World Health Organization

WHO/BS/2016.2290
ENGLISH ONLY

# Guidelines on evaluation of monoclonal antibodies as similar biotherapeutic products (SBPs)

NOTE:

This document has been prepared for the purpose of inviting comments and suggestions on the proposals contained therein, which will then be considered by the Expert Committee on Biological Standardization (ECBS). Publication of this early draft is to provide information about the proposed Guidelines on evaluation of monoclonal antibodies as similar biotherapeutic products (SBPs) to a broad audience and to improve transparency of the consultation process.

**The text in its present form does not necessarily represent an agreed formulation of the Expert Committee. Written comments proposing modifications to this text MUST be received by 16 September 2016 in the Comment Form available separately** and should be addressed to the World Health Organization, 1211 Geneva 27, Switzerland, attention: Department of Essential Medicines and Health Products (EMP). Comments may also be submitted electronically to the Responsible Officer: Dr Kai Gao at email: gaok@who.int.

The outcome of the deliberations of the Expert Committee will be published in the WHO Technical Report Series. The final agreed formulation of the document will be edited to be in conformity with the "WHO style guide" (WHO/IMD/PUB/04.1).

# Contents

# 1. Introduction

Monoclonal antibodies (mAbs) are a major class of recombinant deoxyribonucleic acid (rDNA) technology-derived biotherapeutic products that have achieved outstanding success in treating many life-threatening and chronic diseases. Some of these targeted therapy products are ranked in the top 10 lists of annual global pharmaceutical revenue successes. As patents and data protection measures on mAb products have expired, or are nearing expiry, considerable attention has turned towards producing similar biotherapeutic products (SBPs, also termed biosimilars) of the approved mAb innovator products with a view to making more affordable products that may improve global access to these so-called "blockbusters".

mAbs for therapeutic use are preparations of an immunoglobulin or a fragment of an immunoglobulin with specificity for a target ligand and derived from a single clone of cells. Each full-length molecule of an mAb consists of two heavy and two light polypeptide chains which are linked by disulfide bonds. mAbs have several possible functional domains within a single molecule. The defined specificity of a mAb is based on the binding region for an antigen that is located in the Fab part. For full-length mAbs, their Fc region has the ability to bind to specific receptors, leading to immune effector functions such as antibody-dependent cellular cytotoxicity (ADCC) and complement-dependent cytotoxicity (CDC), etc. Full-length monoclonal antibodies are glycoproteins with glycosylation sites in the Fc portion of the heavy chains, with further glycosylation sites possible depending on the type of molecule. Therefore, mAbs are highly complex biological macromolecules with size and charge variants, various post-translational modifications, including different glycosylation patterns, N and C terminal heterogeneity, long half-lives and the potential for inducing immunogenicity. Each individual mAb may therefore present a unique profile, which must be taken into consideration during the evaluation of these products as biosimilars.

WHO's *Guidelines on evaluation of similar biotherapeutic products (SBPs)* were adopted by the Expert Committee on Biological Standardization (ECBS) in 2009 *(1).* This document provided the scientific principles, including the stepwise approach, which should be applied for demonstration of similarity between the SBP and the reference biotherapeutic product (RBP). High similarity at the quality level is regarded as a prerequisite for enabling the use of a tailored nonclinical and clinical data set for licensure. The goal of the clinical comparability exercise is to confirm the similarity established at previous stages of development and to demonstrate that there are no clinically meaningful differences between the SBP and the RBP, and not to re-establish safety and efficacy because this has been done already for the RBP. The decision on licensure of the SBP should be based on evaluation of the totality of evidence from quality nonclinical and clinical parameters. It should be noted that clinical studies cannot be used to resolve substantial differences in physicochemical characteristics and biological activity between the RBP and the

SBP. If substantial differences in quality attributes are present, a stand-alone licensing approach may be considered.

The set of globally acceptable key principles outlined above on the regulatory evaluation and licensing of SBPs has served well as a basis for setting national requirements for SBPs. However, because of the structural complexity and heterogeneity of mAbs, the quality attributes of mAbs can vary from product to product. Furthermore, one mAb product may have multiple indications. Therefore, comparability studies between a candidate biosimilar mAb and a reference product mAb are challenging for both developers and regulators. Consequently, in 2014 WHO was requested to update the SBP guideline to take into account the technological advances in the characterization of rDNA-derived products, and particularly mAbs. In response, WHO organized an informal consultation in 2015 on the possible amendment of the 2009 SBP guideline with an additional focus on similar biotherapeutic products containing mAbs. The outcome of these discussions was that there was no need to revise the main body of the existing WHO guideline on SBPs. All participants, including national regulatory agencies (NRAs) and industry, recognized and agreed that the evaluation principles described in WHO's SBP guideline are still valid, valuable and applicable for facilitating harmonization of SBP requirements globally. However, it was also agreed that there was a need for additional guidance, rather than an amendment, for the evaluation of biosimilar mAbs.

## 2. Scope

The intention of this class-specific document is to provide special considerations for the evaluation of mAbs developed as similar biotherapeutic products. This WHO guidance covers rDNA-derived biosimilar mAbs used in the treatment of human diseases. The principles discussed in this document also apply to mAb-derived proteins such as mAb fragments and Fc fusion proteins.

From a regulatory perspective, mAb assessment is based on the same principles as those used for the evaluation of other rDNA-derived biotherapeutic proteins. On the other hand, mAb SBP should also comply with the criteria established for demonstration of similarity. Therefore this document should be read in conjunction with WHO's existing *Guidelines on the quality, safety, and efficacy of biotherapeutic protein products prepared by recombinant DNA derived technology* and its' *Guidelines on evaluation of similar biotherapeutic products (SBPs) (1, 2).*

Guidance on various aspects of rDNA-derived medicines, SBPs and mAbs are also available from several other bodies. This WHO guideline is not intended to conflict with, but rather to complement, existing relevant regulatory documents.

## 3. Glossary

**American college of rheumatology 20% improvement criteria (ACR 20):** A combined index that measures disease activity in patients with rheumatoid arthritis, which means at least 20% improvement in both the tender joint count and the swollen joint count and at least 20% improvement in 3 of 5 other score set measures.

**Antibody-dependent cellular cytotoxicity (ADCC):** An immune mechanism through which Fc receptor-bearing effector cells can recognize and kill antibody-coated target cells expressing tumour- or pathogen-derived antigens on their surface.

**Anti-drug antibodies (ADAs):** Antibodies generated by unwanted immunogenicity through an immune response by an organism against a therapeutic antigen (recombinant protein, or monoclonal antibody). This reaction leads to production of ADAs which may or may not inactivate the therapeutic effects of the treatment and, in rare cases, induce adverse effects.

**Area under the curve (AUC):** The area under the curve in a plot of concentration of drug in blood plasma against time.

**$AUC_t$:** The area under the concentration-time curve of drug in blood plasma from zero up to a definite time t.

**$AUC_{tau}$:** The area under the plasma concentration-time curve during a dosage interval.

**Biological activity:** The specific ability or capacity of a product to achieve a defined biological effect.

**Biosimilar mAb:** A monoclonal antibody product that is similar in terms of quality, safety and efficacy to an already licensed reference product.

**$C_{max}$:** The maximum (peak) serum or plasma concentration observed that a drug achieves in a tested area after the drug has been administrated and prior to the administration of a second dose.

**$C_{min}$:** The minimum (peak) serum or plasma concentration observed that a drug achieves in a tested area after the drug has been administrated and prior to the administration of a second dose.

**$C_{trough}$:** The measured serum or plasma concentration of a drug in a tested area at the end of a dosing interval at steady state prior to the administration of the next dose.

**Comparability:** The exercises that are performed after completing process changes or manufacturing site changes to show that a comparable product is being produced pre- and post-change.

**Complement-dependent cytotoxicity (CDC):** The immune process by which an antibody–antigen complex activates complement that ultimately results in the formation of a terminal lytic complex that is inserted into a cell membrane, resulting in lysis and cell death.

**Complete response (CR):** The disappearance of all signs of cancer in response to treatment. This does not always mean the cancer has been cured.

**Disease activity score in 28 joints (DAS28):** A combined index that measures disease activity in patients with rheumatoid arthritis, which assesses the number of swollen and tender joints and also the erythrocyte sedimentation rate, indicating how active the rheumatoid arthritis is at this time.

**Equivalence margin:** A pre-specified value in the equivalence trials, which is the largest difference that can be judged as being clinically acceptable and which should be smaller than differences observed in superiority trials of the active comparator.

**Equivalence trial:** A trial with the primary objective of showing that the response to two or more treatments differs by an amount which is clinically unimportant. This is usually demonstrated by showing that the true treatment difference is likely to lie between a lower and an upper equivalence margin of clinically acceptable differences.

**Mechanism of action (MoA):** The specific biochemical interaction through which a product produces its pharmacological effect.

**Monoclonal antibody (mAb):** Antibody derived from a single clone of cells.

**Non-inferiority trial:** A trial with the primary objective of showing that the response to the investigational product is not clinically inferior to that of a comparative agent (active or placebo control).

**Overall response rate (ORR):** The overall percentage of patients whose cancer shrinks or disappears after treatment; this includes the rate of CR (complete response) and PR (partial response).

**Potency:** The quantitative measure of biological activity based on the attribute of the product which is linked to the relevant biological properties and is expressed in units.

**Similarity:** Absence of a relevant difference in the parameter of interest.

## 4. Special considerations for characterization and quality assessment

WHO's *Guidelines on evaluation of similar biotherapeutic products (SBPs)* provide the principle that the demonstration of similarity of a candidate SBP product with respect to the RBP in terms of quality is a prerequisite for moving forward into comparative nonclinical and clinical studies *(1).* In particular, studies should be comparative in nature and should be performed with an appropriate number of batches of the reference product and of the SBP that is representative of the material intended for clinical use. The RBP should be extensively tested by analysing, if possible, multiple batches over an extended time period (to attempt to evaluate changes during the shelf-life). The minimum number of batches which should be tested will depend on the extent of the variability of the reference product and on assay variability. The number tested should be sufficient to draw meaningful conclusions on the variability of a given parameter for both the SBP and the RBP and on the similarity of both products. To obtain unambiguous results, the methods used should be scientifically valid and suitable for their purpose.

In comparison to many other proteins, mAbs are complex glycoproteins with distinct structural features which contribute to their diverse and variable biological functions. Specific carbohydrates can also have an impact on the biological activity of mAbs. For example, fucose bound by a α1-6 linkage to the core portion of N-linked carbohydrate chains or a non-reduced terminal galactose interferes with the ability of the antibody to bind well to certain Fc receptors, resulting in diminished Fc-mediated activities, including ADCC. Consequently, the assessment of biological activity of mAb biosimilars is particularly important and has some unique characteristics. The expression system used for production of mAbs can, in some cases, considerably affect the structure and function of the mAb product. The general principles for quality assessment of biosimilar mAbs, including physicochemical characterization, are already described in WHO rDNA and SBP guidelines *(1, 2).* Thus the quality part of this document will focus only on specific considerations for the assessment of mAb biological activities and on the impact of the expression system selected for production.

### 4.1 Strategy for assessment of mAb biological activity

Biological activity of mAb products is an important parameter and must be appropriately assessed. Since changes of higher-order structure could alter the biological activity of the mAb and may not be detected by physicochemical methods, the analysis of bioactivity is additionally useful for confirming the comparability of higher-order structure.

An understanding of the mechanism of action (MoA) and receptor interactions of the mAb is important for considering the strategy for biological activity assessment in both the characterization study and the comparability study. MAbs exert their action by various mechanisms ranging from simple binding to antigen, which alone mediates the clinical effect, to binding antigen and mediating one or more immunobiological mechanisms which combine to

produce the overall clinical response. These properties may play a role in the MoA and/or have an impact on product safety and efficacy. Therefore, a detailed analysis of the biological activity of the mAb, demonstrating the MoA (e.g. ligand neutralization, ADCC, CDC, apoptosis), ability for binding to Fc gamma and neonatal Fc receptors, as well as complement binding should be provided as described in Section 5.1.

Although simple antigen binding may seem to be the only mechanism operating to achieve clinical efficacy, other effects may also play a role. In some cases multiple functions of the mAb may be involved in an additive or synergistic manner to produce an overall combined clinical effect, and this may be hard to dissect experimentally to allow a clear understanding of how the mAb mediates its clinical efficacy. Therefore, if intact mAbs are used, care must be taken not to assume that the Fc mediated immunobiological effects of the product are not involved in clinical efficacy, even in situations where simple antigen binding is considered to be the primary MoA. For example, rituximab, which is a chimeric mAb specific for CD20, requires Fc function and especially ADCC for its clinical efficacy. Assessment of Fc function (including ADCC) is therefore paramount for this mAb. For infliximab, a tumor necrosis factor alfa (TNF-alfa) antagonist, neutralization of soluble TNF is the primary MoA while Fc function seems less important. However, as a potential secondary MoA, ADCC as well as other Fc- and Fab-related functions (e.g. reverse signalling) also need to be considered.

The assay for measuring Fc functions can be technically demanding. Both the difference of assay formats and of cell combinations have significant impact on the sensitivities of the assays. Assays for investigating ADCC activity require appropriately responsive target cells and efficient effector cells. Although the use of primary cells may provide a more physiologically relevant model, the criteria of low assay variability and robustness may not be satisfied. Continuously-growing cell lines may overcome these limitations in some cases provided they are more sensitive and more capable of detecting minor differences between the RBP and the SBP. However, identifying or producing a suitable cell line can be difficult and arduous. Furthermore, the clinical relevance of data generated by engineered/artificial cell lines may also be challenged because of over-expression of targets/receptors and the homogenous populations. Therefore, selection of an appropriate assay for the intended purpose should always be considered as a priority in developing the strategy for assessing mAb biological activity. Additional data may be generated by the use of different assay formats and cell combinations in order to obtain results that are more relevant to the physiological/pathophysiological conditions in patients. Although biological assays used in characterization or for demonstrating similarity may not be as robust as release assays, the assays should be qualified for the intended use and should be sufficiently sensitive to detect minor differences between the RBP and the SBP.

## 4.2 Considerations for selection of the expression system

The WHO SBP guideline allows for the use of different expression systems for production of the biosimilar compared to the reference product, so long as the manufacturer can convincingly

demonstrate that the structure of the molecule is not affected or that the clinical profile of the product will not change. However, this may pose a challenge in the context of biosimilar development for a mAb. Therefore, the expression system should be carefully selected, taking into account expression system differences that may result in undesired consequences, such as an atypical glycosylation pattern or a different impurity profile when compared to the RBP.

Differences in glycoforms present on products may or may not have clinical consequences. For instance, production cells based on mouse cell lines, such as SP2/0 and NS0, secrete mAbs with the carbohydrate structure gal-alpha-1, 3-gal present on the carbohydrate moiety. Humans cannot produce the gal-alpha-1, 3-gal structure as they lack the necessary enzyme for its synthesis; however, many humans produce antibodies against this. In a proportion of these individuals the antibodies are of the IgE class and this sensitization can result in anaphylactic reactions (often serious) if they are treated with mouse-cell-line-derived gal-alpha-1, 3-gal containing mAbs. Such pre-existing antibodies, which are particularly evident for cetuximab (an inhibitor of epidermal growth factor receptor, EGFR) may potentially be avoided by using cell substrates of human origin or some clones of Chinese hamster ovary (CHO) cells for mAb production since these cells cannot synthesize the gal-alpha-1, 3-gal antigen. This type of phenomenon can have important implications for biosimilar mAb development. For instance, producing an SBP of cetuximab in mouse cells would probably show the same gal-alpha-1, 3-gal-related anaphylaxis problems as the reference product. However, production of the mAb in CHO cells may avoid the anaphylaxis problem since the gal-alpha-1, 3-gal structure would not be present on the mAb, but the differences in glycosylation, and possibly other modifications, could have an impact on the extent of studies needed for demonstration of biosimilarity. Therefore the selection of an expression system for biosimilar mAb requires careful consideration because various potential issues need to be thoroughly assessed in order to ensure that the expression system difference does not result in a change of critical quality attributes.

## 4.3 International standards for biological assays used in the characterization

Assays for the biological activity of mAbs should be qualified for the intended purpose using WHO international standards or WHO reference reagents or secondary standards calibrated against WHO international standards or WHO reference reagents when available. Importantly, a clear distinction exists between reference products and WHO international standards or reference reagents since they serve different purposes and cannot be used interchangeably. The key difference between their uses reflects the fact that the reference product is the “comparator” for all the comparability studies, whereas WHO international standards and reference reagents are used for calibrating procedures, particularly bioassays, and cannot be used as comparators during SBP development. The distinct roles of reference products and international standards are described elsewhere *(1,3).*

# 5. Special considerations for nonclinical evaluation

As regards nonclinical development, as for all SBPs, a stepwise approach should be applied to evaluate the similarity of biosimilar and reference mAbs. In vitro studies should be conducted first and a decision then made regarding the extent to which, if necessary, in vivo work will be required. When deemed necessary, in vivo nonclinical studies should be performed before initiating clinical trials.

The following approach may be considered and should be tailored on a case-by-case basis to the similar biotherapeutic product concerned. The approach must be fully justified in the nonclinical overview.

## 5.1 In vitro studies

*SBP – general aspects*

In order to assess any difference in biological activity between the SBP and the RBP, data from a number of in vitro studies, some of which may already be available from quality-related assays, should be provided.

As for all SBPs, the following general principles apply to biosimilar mAbs:

- The studies should be sensitive, specific and sufficiently discriminatory to provide evidence that observed differences in quality attributes, as well as possible differences that may not have been detected during the comparative analytical assessment, are not clinically relevant. Functional studies should be comparative and should be designed to be sufficiently sensitive to detect differences in the concentration–activity relationship between the SBP and the RBP.
- Together, these assays should cover the whole spectrum of pharmacological/toxicological aspects with potential clinical relevance for the reference product and for the product class.
- The manufacturer should discuss to what degree the in vitro assays used are representative/ predictive of the clinical situation according to current scientific knowledge.

Since in vitro assays may often be more specific and sensitive to the detection of differences between the biosimilar and the reference product than studies in animals, these assays can be considered as paramount for the nonclinical biosimilar comparability exercise.

*Biosimilar mAbs – specific aspects*

For similar biotherapeutic mAbs, the nonclinical in vitro programme should usually include relevant assays for the following specific topics:

- Binding studies:
  - binding to target antigen(s);
  - binding to representative isoforms of the relevant Fc receptors (i.e. for IgG-based mAbs to FcγRI, FcγRII and FcγRIII), FcRn and complement (C1q).
- Functional studies/biological activities:

- Fab-associated functions (e.g. neutralization of a soluble ligand, receptor activation or blockade);
- Fc-associated functions (e.g. ADCC, CDC, complement activation), as applicable.

These assays are often technically demanding and the models chosen should be appropriately justified by the applicant (see section 4.1). Together these assays should broadly cover the functional aspects of the mAb even though some may not be considered essential for the therapeutic mode of action. However, an evaluation of ADCC and CDC is generally not needed for mAbs directed against non-membrane-bound targets if appropriately justified.

*Additional notes*

As indicated in the ICH S6 (R1) guideline *(4)*, tissue cross-reactivity studies with mAbs are not suitable for detecting subtle changes in critical quality attributes and are thus not recommended for assessing comparability.

### 5.2 In vivo studies

*Determination of the need for in vivo studies*

- As for SBPs in general, on the basis of the totality of quality and nonclinical in vitro data available and the extent to which there is residual uncertainty about the similarity of test mAb and the reference mAb, it is at the discretion of NRAs to waive or not to waive the request for nonclinical in vivo studies. If the quality comparability exercise and the nonclinical in vitro studies are considered satisfactory and no issues are identified which would block direct entrance into humans, an in vivo animal study may not be considered necessary.
- However, some mAbs may mediate effects that cannot be fully elucidated by in vitro studies. Therefore, in such cases nonclinical in vivo studies may be necessary to provide complementary information.

*General aspects to be considered for all SBPs, including biosimilar mAbs*

- If there is a need for additional in vivo information, the availability of a relevant animal species or other relevant models (e.g. transgenic animals, transplant models) should be considered.
- If a relevant in vivo animal model is not available, the manufacturer may choose to proceed to human studies taking into account principles to mitigate any potential risk.
- When the need for additional in vivo nonclinical studies is evaluated, factors to be considered include but are not restricted to:
  - the presence of potentially relevant quality attributes that have not been detected in the reference product (e.g. new post-translational modification structures);
  - the presence of potentially relevant quantitative differences in quality attributes between the SBP and the RBP;

  - relevant differences in formulation (e.g. use of excipients not widely used for mAbs). Although each of the factors mentioned here do not necessarily warrant in vivo testing, these factors should be considered together to assess the level of concern and whether there is a need for in vivo testing.
- If product-inherent factors that have an impact on pharmacokinetics (PK) and/or biodistribution, such as extensive glycosylation, cannot sufficiently be characterized on a quality and in vitro level, the manufacturer should carefully consider if in vivo animal PK and/or pharmacodynamics (PD) studies should be performed in advance of clinical PK/PD testing.

*Performance of in vivo studies*

The following explanations apply to all SBPs, including similar biotherapeutic mAbs.

*General aspects*

If an in vivo evaluation is deemed necessary, the focus of the study/studies (PK and/or PD, and/or safety) depends on the need for additional information to address residual uncertainty from the quality and in vitro nonclinical evaluation.

Animal studies should be designed to maximize the information obtained. The duration of the study (including observation period) should be justified, taking into consideration the PK behaviour of the reference mAb, the time to onset of formation of anti-drug antibodies (ADAs) in the test species and the clinical use of the reference mAb.

Effects of SBPs are often species-specific. In accordance with ICH S6(R1) and WHO's *Guidelines on the quality, safety and efficacy of biotherapeutic protein products prepared by recombinant DNA technology (2,4)*, in vivo studies should be performed only in relevant species – i.e. species which is/are pharmacologically and/or toxicologically responsive to the SBP.

*PK and/or PD studies*

When the model allows, the PK and/or PD of the SBP and the RBP should be compared quantitatively, including by, if feasible, a dose–response assessment that includes the intended exposure in humans.

In vivo assays may include the use of animal models of disease to evaluate functional effects on PD markers or efficacy measures.

*Safety studies*

In case in vivo safety studies are deemed necessary on the basis of the need for additional information, a flexible approach should be considered. The conduct of repeated dose toxicity studies in non-human primates is usually not recommended. If appropriately justified, a repeated dose toxicity study with refined design (e.g. using just one dose level of SBP and RBP and/or just

one gender and/or no recovery animals) and/or an in-life evaluation of safety parameters (such as clinical signs, body weight and vital functions) may be considered. Depending on the endpoints needed, it may not be necessary to sacrifice the animals at the end of the study.
For repeated dose toxicity studies, where only one dose is evaluated and the focus of the study is an evaluation of potential qualitative differences in the toxicity profile between RBP and SBP, the dose would usually be selected at the high end of the known dosing range of the RBP. In case the focus of the study is an evaluation of potential quantitative differences with regard to the known toxicity profile of the RBP, the dose level most likely to reveal differences between RBP and SBP should be chosen as justified on the basis of the known toxicity and/or pharmacodynamic response of the RBP.

The conduct of toxicity studies in non-relevant species (i.e. to assess unspecific toxicity only, based on impurities) is not recommended. Because of the different production processes used by the biosimilar and reference product manufacturers, qualitative differences of process-related impurities will occur (e.g. host cell proteins). Such impurity should be kept to a minimum in order to minimize any associated risk.

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

However, while immunogenicity assessment in animals is generally not predictive of immunogenicity in humans, it may be needed for PK/toxicokinetics (TK) interpretation of in vivo animal studies. Therefore, adequate blood samples should be taken and stored for future evaluations if needed.

*Local tolerance studies*
Studies on local tolerance are usually not required. If excipients are introduced for which there is little or no experience with the intended clinical route, local tolerance may need to be evaluated. If other in vivo studies are performed, the evaluation of local tolerance may be part of the design of those studies rather than performing separate local tolerance studies.

*Other studies*
Studies regarding safety pharmacology and reproduction toxicology are usually not required for nonclinical testing of biosimilar mAbs.

In accordance with ICH S6 (R1) and WHO's *Guidelines on the quality, safety and efficacy of biotherapeutic protein products prepared by recombinant DNA technology (2,4),* genotoxicity and (rodent) carcinogenicity studies are not required for (similar) biotherapeutic products. These criteria also apply to biosimilar mAbs.

# 6. Special considerations for clinical evaluation

The clinical comparability exercise is a stepwise procedure that should begin with PK/PD studies and usually continues with one pivotal clinical trial. In exceptional circumstances, data obtained in clinical PK/PD studies may suffice to confirm biosimilarity established in preceding steps. In general, the goal of the clinical evaluation programme is to confirm that any residual uncertainty about quality attributes would not result in clinically meaningful differences and not to establish the product's efficacy and safety in a particular indication.

If relevant differences between the SBP and the RBP are detected at any stage, the reasons for the differences must be explored and justified. In coming to a conclusion on whether a product qualifies as an SBP, the totality of evidence must be considered.

## 6.1 Pharmacokinetic studies

### 6.1.1 Aim of comparative pharmacokinetic studies

Comparative clinical PK studies are always required and should be used to further confirm the comparability of a mAb SBP to the RBP that has been established through comparative structural, functional and nonclinical studies. In general, factors to consider include whether the mAb is targeting a soluble antigen or a membrane-bound antigen, and whether it is dependent on FcRn binding and/or dependent on target-mediated clearance or non-target-mediated clearance. For instance, a mAb SBP may differ in its affinity for FcRn receptors from its RBP which may lead to either a shorter or longer half-life. As a consequence of a shorter half-life, drug exposure would be reduced which may lead to a lower efficacy *(5).* Comparative PK studies may be useful in monitoring the impact of the formation of ADAs on efficacy and safety, whereas exploration the impact of ADAs on PK in return is also necessary. Both approaches contribute to establishing evidence in support of extrapolation. It is not necessary to study the PK of the mAb SBP in every indication that is being sought. In general, one comparative PK study should be sufficient to bridge across the indications for which the reference mAb has been authorized. The design of the comparative PK studies depends on various factors, including clinical context, safety profiles and PK characteristics of the reference product (target-mediated disposition, linear or non-linear PK, time-dependency, half-life, etc.).

### 6.1.2 Study design and population

A single-dose PK study in healthy volunteers is generally recommended as they can be considered as a sensitive and homogenous study population *(6).* A parallel group design, which generally requires a higher number of subjects, is usually required for mAbs since a single-dose

crossover design may not be appropriate because of the long half-lives of mAbs and the potential influence of immunogenicity on the PK profile. However, for mAb fragments or mAbs which are not administered systemically, alternative approaches may also be applied.

Regarding the use of healthy volunteers to study the PK of mAbs, some key considerations should be taken into account. First, healthy subjects are generally preferred, if possible, because of their higher sensitivity and homogeneity as compared to patient populations. Second, a clinically relevant dose of some mAbs (e.g. bevacizumab) may not be considered ethical in healthy volunteers because of safety concerns and, in these cases, a subtherapeutic dose which is on the linear part of the dose–response curve may be required. Third, it may be necessary to perform the PK study for mAb SBPs in a sensitive patient population rather than in healthy volunteers for safety reasons, as in the case of rituximab. Unnecessary exposure to risk (because of safety or medical reasons) would be viewed as unethical. Fourth, it may sometimes be necessary to perform the PK study in a different population to that selected in order to establish similar clinical efficacy. In such scenarios, population PK measurements should be collected during the clinical efficacy trial since such data may add relevant information on similarity. Measurement of PK parameters, especially trough levels, along with sampling for immunogenicity, is also recommended for the evaluation of clinical correlates of possible anti-drug antibodies. Further, the choice of a particular population for PK analysis also depends on the range of therapeutic indications of the mAb under development. For instance, if a reference mAb is authorized both as an anti-inflammatory agent and as an anti-cancer antibody (e.g. as with rituximab), PK data in one therapeutic area may complement clinical data obtained in another therapeutic area and thus can also strengthen the evidence for indication extrapolation.

### 6.1.3 Regimen

mAbs are often indicated both for monotherapy and as a part of combination regimens that incorporate immune-suppressants or chemotherapeutics. It may be sensible to study the comparative PK in the monotherapy setting in order to minimize the sources for variability. When concomitant therapy alters PK, it may be appropriate to study comparative PK both in the monotherapy setting and in combination, particularly if one cannot exclude differences with regard to quality attributes that might specifically have an impact on how the drug was cleared when used in combination.

### 6.1.4 Pharmacokinetic characteristics of the reference mAb

The PK of the mAb may be affected by factors such as the antigen/receptor level (e.g. related to tumour burden in oncology), the existence of target-mediated clearance, and/or receptor shedding which has an impact on variability of PK measurements. These factors should be considered when selecting the population in which to compare the PK of the SBP to the reference.

### 6.1.5 Doses

A dose should be selected that will enable the detection of potential PK differences between the mAb SBP and the reference mAb. mAbs generally possess a high degree of target selectivity, with many exhibiting nonlinear distribution and elimination, influenced by binding to their target. In general, it is recommended that the PK profiles should be compared using the lowest recommended therapeutic dose. A higher, or the highest, therapeutic dose may be required where the non-specific clearance mechanism dominates. For mAbs that are eliminated by target-mediated disposition, a low dose (i.e. one at which target-mediated disposition, or TMD, is not saturated) may be particularly useful for detecting differences in PK *(7).*

### 6.1.6 Routes of administration

Administration via a route that requires an absorption step is preferred unless intravenous administration only is intended. Where the route of administration requires an absorption step, such as the subcutaneous route, standard comparisons of $C_{max}$, $AUC_t$ and $AUC_{0\text{-}inf}$ may be used to assess PK comparability.

### 6.1.7 Sampling times and parameters

Primary PK comparability studies should include sufficient sampling to accurately characterize $C_{max}$ and should adequately characterize the late elimination phase so that the terminal disposition rate constant can be reliably estimated, including at time points, to sufficiently characterize any ADA response. In single-dose studies, optimal sampling should continue past the expected last quantifiable concentration ($AUC_t$), and the concentration-time curve should cover at least 80% of $AUC_{0\text{-}inf}$.

If a multiple-dose study is performed in patients, sampling should be carried out at first dose and at steady state. Steady state is typically reached after five half-lives of the mAb. PK parameters that should be evaluated include $AUC_{tau}$, $C_{max}$ and $C_{trough}$, clearance and half-life. For mAbs that are administered only intravenously, the aforementioned parameters should be compared, as should parameters that reflect the clearance of the product.

### 6.1.8 Specific assays for serum drug concentration

It is preferable to have a single, validated bio-analytical assay to detect both the mAb SBP and the reference mAb. The bio-analytical assay should be appropriate for the detection and quantification of mAbs and should be demonstrated to be bio-analytically comparable with respect to its ability to quantify precisely and accurately both the mAb SBP and the reference *(8).* The production of ADAs may interfere with assays for test products. Therefore ADAs should be measured in parallel with PK assessment, using the most appropriate sampling time points and a subgroup analysis by ADA status should be performed. The PK analysis on the ADA negative samples is of particular interest, providing the clearest picture of PK similarity.

### 6.1.9 Equivalence margin

In general, a comparability margin of 80–125% for the primary parameters may be acceptable but must be justified. In some circumstances, narrowing or widening may be required and should be justified.

## 6.2 Pharmacodynamic (PD) studies

In general, it is advisable to include PD markers as part of the clinical comparability exercise (PK and pivotal efficacy trial).

### 6.2.1 PD markers and PD assay

Characteristics of PD markers that would support clinical efficacy, and that manufacturers should pay attention to, are given below *(6):*

- The PD marker should be sufficiently sensitive to detect relevant differences, and should be measurable with sufficient precision.
- The use of multiple PD markers, if they exist, is recommended.
- The study dose–concentration–response relationships or time–response relationships of the selected doses should be within the linear part of the established dose–response curve of the RBP.
- A clear dose–response relationship is shown.
- The PD marker is an acceptable surrogate marker and is related to patient outcome.
- An equivalence margin should be predefined and justified.
- The PD assay should at least be relevant to a pharmacological effect of the biological product (PD assay is highly dependent on the pharmacological activity of the product; the approach for assay validation and the characteristics of the assay performance may differ depending on the specific PD assay).

For some mAbs it may even be possible to perform confirmatory PD studies instead of comparative clinical efficacy studies with conventional clinical outcome measures. When clinical studies using PD markers are planned to provide the pivotal evidence to establish similarity, it is recommended to discuss such an approach with the regulatory authorities.

## 6.3 Comparative clinical efficacy study

The confirmatory efficacy trial is the last step of the comparability exercise, thus confirming that the clinical performance of the SBP and the RBP are comparable. Typically, one clinical efficacy study which is randomized, adequately powered and preferably double-blinded should be performed.

The manufacturer of a biosimilar mAb should perform a thorough analysis of the available clinical data in the public domain for the reference product in order to determine the population

endpoint combination that is likely to provide a relevant and sensitive model for detecting clinically meaningful differences in efficacy and safety and for extrapolating efficacy and safety to therapeutic indications that are not investigated. The type of comparative clinical trial required for the proposed mAb SBP could be influenced by several factors, including:

- the nature and complexity of the mAbs and derived products;
- the extent of clinical experience with the reference mAb in terms of efficacy, safety and immunogenicity;
- the degree of understanding of the MoA of the mAb and disease pathology, and the extent to which they vary in different indications – including MoA, site of action, antigen load, drug administration (dose, route, regimen, duration), concomitant medications, and target population sensitivity to drug effects.

The clinical data obtained in a sensitive model can also be used to support extrapolation to other indications in which the proposed mAb SBP has not been tested.

### 6.3.1 Clinical trial design

Clinical trial design and statistical analysis of equivalence and non-inferiority trials that are already addressed in WHO's *Guidelines on evaluation of similar biotherapeutic products (SBPs)* clearly apply to biosimilar mAbs *(1).* The guidelines stress the importance of clearly stating the specific design selected for a given study, and include details on the determination of the comparability margin, sample size determination and statistical analyses. For biosimilar mAbs, extrapolation to other indications is especially important, and additional considerations are required in order to design a meaningful trial to support additional indications.

Although equivalence or non-inferiority studies would be acceptable for the comparative clinical studies of the biosimilar mAb to the RBP, equivalence trials are generally preferred. Detailed explanations of the advantages and disadvantages of equivalence/non-inferiority designs for similar biotherapeutic products are provided in the WHO guidelines and are also available from several other bodies on SBPs *(1,9,10,11).* Special considerations for clinical trial design of mAbs developed as biosimilars are provided below.

A demonstration of equivalence, as opposed to non-inferiority, is especially important given that extrapolation to other indications may be one of the goals of the development programme for the biosimilar mAb. Non-inferiority trials are one-sided and hence do not exclude the possibility that the biosimilar mAb could be found to be superior to the RBP. Such a finding would create challenges in providing a justification for extrapolation to other indications of the RBP. From a statistical perspective, assay sensitivity is important in order to provide some confidence that the trial, as planned and designed, will have the ability to detect differences between the biosimilar mAb and the RBP if such differences exist *(12).* A trial that lacks sensitivity could lead to the erroneous conclusion of equivalence of the biosimilar mAb to the RBP. The selected study population should not only be representative of the approved therapeutic indications of the RBP,

but should also be sufficiently sensitive to detect potential differences between the biosimilar mAb and the RBP. Hence, historical scientific evidence should be provided which shows that appropriately designed and conducted trials with the RBP against placebo for the approved indication have reliably demonstrated the superiority of the RBP over placebo.

Study population or study endpoints may deviate from those leading to approval of the RBP for the specific indication as long as the primary endpoints are sensitive to detection of clinically meaningful differences between the biosimilar mAb and the RBP. Whatever approach is taken, applicants should always justify their selection of endpoints, time points for analysis, and the proposed margin, irrespective of whether this follows the RBP approach or not, and may wish to consult relevant regulatory authorities during the planning and design stage of the trial if in doubt.

The efficacy of the RBP compared to placebo has been demonstrated previously. Therefore, it is considered clinically important to ensure that the biosimilar mAb retains a substantial fraction of the effect of the RBP. As a consequence, an equivalence margin that preserves a fraction of the smallest effect size that the RBP can be expected to have relative to a placebo control is the most suitable. The fraction of the effect size of the RBP that must be retained by the biosimilar mAb should be clearly justified in each case, and should take into account the smallest clinically important difference in a given setting. Once the margin has been selected, the determination of the required sample size should be based on methods specifically designed for equivalence/non-inferiority trials.

Statistical analysis of data from equivalence trials is typically based on the indirect confidence interval comparison which requires specification of the equivalence limits *(13).* Equivalence is demonstrated when the confidence interval for the selected metric of the treatment effect falls entirely within the lower and upper equivalence limits. If a p-value approach is used, then the p-values should be computed on the basis of the Two One-Sided Test (TOST) procedure, testing simultaneously the null hypotheses of inferiority and superiority. In using the TOST procedure, equivalence is demonstrated when the p-values obtained are less than the significance level used.

### 6.3.2 Study population

In order to detect differences between the mAb SBP and the reference mAb, clinical trials for the mAb SBP should be carried out in an appropriately sensitive patient population using endpoints that are sensitive to detection of clinically meaningful differences between the biosimilar and the reference product in the indication (see section 6.3.3). The rationale for the selected study population should be provided. In general, a homogeneous population of patients (e.g. same line(s) of therapy, severity or stage of disease progression) would minimize inter-patient variability and thus increase the likelihood of detecting differences between the mAb SBP and the reference mAb if such differences exist. Patients who have not received previous treatment

(e.g. first-line therapy) are considered to be more homogeneous than patients who have previously received several or different lines of therapy. Ideally, the observed clinical effects should be triggered by the direct action of the mAb SBP/reference mAb without interference of other medications because concomitant medications may affect or mask differences in PK/PD, efficacy, safety and/or immunogenicity of the tested products. To validate the effect of the reference mAb and the sensitivity of the study in the chosen study population, historical data should be used to justify the selection of the study population and the equivalence margin. This could generally be done through a meta-analysis or a systematic review.

mAbs can function through various mechanisms of action, such as agonist activity or receptor blockade (e.g. vascular endothelial growth factor, or VEGF, and EGFR), induction of apoptosis, delivery of a drug or cytotoxic agent, and immune-mediated mechanisms (e.g. CDC, ADCC and regulation of T cell function). Because the mechanisms involved in one disease may differ from those involved in another, extensive consideration should be given to the setting in which clinical comparability is to be tested, especially where it is known that extrapolation to other indications and uses will be sought.

Clinical studies in an unauthorized population (e.g. line of therapy, combination therapy, disease severity, indication authorized in some but not all jurisdictions) may be acceptable to demonstrate "no clinically meaningful differences" for biosimilar mAbs. Manufacturers of biosimilar mAbs should consult relevant regulatory authorities prior to conducting such studies.

### 6.3.3 Primary study endpoint

Clinically relevant and sensitive study endpoints within a sensitive population should be selected to improve the detection of potential differences between the mAb SBP and the reference product. In general, clinical outcomes, surrogate outcomes or a combination of both can be used as primary endpoints in mAb SBP trials. The same study endpoints for the innovator products may be used because a large body of historical data is generally available in the public domain for setting the equivalence margin and calculating the sample size. The study endpoints may be different from those traditionally used or from the endpoints recommended by study guidelines for the innovator products, as these endpoints may not be considered as the most sensitive endpoints for detecting clinically meaningful differences in an equivalent trial setting. A surrogate endpoint can be used as the primary endpoint when the surrogacy to the clinical outcome is well-established or generally accepted, as seen with pathological complete response (pCR) in neoadjuvant treatment of breast cancer. The choice of the study endpoint should always be scientifically justified. More sensitive clinical endpoints could be used as secondary endpoints for the innovator product, primary or secondary endpoints for the innovator products at different time points of analysis, and/or new surrogates. For example, overall response rate (ORR) or complete response (CR) rate as endpoints for clinical efficacy studies of mAb SBPs in oncology trials are recommended because these endpoints may be more sensitive and are not time-related. However, if progression-free survival (which is one of the endpoints that is frequently used for

clinical efficacy-testing for innovator products) is considered more sensitive than ORR, then this may be the preferred option. Likewise, in rheumatoid arthritis (RA), both continuous outcomes – e.g. changes in disease activity score of 28 joints (DAS28) from baseline – and dichotomous outcomes – e.g. American College of Rheumatology score 20 (ACR20) – are considered acceptable in RA trials for determining clinical comparability *(14).*

When the primary efficacy endpoints that were used for the RBP cannot be used for SBP, it is advisable to include some common endpoints as secondary endpoints to facilitate comparisons between the mAb SBP and the mAb RBP. The role of these secondary endpoints in the overall interpretation of the study results should be clearly defined, particularly in terms of whether the secondary endpoints are used to support or to confirm equivalency or similarity.

National regulatory authorities (NRAs) may not always agree on the choice of study endpoints. For a biosimilar manufacturer with a global development programme that is guided or required by various NRAs to fulfil local regulatory or clinical practice requirements, it may be possible to pre-specify different primary study endpoints with the statistical power in the same trial to comply with various regulatory requirements.

### 6.3.4 Safety

#### 6.3.4.1 General considerations

Comparative safety data should normally be collected pre-authorization. The extent of collection depends on the type and severity of safety issues known for the reference product. The SBP study population should be followed to provide information on safety events of interest according to the experiences with the reference mAb. Care should be given to compare the nature, severity and frequency of the adverse events between the mAb SBP and the reference product in clinical trials that enrolled a sufficient number of patients treated for an acceptable period of time. Clinical safety issues should be captured throughout clinical development during initial PK and/or PD evaluations and also in the primary clinical study establishing comparability.

#### 6.3.4.2 Immunogenicity

Therapeutic mAbs, like other rDNA-derived biotherapeutics, are recognized by the human immune system. This immune recognition may lead to an unwanted immune response. Monoclonal antibodies are often immunogenic in patients. Therefore, the goal of the development of a biosimilar mAb is to demonstrate comparable immunogenicity to the reference product. There are some special considerations for the immunogenicity of mAbs as compared to other biotherapeutics. For instance, mAbs do not evoke cross-reacting antibodies against the body's endogenous proteins as some growth factors and proteins for replacement therapy do. However, the development of assays for testing of anti-mAb-antibodies can be challenging.

From the regulatory point of view, animal data are not sufficiently predictive of the human immune response against a therapeutic protein. Thus, immunogenicity generally needs to be investigated as part of the clinical trial programme of a biosimilar mAb. The analysis of the immunogenicity of DNA-derived biotherapeutics is outlined in WHO's *Guidelines on the quality, safety, and efficacy of biotherapeutic protein products prepared by recombinant DNA technology* and *Guidelines on evaluation of similar biotherapeutic products (SBPs) (1,2).* These general guidances should be taken into account in assessing biosimilar mAbs. In addition, details regarding the advantages and disadvantages of particular assays, as well as some considerations for interpretation of the results and the decision-making process, are provided in several review articles *(15,16,17,18).*

*Immunogenicity of a biosimilar mAb*

The basic data package contains the incidence, titre, neutralization ability and persistence of product antibodies determined by appropriate assays, as well as their pharmacokinetic and clinical correlations. The immunogenicity programme needs to be tailored to each product. Thus, the evaluation of immunogenicity requires a multidisciplinary approach, including considerations of product-, process-, patient- and disease-related factors that will be the basis of a risk-based immunogenicity programme. It is recommended that the application for marketing authorization of a mAb will include a summary of the immunogenicity programme in support of the selected approach to immunogenicity. This summary should address the following topics as appropriate:

*Risk assessment*

- previous knowledge of the immunogenicity of the reference product, such as the presence of immunogenic structures in the active substance as well as the incidence, type, persistence and clinical correlations of the antibodies;
- findings of the physicochemical and structural comparisons between the biosimilar mAb and its reference product, including process-related impurities and aggregates;
- differences in formulation and packaging (e.g. potential impurities and leachables);
- route and/or mode of administration of the product;
- patient- and disease-related factors such as the state of the immune system, concomitant immune-modulatory therapy and potential pre-existing immunity, antigenicity and sensitivity.

*Risk-based immunogenicity programme*

The manufacturer should present a risk-based immunogenicity assessment programme:

- The basis of the immunogenicity assessment is the testing of patient samples pre-treatment, during treatment and, if needed, post-treatment in an appropriate set of assays that are suitable for the product in question. The measurement of antibodies to mAbs is methodologically challenging since standard assay formats involving anti-immunoglobulin reagents are inappropriate for this product class; therefore alternative methods should be used. A multi-tiered approach for assessment is needed, as for other biotherapeutic products. The developer has to validate assays for screening, confirmation and neutralization ability. Special attention

should be paid to the choice of the control matrix, determination of cut-off points, and the estimation of interference by matrix components, including the target and the residual drug in the sample. To mitigate the potential interference, corrective measures should be implemented. For example, drug interference may be overcome by allowing time for clearance of the drug from the circulation prior to sampling, or by dissociating immune complexes, and/or by removal of the drug. Inclusion of any of these measures must not compromise the detection of antibodies or patient treatment.

- With regard to the integration of the product antibody testing into the comparative clinical trials, it is particularly important to synchronize the sampling schedule and duration of the follow-up for product antibody determination and pharmacokinetic measurements, as well for assessments of safety and efficacy.
- Special emphasis should be placed on the potential association of product antibodies with loss of efficacy, with infusion reactions, and with acute and delayed hypersensitivity. The manufacturer should systematically use terminology and definitions to characterize potentially immune-mediated symptoms, in accordance with relevant publications *(19,20).*
- The manufacturer should take into account the dose and dosing schedule, including re-administration, after discontinuation of treatment.
- The vulnerability of the patient population(s) and the expected risks of immunogenicity should be taken into account in planning for the intensity of monitoring.
- The manufacturer should give a description and analysis of the use of pre-medication or de-immunization measures to mitigate acute infusion/injection-related reactions and other possibly immune-mediated reactions.

After discontinuation of the therapy, it is important to investigate the persistence of product antibodies formed during drug administration, as well as the emergence of product antibodies that may have escaped detection because of the immunosuppressive action of the product or because of technical problems (notably drug interference). The timing of the post-treatment samples should be justified.

*Comparative immunogenicity*

The lack of standardization and rapid evolution of the assay methodology makes it difficult to compare immunogenicity studies. Therefore, pre-licensing comparative immunogenicity data are generally needed in the development of SPBs *(1,11).* Immunogenicity testing of the SBP and the reference product should be conducted within the biosimilar comparability exercise by using the same assay format and sampling schedule. A parallel group design is recommended because of the long half-life of antibodies and because it may be difficult to interpret immunogenicity after a switch.

*Assays and mAb characterization*

ADA assays should ideally be capable of detecting all antibodies against both the biosimilar and the reference molecule. Thus, assays can be performed with both the reference and biosimilar molecule as the antigen/capture agent in parallel in order to measure the immune response against the product that was received by each patient. The challenge is to develop two assays with similar sensitivity. Cross-testing all serum samples by both tests is useful to explore assay performance and antigen epitopes. The use of a single assay with the active substance of the biosimilar as the antigen/capturing agent for evaluation of all samples (including those from reference-product-treated patients) will be able, in principle, to detect all antibodies developed against the biosimilar molecule (i.e. it is, in principle, conservative). However, this assumption needs to be justified by the manufacturer.

Following identification of confirmed antibody-positive samples, characterization of the antibodies is required. Determination of their neutralizing potential is essential and deviation from this requires justification. Although a functional (usually cell-based) bioassay or a binding (e.g. a competitive ligand-binding, or CLB) assay can be used alone, the latter should be used only if relevant to the MoA of the product. For example, a CLB assay is appropriate in a scenario where a therapeutic mAb acts by binding to a soluble ligand, thereby blocking it from interacting with its receptor and thus inhibiting the biological action of the ligand. Since the assay procedure measures binding to the target and inhibition of the binding activity if neutralizing antibodies are present, it is reflective of the MoA of the therapeutic mAb. For intact mAbs where effector functions are likely to contribute to the clinical effect, functional cell-based bioassays are recommended because the mechanism of action cannot be reflected adequately in a CLB assay. Nevertheless, such cell-based assays may not be sufficiently sensitive and a CLB assay may give a more accurate assessment of the incidence of neutralizing antibody induction.

Additional studies beyond the standard data package, such as immunoglobulin class, epitope mapping and IgG subclass, may be useful in specific situations (e.g. occurrence of anaphylaxis or use of certain assay formats). It may also be necessary to locate the antigenic sites (e.g. antigen-binding region versus constant region of antibody molecule). Banking of patient samples is necessary in order to have the possibility for re-testing in case of technical problems in the original assay.

*Clinical immunogenicity assessment*

The selected patient population should be sensitive for the detection of differences in immunogenicity. It is also important that the pivotal safety and efficacy study will include both immunogenicity and PK measurements (especially $C_{trough}$ levels) in order to establish the clinical impact of immunogenicity. If the study includes patients previously treated with the reference mAb, a subgroup analysis of previously-treated patients should be performed. The sampling schedule should be optimized for the demonstration of similar onset and persistency of antibodies to the test and the reference product.

The duration of follow-up of immunogenicity depends on the duration of exposure and should be sufficient to demonstrate similar persistence and clinical impact of the antibodies. In chronic administration, the minimum follow-up is 6 months.

Immunogenicity should be followed after licensing by monitoring possible immune-mediated adverse effects. Special immunogenicity studies may be necessary in high-risk situations (e.g. when the reference product is known to have serious but rare immune-mediated effects, such as anaphylaxis).

Evaluation of immunogenicity includes the incidence, titre, neutralization capacity and persistency, as well as correlations to exposure, safety and efficacy. For the time-being, there is no generally accepted statistical methodology that could be used to define the limits of comparable immunogenicity. In general, differences in immunogenicity of an SBP are not compatible with the biosimilarity principle unless the sponsor can show that the product antibodies have no clinical relevance and that the underlying difference between the SBP and the reference product does not signal an otherwise important problem.

### 6.4 Indication extrapolation

Indication extrapolation is the leveraging of efficacy and safety data from clinical studies in one therapeutic indication to support the authorization of other indications in which the biosimilar has not been studied, but for which the reference product is authorized and well-characterized. The starting point for extrapolation is that the physicochemical and structural analyses, nonclinical tests and clinical studies have demonstrated comparability. Thus, extrapolation should be considered in the light of the totality of evidence of comparability. As part of guidance for SBP development, the WHO guidance document has provided recommendations regarding extrapolation of clinical data across indications. The principles recommended for indication extrapolation for SBPs in the WHO guidance apply to biosimilar mAbs. The extrapolation is possible when the following requirements are fulfilled:

- A sensitive clinical test model has been used that is able to detect potential differences between the SBP and the RBP.
- The clinically relevant MoA and/or involved receptor(s) are the same.
- Safety and immunogenicity of the SBP have been sufficiently characterized and there are no unique/additional safety issues expected for the extrapolated indication(s).

mAbs have both Fab and Fc effector functions and may exert their clinical effect through a variety of mechanisms – e.g. ligand blockade, receptor blockade, receptor down-regulation, cell depletion (via ADCC, CDC, apoptosis), and signalling induction. A particular mAb may act through one or a combination of these or other mechanisms. Where a therapeutic mAb is indicated for a variety of diseases, various MoAs may be important, depending on the indication

in question. Therefore it is essential that the basic functions of the antibody are considered when relevant. The tests should be selected according to their relevance for a particular product and therapeutic indication and, if possible, tailored (e.g. ADCC) accordingly. If differences are found, additional data are needed in order to address any residual uncertainty. Such data may include quality or PK/PD data; confirmatory safety and efficacy studies are rarely needed. Special post-marketing measures may be used to monitor aspects of safety and/or immunogenicity in the extrapolated therapeutic indications.

### 6.5 Pharmacovigilance and post-approval consideration

A risk management plan (RMP) should be in place once a biosimilar mAb is approved, in order to ensure its long-term safety and efficacy. The general requirements for pharmacovigilance are the same as for any approved new drug. As described in WHO documents, it is essential to record the product brand name, batch number, manufacturers' name and INN (where it exists) *(1,2).* In many cases, clinically important adverse events occur at a relatively low frequency and the probability of them occurring during the time frame of the clinical trial is also low. Additionally, because of the relatively small sample size, mAb SBP clinical trials may have the statistical power only to detect common adverse events. Thus, as for any biological medicines, pharmacovigilance is essential for the identification and assessment of potential post-marketing risks in order to enable detection of potential overt new or rare mAb SBP specific safety findings.

## Authors and acknowledgements

The first draft of these guidelines was prepared by the following authors: Dr K. Gao, World Health Organization, Geneva, Switzerland; Dr E. Griffiths, Consultant, Kingston upon Thames, United Kingdom; Dr H.K. Heim, Federal Institute for Drugs and Medical Devices, Bonn, Germany; Dr I. Knezevic, World Health Organization, Geneva, Switzerland; Dr P. Kurki, Finnish Medicines Agency, Helsinki, Finland; Dr N. Ekman, Finnish Medicines Agency, Helsinki, Finland; Dr C. Njue, Health Canada, Ottawa, Canada; Dr R. Thorpe, Consultant, United Kingdom; Dr M. Wadhwa, National Institute for Biological Standards and Control, Potters Bar, United Kingdom; Dr J. Wang, Health Canada, Ottawa, Canada; and Dr E. Wolff-Holz, Paul Ehrlich Institute, Langen, Germany. The drafting group took into consideration proposals from the WHO informal consultation on monoclonal antibodies as similar biotherapeutic products (SBPs) held in Geneva in April 2015 with following participants:

Temporary advisers: Dr M.C. Bielsky, Medicines and Healthcare Products Regulatory Agency, London, United Kingdom; Dr E. Griffiths, Consultant, Kingston upon Thames, United Kingdom; Dr H.K. Heim, Federal Institute for Drugs and Medical Devices, Bonn, Germany; Dr J. Joung, Ministry of Food and Drug Safety, Chungcheongbuk-do, Republic of Korea; Dr C. Njue, Health Canada, Ottawa, Canada; Dr R. Thorpe, Consultant, United Kingdom; Dr M. Wadhwa, National Institute for Biological Standards and Control, Potters Bar, United Kingdom; Dr J. Wang, Health Canada, Ottawa, Canada; Dr J. Wang, National Institutes for Food and Drug Control, Beijing, China; Dr T. Yamaguchi, Pharmaceuticals and Medical Devices Agency, Tokyo, Japan.

Representatives of national regulatory authorities: Mrs A. Abas, Ministry of Health Malaysia, Putrajaya, Malaysia; Dr A. Abdelaziz, Jordan Food and Drug Administration, Amman, Jordan; Dr K. Bangarurajan, Central Drug Standards Control Organization, Delhi, India; Dr B. Boonyapiwat, Ministry of Public Health, Bangkok, Thailand; Ms J. Dahlan, National Agency of Drug and Food Control, Jakarta, Indonesia; Mrs D. Darko, Food and Drug Authority, Accra, Ghana; Mr D. Goryachev, Ministry of Health, Moscow, Russian Federation; Dr C. Ilonze, National Agency For Food and Drug Administration and Control, Abuja, Nigeria; Mrs W. Jariyapan, Ministry of Public Health, Bangkok, Thailand; Dr D. Khokal, Health Sciences Authority, Singapore; Dr P. Kurki (on behalf of European Medicines Agency Biosimilar Medicinal Products Working Party), Finnish Medicines Agency, Helsinki, Finland; Ms I. Lyadova, Ministry of Health, Moscow, Russian Federation; Mrs S. Marlina; National Agency of Drug and Food Control, Jakarta, Indonesia; Dr J. Meriakol; Ministry of Health, Nairobi, Kenya; Mr K. Nam, Ministry of Food and Drug Safety, Chungcheongbuk-do, Republic of Korea; Mrs Y.H. Nunez, Centro para el Control Estatal de la Calidad de los Medicamentos, Havana, Cuba; Dr C. Gonzalez, Ministry of Health, Bogota, Colombia; Mrs B. Valente, National Health Surveillance Agency, Brasilia, Brazil; Ms N. Vergel, National Institute of Drugs and Food, Bogota, Colombia; Dr L. Wang, National Institutes for Food and Drug Control, Beijing, China; Dr E. Wolff-Holz, Paul Ehrlich Institute, Langen, Germany; Dr S. Xie, China Food and Drug Administration, Beijing, China.
Representatives of the Latin American Association of Pharmaceutical Industries (ALIFAR): Dr E. Spitzer, Argentina.

Representatives of the Developing Countries Vaccine Manufacturers' Network (DVCMN): Dr A. Qu, Shanghai Institute of Biological Products Company Limited, China; Ms L. Susanti, Bio Farma, Indonesia; Dr A. Zhang, China National Biotec Group Company Ltd., China.

Representatives of the International Generic Pharmaceutical Alliance (IGPA)/European Generic Medicines Association (EGA): Dr S. Bukofzer, Hospira (IL), USA; Mrs S. Kox, European Generic Medicines Association, Brussels, Belgium; Dr M. Schiestl, Sandoz GmbH, Kundl, Austria; Dr C. Webster, Baxter, Cambridge (MA), USA.

Representatives of the International Federation of Pharmaceutical Manufacturers and Associations (IFPMA): Mrs J. Bernat, International Federation of Pharmaceutical Manufacturers & Associations, Geneva, Switzerland; Dr K. Heidenreich, Novartis, Basel, Switzerland; Dr J. Macdonald, Pfizer, Tadworth, United Kingdom; Dr S. Ramanan, Amgen (CA), USA; Dr T. Schreitmueller, F. Hoffmann–La Roche Ltd., Basel, Switzerland; Dr J. Siegel, Johnson & Johnson, Washington (DC), USA; Dr K. Watson, AbbVie Ltd, Maidenhead, United Kingdom.

Representatives of individual manufacturers: Mr R.Ivanov, Biocad, St. Petersburg, Russian Federation; Dr A. Kudrin, Celltrion, Incheon, Republic of Korea; Mr B. Kim, Celltrion, Incheon, Republic of Korea.

Representatives of WHO Regional Offices: Ms H.B. Pedersen, World Health Organization Regional Office for Europe, Copenhagen, Denmark; Dr M.L. Pombo, Pan American Health Organization, Washington (DC), USA.

WHO Secretariat: Dr K. Gao, Dr H. Kang, Dr I. Knezevic and Dr David Wood, World Health Organization, Geneva, Switzerland.

## References

1. Guidelines on evaluation of similar biotherapeutic products (SBPs). In: WHO Expert Committee on Biological Standardization: sixtieth report. Geneva: World Health Organization; 2013: Annex 2 (WHO Technical Report Series, No. 977; http://www.who.int/entity/biologicals/expert_committee/TRS_977_60th_report.pdf?ua=1, accessed 6 July 2016).

2. Guidelines on the quality, safety and efficacy of biotherapeutic protein products prepared by recombinant DNA technology. In: WHO Expert Committee on Biological Standardization: sixty-fourth report. Geneva: World Health Organization; 2014: Annex 4 (WHO Technical Report Series, No. 987; http://apps.who.int/iris/bitstream/10665/129494/1/TRS_987_eng.pdf?ua=1&ua=1, accessed 6 July 2016).

3. Thorpe R, Wadhwa M. Intended use of reference products and WHO International Standards/Reference Reagents in the development of similar biological products (biosimilars). Biologicals. 2011;39(5):262–5.

4. Preclinical safety evaluation of biotechnology-derived pharmaceuticals. ICH Guideline S6 (R1). Geneva: International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use; 2011.

5. Wang W, Lu P, Fang Y, Hamuro L, Pittman T, Carr B et al. Monoclonal antibodies with identical fc sequences can bind to FcRn differentially with pharmacokinetic consequences. Drug Metab Dispos. 2011;39(9):1469–77.

6. Committee for Medicinal Products for Human Use. Guideline on similar biological medicinal products containing monoclonal antibodies – nonclinical and clinical issues. London: European Medicines Agency; 2012 (Document EMA/CHMP/BMWP/403543/2010).

7. Pippig SD, et al. Biosimilar monoclonal antibodies. In: Dübel S, Reichert JM, editors. Handbook of therapeutic antibodies. Weinheim: Wiley-VCH Verlag; 2014. doi: 10.1002/9783527682423.ch24. 2014:681-704.

8. Marini JC, et al. Systemic verification of bioanalytical similarity between a biosimilar and a reference biotherapuetic. Committee recommendations for the development of a single ligand binding assay to support PK assessments. The AAPS Journal. 2014;16(6):1149–58.

9. Guidance for sponsors: information and submission requirements for subsequent entry biologics (SEBs). Ottawa: Health Canada; 2010.

10. Guidance for industry: scientific considerations in demonstrating biosimilarity to a reference product. Silver Spring (MD): Food and Drug Administration; 2015.

11. Committee for Medicinal Products for Human Use. Guideline on similar biological medicinal products containing biotechnology-derived proteins as active substance: nonclinical and clinical issues. London: European Medicines Agency; 2015.

12. Choice of control group and related issues in clinical trials. ICH Guideline E10. Geneva: International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use; 2000.

13. Shein-Chung C, Jen Pei L. Design and analysis of clinical trials: concepts and methodologies. Hoboken (NJ); John Wiley & Sons; 2004.

14. Kudrin A, et al. Case studies on clinical evaluation of biosimilar monoclonal antibody: Scientific considerations for regulatory approval. Biologicals. 2015;43(1):1–10.

15. Wadhwa M, et al. Immunogenicity assessment of biotherapeutic products: An overview of assays and their utility. *Biologicals,* 2015, 43(5):298–306.

16. Knezevic I, Joung J, Kang HN. Immunogenicity assessment of monoclonal antibody products: a simulated case study correlating antibody induction with clinical outcomes. *Biologicals,* 2015 43(5):307–317.

17. Gupta S, Devanarayan V, Finco D, Gunn GR, Kirshner S, Richards S, et al. Recommendations for the validation of cell-based assays used for the detection of neutralizing antibody immune responses elicited against biological therapeutics. J Pharm Biomed Anal. 2011;55(5):878–88.

18. Shankar G, Devanarayan V, Amaravadi L, Barrett YC, Bowsher R, Finco-Kent D, et al. Recommendations for the validation of immunoassays used for detection of host antibodies against biotechnology products. J Pharm Biomed Anal. 2008;48(5):1267–81.

19. Kang SP, Saif MW. Infusion-related and hypersensitivity reactions of monoclonal antibodies used to treat colorectal cancer: identification, prevention, and management. J Support Oncol. 2007;5(9):451–7.

20. Rup B, Pallardy M, Sikkema D, Albert T, Allez M, Broet P, et al. Standardizing terms, definitions and concepts for describing and interpreting unwanted immunogenicity of biopharmaceuticals: recommendations of the Innovative Medicines Initiative ABIRISK consortium. Clin Exp Immunol. 2015;181(3):385–400.

= = =