directly impact the MoA, or which could lead to an altered safety profile, are not compatible with the biosimilarity concept.

Where the similarity criteria for all QAs and prerequisites formulated in Section 3.1.3 are fulfilled, tailoring or reduction of the pivotal CES could be justified. However, in practice, the probability of differences in at least one QA not only depends on the variability in analytical method and the magnitude of acceptable differences between products for each individual attribute but also with the number of QAs tested (multiplicity). In addition, a real difference in one or more QAs could be present. Consequently, an expectation that similarity criteria are met for all QAs could require infeasibly large numbers of independent batch samples from both the reference medicinal product and the biosimilar candidate.

Therefore, the fact that some data points fail to meet similarity criteria (e.g., fall outside the biosimilarity range) for some QAs does not *a priori* preclude approval as a biosimilar, nor does it invalidate the use of a tailored clinical development programme with limited or no CES. Nonetheless, since biosimilars are approved based on the totality of data, the availability of CES data has added supportive weight to assuage any remaining uncertainties in the quality package. In the absence of CES, the presence of (minor) differences may increase the overall uncertainty, which needs to be considered in the conclusion on biosimilarity. If the similarity criteria 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 reference medicinal product. Otherwise, a supportive CES may be necessary to provide sufficient assurance that the clinical performance of the biosimilar is comparable to the reference medicinal product. However, CES cannot be used to justify substantial differences in QAs.

For attributes that fail to meet similarity criteria, 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. The approach for addressing CQAs for which similarity criteria cannot be met should be pre-specified in the similarity assessment protocol as far as possible, to avoid reliance on post-hoc justifications of differences. 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 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 differences are due to analytical variability or true batch-to-batch variability should be carefully considered in any discussion of analytical data. Where applicants consider that analytical variability is the underlying reason for anomalous results, every effort should be made to improve the precision of results, for example by multiple repeat testing of the same batch/sample. Unsubstantiated claims about analytical variability would not be sufficient in this respect. It is also important to recognise that differences detected using a sensitive assay typically cannot be overcome by providing supportive data from a less sensitive assay. 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., reanalysis of the batches at different time points, 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 fail to meet similarity criteria, characterisation data using orthogonal assays can provide supportive evidence. The final data package should be such that residual uncertainty does not hamper the benefit/risk decision. 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 reference medicinal product. There are several approaches which could be included in the pre-specified similarity assessment protocol to address the situation where unanticipated differences in QAs are found. This may help to avoid rejection of the application or the need to carry out confirmatory CES. 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 confirm binding affinity of 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 safety. Even minor differences in higher order structure can have significant implications for the biosimilarity claim. High resolution structural analysis is needed to characterise any small changes in conformation that may result in potential differences in efficacy. 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 any that differences observed have no impact on binding 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 appropriate (see EMA Clinical pharmacology and pharmacokinetics: questions and answers for further details).

CHMP has encountered several examples where PK trials were conducted using a batch(es) of biosimilar where the protein concentration was subsequently found to be slightly different from the RMP. In some cases, this led to difficulties in demonstrating a comparable PK profile. Applicants should not rely on the label claim of the RMP; instead, the extinction coefficient of the biosimilar and the RMP should be experimentally determined early in development in order to make an accurate determination of the true protein concentration. Reference to a published extinction coefficient of the RMP is not considered sufficient.

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 criteria for biological activity, conclusion of biosimilarity is unlikely. Nonetheless, there may be scenarios where a panel of orthogonal assays can be used to interrogate biological activity. In such cases, a minor difference in a single assay might not preclude approval in the absence of CES; however, as for any QA, such a scenario would need to be appropriately justified *a priori* in the similarity assessment protocol.

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 clearly explain 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. Applicants may provide data from samples treated with enzymes such as carboxypeptidase to provide supportive experimental evidence.

3.1.7.5. Glycosylation

Based on experience to date, differences in glycosylation 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 consider the glycosylation profile of the RMP during the early development of their biosimilar candidate and make every effort to closely match the biosimilar with this glycosylation profile to minimise the risk of rejection of the claim of biosimilarity, particularly in the absence of CES.

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; this should be outlined in the similarity assessment protocol.

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 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 with effector functions where there are clear differences in ADCC or any other relevant Fc-functions between the biosimilar and the RMP, approval as a biosimilar may not be possible. In such cases, adapting the manufacturing process to achieve a more consistent glycoprofile should be pursued.

Products such as recombinant hormones and enzymes may have complex glycosylation profiles and multiple N-linked and O-linked sites of glycosylation. 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 and other size variants may increase the likelihood of product immunogenicity. Where differences have been observed in impurity levels, experience has shown that further characterisation data has been sufficient in many cases to 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. Data from prior knowledge of other products or additional non-clinical or PK studies can also be helpful in determining whether a particular impurity is a relevant safety concern. However, reducing a novel impurity (i.e., one that is not present in the reference medicinal product) 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.

3.1.8. Final reflection on Quality aspects

Reconsidering the need for a CES may be possible if the Quality data package provides solid evidence for similarity. Such a data package entails at least:

- MoA and structure-function relationship(s) are well understood; as a consequence, CQAs are well known and can be reliably assessed in a Quality Risk assessment.
- Sensitive analytical methods are used.
- Sufficient number of RMP batches have been characterised to properly estimate batch-to-batch variability. The quality target product profile (QTPP) applied during development of the biosimilar is strongly linked to the variability seen in the RMP.
- The similarity assessment has been preplanned and is captured in the similarity assessment protocol. Similarity conditions and similarity criteria are defined and applied for the assessment.

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. While 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.

It is reminded that a CES should not be used as a substitute for incomplete Quality development, or to justify the presence of large differences in QAs.

3.2. Clinical

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

In the European Union Regulatory Framework for biosimilars, Comparative Clinical Efficacy Studies (CES) that also include supportive safety data have historically played an important role. CES are intended to address uncertainties regarding biosimilarity following the analytical comparison of a biosimilar candidate and its reference medicinal product, 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.

As addressed in the sections above, in general, analytical tools are considered sensitive enough to detect differences between a biosimilar and its reference medicinal product, and CES may not add essential scientific knowledge in the decision for biosimilar approval (Guillen *et al*., Kirsch-Stefan *et al*., Bielsky MC *et al*., IPRP workshop report 2024). CES, however, may still be important in cases where a biological is not well-characterisable and/or has an unknown or poorly understood MoA, structure-function relationship, or if the impact of observed differences on clinical outcomes is unclear. In such cases, it would be challenging to fully rely on comparative analytical data for the demonstration of similar efficacy and safety. The above mentioned criteria would inhibit an assumption of similar clinical performance to the originator based on quality and PK alone, as quality comparability would be associated with a higher degree of uncertainty.

In addition, a CES would 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, in combination with a comprehensive analytical comparison, are essential elements of a biosimilar development.

Generally, data 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) apply, but some adjustments may be necessary to fully complement the analytical data and together form an acceptable amount of evidence to conclude on biosimilarity (see section: Safety and Immunogenicity). If available, relevant PD endpoints, especially those reflecting the MoA of the biological, may be added to the PK study to address minor differences in quality attributes related to the MoA of observed *in vitro* and to further strengthen a conclusion of biosimilarity (see section: Pharmacodynamics).

Traditionally, PK studies have not been pivotal in answering questions related to safety and immunogenicity in biosimilar development programmes. These aspects have instead been addressed as part of the CES. The main reasons are that the sample size of the PK trial is usually small and that the trial duration is short in comparison with a CES. Therefore, the PK trial in a tailored approach without a CES will not be able to draw robust conclusions about the overall safety profile, and similar safety mainly needs to be inferred from a thorough analytical comparability exercise and similar PK and potentially PD profiles.

It is envisaged that in a tailored approach the comparative PK trial will be adapted to address residual uncertainty regarding comparability in exposure as well as safety and immunogenicity.

In cases where biosimilarity cannot be robustly concluded from state-of-the-art comparative analytical and PK studies and where accepted surrogate PD endpoints are not available, CES are still required. All the considerations regarding study design, endpoints and extrapolation laid down in the Guideline on similar biological medicinal products containing biotechnology-derived proteins as active substance: non-clinical and clinical issues () should be complied with unless justified, in case a CES is needed.

3.2.3. Pharmacodynamics (PD)

In biosimilar development, the evaluation of an accepted PD surrogate endpoint has previously been considered an integral part to waive a CES. PD assessments offer insights into the biological effects of the biosimilar and may confirm its MoA and therapeutic potential. However, when the structure and function of the molecule in question are well-characterised and shown to be highly similar, the necessity of PD comparability is debatable.

For a biosimilar molecule that is well-characterised and shown to be similar on a quality level, demonstrating comparable structural and functional attributes to the RMP, demonstration of PD comparability may not be needed. The extensive characterisation at the quality level ensures that the biosimilar mirrors the reference medicinal product closely, diminishing the need for PD evaluations.

Nonetheless, even if not essential, PD comparability evaluations may provide additional layers of confidence and assurance in the biosimilar's clinical performance. If relevant PD endpoints can be easily measured within the PK study, applicants are encouraged to include them. If an equivalence criterion has to be fulfilled also for the PD endpoints, this needs to be considered in the sample size calculation of the PK/PD study. It should be considered that PD endpoints may not be meaningfully interpretable or sensitive enough in healthy volunteers.

The acceptability of a tailored clinical approach should mainly depend on the product understanding, including the MoA and the ability to extensively characterise the structure and function of the molecule, rather than the availability of meaningful PD markers.

3.2.4. Safety and Immunogenicity

While comparative PK studies primarily focus on establishing equivalence in drug exposure between the biosimilar and the reference medicinal product, they can also provide supportive safety and immunogenicity data that help ascertain similarity in immunological responses between the biosimilar and the reference medicinal product. In cases with a comprehensive quality package showing close analytical similarity and high purity of the biosimilar, a limited but well-defined set of comparative safety and immunogenicity data as part of the PK study could provide sufficient confidence in the biosimilar's safety and immunogenicity profile. If relevant uncertainties remain from the quality package, longer and/or larger studies may be needed to ensure the absence of a clinically relevant impact.

In case relevant uncertainties remain, longer and/or larger studies may be needed to ensure no clinical meaningful impact (see also 3.1.7.5, 3.1.7.6.).

3.2.4.1. Extended PK studies with more than one dosing

In some cases, immunogenicity data from a single-dose PK study may not be enough,. especially if anti-drug antibodies (ADAs) are known to exert relevant effects on efficacy (e.g., due to neutralising antibodies) or safety (e.g., serious infusion reactions) developing later in the treatment course. In such cases, two or even more administrations may be necessary in an appropriate healthy volunteer or patient population. The applicant should assess the timeframe of ADA development and the immunogenic risk of the reference medicinal product to design a comparative PK study of adequate duration.

3.3. Conclusion

Taken together, biosimilars may be approved without providing CES or even PD data if similar clinical efficacy and safety pharmacology can be inferred from a sufficiently stringent evaluation of analytical comparability, *in vitro* pharmacology, and a comparative clinical PK trial. Whether a development programme without a CES could be envisaged depends on the ability to extensively characterise the structure and function of the RMP, and understanding whether the differences in particular QAs have a meaningful impact on clinical outcomes (see prerequisites for similarity assessment).

In any case, a well-defined comparative human PK study would still be required.

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 Q9 on quality risk management (EMA/CHMP/ICH/24235/2006).

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

Relevant Literature References

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

WHO 2022 "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