Annex 2

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

1. Introduction 97
2. Purpose and scope 98
3. Terminology 99
4. Special considerations for characterization and quality assessment 101
   4.1 Strategy for assessment of mAb biological activity 101
   4.2 Considerations for selection of the expression system 103
   4.3 International standards for biological assays used in the characterization 104
5. Special considerations for nonclinical evaluation 104
   5.1 In vitro studies 104
   5.2 In vivo studies 106
6. Special considerations for clinical evaluation 109
   6.1 Pharmacokinetics studies 109
   6.2 Pharmacodynamics studies 112
   6.3 Comparative clinical efficacy study 113
   6.4 Indication extrapolation 122
   6.5 Pharmacovigilance and post-approval consideration 123

Authors and acknowledgements 123
References 125
Guidelines published by the World Health Organization (WHO) are intended to be scientific and advisory in nature. Each of the following sections constitutes guidance for national regulatory authorities (NRAs) and for manufacturers of biological products. If an NRA so desires, these WHO Guidelines may be adopted as definitive national requirements, or modifications may be justified and made by the NRA.
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACR20</td>
<td>American College of Rheumatology 20% improvement criteria</td>
</tr>
<tr>
<td>ADA</td>
<td>anti-drug antibody</td>
</tr>
<tr>
<td>ADCC</td>
<td>antibody-dependent cellular cytotoxicity</td>
</tr>
<tr>
<td>ADCP</td>
<td>antibody-dependent cellular phagocytosis</td>
</tr>
<tr>
<td>AUC</td>
<td>area under the curve</td>
</tr>
<tr>
<td>CDC</td>
<td>complement-dependent cytotoxicity</td>
</tr>
<tr>
<td>CHO</td>
<td>Chinese hamster ovary</td>
</tr>
<tr>
<td>CLB</td>
<td>competitive ligand-binding (assay)</td>
</tr>
<tr>
<td>CR</td>
<td>complete response</td>
</tr>
<tr>
<td>CRP</td>
<td>C-reactive protein</td>
</tr>
<tr>
<td>DAS28</td>
<td>disease activity score in 28 joints</td>
</tr>
<tr>
<td>DCVMN</td>
<td>Developing Countries Vaccine Manufacturers Network</td>
</tr>
<tr>
<td>EGFR</td>
<td>epidermal growth factor receptor</td>
</tr>
<tr>
<td>ESR</td>
<td>erythrocyte sedimentation rate</td>
</tr>
<tr>
<td>IFPMA</td>
<td>International Federation of Pharmaceutical Manufacturers &amp; Associations</td>
</tr>
<tr>
<td>IgE</td>
<td>immunoglobulin E</td>
</tr>
<tr>
<td>IgG</td>
<td>immunoglobulin G</td>
</tr>
<tr>
<td>IGPA</td>
<td>International Generic Pharmaceutical Alliance</td>
</tr>
<tr>
<td>mAb</td>
<td>monoclonal antibody</td>
</tr>
<tr>
<td>MOA</td>
<td>mechanism of action</td>
</tr>
<tr>
<td>NRA</td>
<td>national regulatory authority</td>
</tr>
<tr>
<td>ORR</td>
<td>overall response rate</td>
</tr>
<tr>
<td>pCR</td>
<td>pathological complete response</td>
</tr>
<tr>
<td>PD</td>
<td>pharmacodynamics</td>
</tr>
<tr>
<td>PK</td>
<td>pharmacokinetics</td>
</tr>
<tr>
<td>RBP</td>
<td>reference biotherapeutic product</td>
</tr>
<tr>
<td>rDNA</td>
<td>recombinant deoxyribonucleic acid</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>------------------------------------------------</td>
</tr>
<tr>
<td>SBP</td>
<td>similar biotherapeutic product</td>
</tr>
<tr>
<td>TK</td>
<td>toxicokinetics</td>
</tr>
<tr>
<td>TMD</td>
<td>target-mediated disposition</td>
</tr>
<tr>
<td>TNF</td>
<td>tumour necrosis factor</td>
</tr>
<tr>
<td>TOST</td>
<td>two one-sided test</td>
</tr>
<tr>
<td>VEGF</td>
<td>vascular endothelial growth factor</td>
</tr>
</tbody>
</table>
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 sources. 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”) based upon the approved mAb innovator products, with a view to making more affordable products that could improve global access to these so-called blockbusters.

Therapeutic mAbs are preparations of an immunoglobulin or a fragment of an immunoglobulin with specificity for a target ligand and are derived from a single clone of cells. Each full-length molecule of a 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 antigen-binding fragment (Fab) region. For full-length mAbs, their crystallizable fragment (Fc) region has the ability to bind to specific receptors, potentially leading to immune effector functions such as antibody-dependent cellular cytotoxicity (ADCC), complement-dependent cytotoxicity (CDC) and antibody-dependent cellular phagocytosis (ADCP). Full-length mAbs are glycoproteins with glycosylation sites in the Fc region of the heavy chains, with further possible glycosylation sites 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 and N- and C-terminal heterogeneity, long half-lives and the potential to induce immunogenicity. Each individual mAb may therefore present a unique profile, which should be taken into consideration during the evaluation of such products as SBPs.

The WHO Guidelines on evaluation of similar biotherapeutic products (SBPs) were adopted by the WHO Expert Committee on Biological Standardization in 2009 (1). This document set out the scientific principles, including the stepwise approach, which should be applied for the demonstration of similarity between an 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 programme 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, as this has been done already for the RBP. The decision on
licensure of the SBP should be based upon 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 for 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, their quality attributes can vary from product to product. Furthermore, one mAb product may have multiple indications. Therefore, biosimilar 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 its 2009 SBP Guidelines to take into account 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 Guidelines, with an additional focus placed on SBPs containing mAbs. All participants, including national regulatory authorities (NRAs) and industry, recognized and agreed that the evaluation principles described in the WHO Guidelines were still valid, valuable and applicable in facilitating the harmonization of SBP requirements globally. It was therefore concluded that there was no need to revise the main body of the existing WHO Guidelines on SBPs. However, it was also agreed that, rather than an amendment, there was a need for additional guidance on the evaluation of biosimilar mAbs.

2. Purpose and scope

The intention of this class-specific document is to set out the specific considerations involved in the evaluation of mAbs developed as SBPs. These WHO Guidelines cover rDNA-derived biosimilar mAbs used in the treatment of human diseases. The principles discussed in this document also apply to mAb-derived proteins – for example, 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, biosimilar mAbs should also comply with the criteria established for demonstration of similarity. Therefore this document should be read in conjunction with both the WHO Guidelines on evaluation of similar biotherapeutic products (SBPs) and the WHO Guidelines on the quality, safety and efficacy of biotherapeutic protein products prepared by recombinant DNA technology (1, 2).
Guidance on various aspects of rDNA-derived medicines, SBPs and mAbs is also available from several other bodies. These WHO Guidelines are not intended to conflict with, but rather to complement, existing relevant regulatory documents.

3. Terminology

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

**American College of Rheumatology 20% improvement criteria (ACR20):** a combined index that measures disease activity in patients with rheumatoid arthritis, and which corresponds to at least a 20% improvement in both the tender joint count and the swollen joint count, and at least a 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.

**Antibody-dependent cellular phagocytosis (ADCP):** an immune mechanism which relies on Fc receptors, especially FcγRIIa, on macrophages or other phagocytic cells which bind to antibodies that are attached to target cells, followed by the phagocytosis and destruction of target cells, including tumour cells.

**Anti-drug antibodies (ADAs):** host antibodies that are capable of binding to a therapeutic antigen (recombinant protein or mAb). This may or may not inactivate the therapeutic effects of the treatment and, in rare cases, induce serious adverse effects.

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

**AUC<sub>t</sub>:** the area under the concentration-time curve of drug in serum or plasma from zero up to a definite time t.

**AUC<sub>tau</sub>:** the area under the concentration-time curve of drug in serum or plasma during a dosage interval.

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

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

**C<sub>max</sub>:** the maximum (peak) serum or plasma concentration observed that a drug achieves in a tested area after the drug has been administered and prior to the administration of a second dose.
**C\textsubscript{min}**: the minimum serum or plasma concentration observed that a drug achieves in a tested area after the drug has been administered and prior to the administration of a second dose.

**C\textsubscript{trough}**: the measured serum or plasma concentration of a drug in a tested area at the end of a dosing interval prior to the administration of the next dose.

**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 the erythrocyte sedimentation rate (ESR) or C-reactive protein (CRP) levels indicating how active the rheumatoid arthritis is, along with a patient's global assessment of their health.

**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.

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

**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

The WHO Guidelines on evaluation of similar biotherapeutic products (SBPs) set out the principle that demonstrating the similarity of a candidate SBP with respect to the RBP in terms of quality is a prerequisite for moving forward to 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 multiple batches, preferably over an extended period, in order to detect possible changes in the quality profile of the RBP over time. The minimum number of batches that 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 for drawing 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 sufficiently sensitive, 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 an α1–6 linkage to the core portion of N-linked carbohydrate chains interferes with the ability of the antibody to bind well to certain Fc receptors, resulting in diminished Fc-mediated activities, including ADCC, whereas an increase in non-reduced terminal galactose can enhance FcγRIIIa binding and ADCC activity. Consequently, the assessment of biological activity of biosimilar mAbs is particularly important and has some unique characteristics. The expression system used for the 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 the WHO SBP and rDNA guidelines (1, 2). Thus the quality aspects covered in this document will focus only on specific considerations for the assessment of mAb biological activity 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 should 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 when 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 that 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 (including Fab- and/or Fc-mediated functions) and ability to bind to Fγ and neonatal Fc receptors (as well as to complement C1q) – should be provided (see section 5.1.2 below).

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. This combined effect may be hard to dissect experimentally when seeking a clear understanding of how the mAb mediates its clinical efficacy. Therefore, if intact MAbs are used, care should 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 (a chimeric mAb specific for CD20) requires Fc function, including ADCC, for its clinical efficacy. Assessment of Fc functions is therefore paramount for this mAb. For infliximab (a tumour necrosis factor alpha (TNFα) antagonist) the neutralization of soluble TNF is the primary MOA while Fc function seems less important. However, ADCC along with other Fc- and Fab-related functions (for example, reverse signalling) also need to be considered as potential secondary MOAs.

Assays for measuring Fc functions can be technically demanding. Differences in both assay formats and cell combinations have significant impact on assay sensitivities. 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 the use of a homogeneous cell population over-expressing the targets/receptors.
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 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 Guidelines (1) allow for the use of different expression systems for production of the SBP compared to the reference product, as 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 mAb development. 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 example, production cells based on mouse cell lines (such as SP2/0 and NS0) secrete mAbs with the carbohydrate structure alpha-gal-1,3-gal present on the carbohydrate moiety. Humans cannot produce the alpha-gal-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 immunoglobulin E (IgE) class and this sensitization can result in anaphylactic reactions (often serious) if they are treated with mouse-cell-line-derived mAbs containing alpha-gal-1,3-gal. Such pre-existing antibodies are particularly evident for cetuximab – an inhibitor of epidermal growth factor receptor (EGFR), which contains an additional glycosylation site on the Fab region that is accessible for IgE binding. Anaphylactic responses may potentially be avoided by using cell substrates of human origin or selected clones of Chinese hamster ovary (CHO) cells for mAb production since these cells normally cannot synthesize alpha-gal-1,3 gal. This type of phenomenon can have important implications for biosimilar mAb development. For example, producing an SBP of cetuximab in mouse cells would probably show the same alpha-gal-1,3-gal-related anaphylaxis problems as the reference product. Although production of the mAb in CHO cells may avoid the anaphylaxis problem (since the alpha-gal-1,3-gal structure would not be likely to be present on the mAb) 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 a biosimilar mAb requires careful consideration, with various potential issues needing to be thoroughly assessed to ensure that an expression system difference does not result in changes to critical quality attributes.

4.3 International standards for biological assays used in the characterization

The development of assays for the determination of biological activity of mAbs will be facilitated by 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 RBP is used for all the comparability studies, whereas WHO International Standards and Reference Reagents are used for calibrating procedures, particularly bioassays, and cannot be used as RBPs. The distinct roles of reference products and international standards are described elsewhere (1, 3).

5. Special considerations for nonclinical evaluation

As with all SBPs undergoing nonclinical evaluation, 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 studies 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 SBP concerned. The approach used should be scientifically justified in the nonclinical overview.

5.1 In vitro studies

5.1.1 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 for detecting differences between the SBP and the reference product than studies in animals, these assays can be considered as paramount for the nonclinical biosimilar comparability exercise.

5.1.2 Biosimilar mAbs – specific aspects
For biosimilar mAbs, the nonclinical in vitro programme should usually include relevant assays for the following specific evaluations:

- Binding studies:
  (a) binding to soluble and/or membrane-bound target antigen(s); and
  (b) binding to representative isoforms of the relevant Fc receptors (that is, for immunoglobulin G (IgG)-based mAbs to FcγRI, FcγRII and FcγRIII), FcRn and complement (C1q).

- Functional studies/biological activities:
  (a) Fab-associated functions (for example, neutralization of a soluble ligand, receptor activation or blockade, reverse signalling via activation of membrane-bound antigen); and
  (b) Fc-associated functions (for example, ADCC, ADCP, CDC), as applicable.

These assays are often technically demanding and the models chosen should be appropriately justified by the applicant (see section 4.1 above). Together, these assays should broadly cover the functional aspects of the mAb even though some may not be considered essential for the therapeutic MOA. However, an evaluation of ADCC, ADCP and CDC may be waived for mAbs directed against non-membrane-bound targets if appropriately justified.
Additional note: as indicated in the ICH Guideline S6(R1) (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 biosimilar comparability.

5.2 In vivo studies

5.2.1 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 the test mAb and the reference mAb, it is at the discretion of NRAs to waive or not to waive a requirement for nonclinical in vivo studies. If the biosimilar quality-comparability exercise and nonclinical in vitro studies are considered satisfactory, and no issues are identified that would block direct entrance into humans, an in vivo animal study may not be considered necessary.

5.2.2 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 (for example, transgenic animals or 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 the principles for mitigating any potential risk.
- When the need for additional in vivo nonclinical studies is evaluated, the factors to be considered include but are not restricted to:
  (a) (the presence of potentially relevant quality attributes that have not been detected in the reference product (for example, new post-translational modification structures);
  (b) the presence of potentially relevant quantitative differences in quality attributes between the SBP and the RBP; and
  (c) relevant differences in formulation (for example, use of excipients not widely used for mAbs).

Although not all of the factors mentioned here necessarily warrant in vivo testing, these factors should be considered together to assess the level of concern and to determine whether or not there is a need for in vivo testing.
- If product-inherent factors that have an impact on pharmacokinetics (PK) and/or biodistribution (such as 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.

5.2.3 Performance of in vivo studies
The following guidance applies to all SBPs, including biosimilar mAbs.

5.2.3.1 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.

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

5.2.3.2 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, if feasible, through 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.

5.2.3.3 Safety studies
If in vivo safety studies are deemed necessary on the basis of a 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 (for example, using just one dose level of SBP and RBP, and/or just one biological sex 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 end-points needed it may not be necessary to kill 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. Where 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 the 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 (that is, to assess nonspecific toxicity only, based on impurities) is not recommended. Because of the different production processes used by the SBP and reference product manufacturers, qualitative differences in process-related impurities will occur (for example, host cell proteins). Such impurities should be kept to a minimum in order to minimize any associated risk.

5.2.3.4 Immunogenicity studies

Qualitative or quantitative difference(s) in product-related variants (for example, glycosylation patterns, charge variants, aggregates and impurities such as host-cell proteins) may have an effect on immunogenic potential and on the potential to cause hypersensitivity. These effects are usually difficult to predict from animal studies and 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 the PK/toxicokinetics (TK) interpretation of in vivo animal studies. Therefore, adequate blood samples should be identified and stored for future evaluations if needed.

5.2.3.5 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 to avoid the need for separate local tolerance studies.

5.2.3.6 Other studies

In general, safety pharmacology and reproductive and development toxicity studies are not warranted in the nonclinical testing of biosimilar mAbs.

In accordance with the WHO Guidelines on the quality, safety and efficacy of biotherapeutic protein products prepared by recombinant DNA technology (2) and ICH Guideline S6(R1) (4), genotoxicity and (rodent) carcinogenicity studies are not required for (similar) biotherapeutic products. This guidance also applies to biosimilar mAbs.
6. Special considerations for clinical evaluation

In general, the goal of the clinical evaluation programme is to confirm that any residual uncertainty about quality attributes or related to the preclinical assessment would not result in clinically meaningful differences – and not to establish the product’s efficacy and safety in a particular indication. The clinical comparability exercise is a stepwise procedure that should begin with PK/PD studies and usually continues with one controlled clinical trial addressing comparative safety and efficacy. In exceptional circumstances, data obtained in clinical PK/PD studies may suffice in confirming biosimilarity established in preceding steps (see section 6.2.1 below). If relevant differences between the SBP and the RBP are detected at any stage, the reasons for the differences should be explored and justified. In reaching a conclusion as to whether a product qualifies as an SBP, the totality of evidence should be considered.

If the original development programme demonstrated that the reference product performed the same in different ethnic groups then there is no scientific rationale for conducting a comparative clinical study of the SBP in each ethnic group.

6.1 Pharmacokinetics studies

6.1.1 Aim of comparative PK studies

Comparative clinical PK studies are always required and should be used to further confirm the similarity of a biosimilar mAb to the RBP already 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 example, a biosimilar mAb 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 lower efficacy (5). Comparative PK studies may be useful in monitoring the impact of the formation of ADAs on efficacy and safety – while exploring the impact of ADAs on PK is also necessary. Both approaches contribute to establishing evidence in support of extrapolation. It is not necessary to study the PK of the biosimilar mAb in every indication that is being sought. In general, one comparative PK study under sensitive conditions that allow any potential differences between the SBP candidate and the RBP to be detected should be sufficient to bridge across the indications for which the reference mAb has been authorized. The design of comparative PK studies depends on various factors, including clinical context, safety profiles and PK characteristics of the reference product (for example, target-mediated disposition (TMD), linear or nonlinear PK, time-dependency and half-life).
6.1.2 **Study design and population**

A single-dose PK study in healthy volunteers is generally recommended as they can be considered a sensitive and homogeneous study population (6). A parallel-group design, which generally requires a higher number of subjects, is usually required for mAbs since a single-dose cross-over design may not be appropriate due to the long half-lives of mAbs and the potential influence of immunogenicity on the PK profile. However, for mAb fragments or mAbs that are not administered systemically, alternative approaches may also be applied.

A number of key issues should be taken into account regarding the use of healthy volunteers to study the PK of mAbs. First, healthy subjects are generally preferred, if possible, because of their higher sensitivity and homogeneity as compared to patient populations. Second, administering a clinically relevant dose of some mAbs (for example, bevacizumab) may not be considered ethical in healthy volunteers because of safety concerns, and in these cases a sub-therapeutic dose on the linear part of the dose–response curve may be required. Third, it may be necessary to perform the PK study for some biosimilar mAbs (for example, rituximab) in a sensitive patient population rather than in healthy volunteers for safety reasons. 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 for the comparative clinical efficacy study 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 ADAs. Furthermore, the choice of a particular population for PK analysis also depends on the range of therapeutic indications of the mAb under development. For example, if a reference mAb is authorized both as an anti-inflammatory agent and as an anticancer antibody (as for example with rituximab) then 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 immunosuppressants or chemotherapeutics. It may be sensible to study the comparative PK in the monotherapy setting in order to minimize sources of variability. When concomitant therapy alters PK, it may be appropriate to study comparative PK both in the monotherapy setting and in combination, particularly where differences cannot be excluded with regard to
quality attributes that might specifically have an impact on how the drug was cleared when used in combination.

6.1.4 **PK characteristics of the reference mAb**
The PK of the mAb may be affected by factors such as the antigen/receptor level (for example, related to tumour burden in oncology), the existence of target-mediated clearance, and/or receptor shedding which has an impact on the 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 product.

6.1.5 **Doses**
A dose should be selected that will enable detection of potential PK differences between the biosimilar mAb 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 nonspecific clearance mechanism dominates. For mAbs that are eliminated by TMD, a low dose (that is, one at which 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 $C_{\text{max}}$, 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 early time points to accurately measure $C_{\text{max}}$ and should also include sufficient sampling time points in the later phases to adequately characterize the late elimination phase. This will allow for reliable estimation of the terminal disposition rate constant and sufficient characterization of 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\textsubscript{0–t}, AUC\textsubscript{tau}, C\textsubscript{max} and C\textsubscript{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 bioanalytical assay to detect both the biosimilar mAb and the reference mAb. The bioanalytical assay should be appropriate for the detection and quantification of mAbs, and should be demonstrated to be bioanalytically comparable with respect to its ability to quantify precisely and accurately both the biosimilar mAb and the reference mAb (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. PK analysis on the ADA-negative samples is of particular interest as it provides 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 should be justified. In some circumstances, narrowing or widening of this margin may be required and this too should be justified.

6.2 **Pharmacodynamics studies**

In general, it is advisable to include PD markers as part of the clinical comparability exercise.

6.2.1 **PD markers and PD assay**

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

The characteristics of PD markers that would support clinical efficacy, and that manufacturers should pay attention to, are (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).

In general, the principles regarding study design, conduct, analysis and interpretation that are relevant to equivalence trials with a clinical outcome as the primary end-point are applicable to equivalence trials with a PD marker as the primary outcome.

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 randomized, adequately powered and preferably double-blinded clinical efficacy study should be performed.

The manufacturer of a biosimilar mAb should perform a thorough analysis of the publicly available clinical data for the reference product to determine the most appropriate study population and primary end-point combination 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 biosimilar mAb could be influenced by several factors, including:

- the nature and complexity of the mAb and derived products;
- the behaviour of the reference product in the clinic;
- the degree of understanding of the MOA of the mAb and disease pathology, and the extent to which these vary in different indications – including MOA, site of action, antigen load, drug administration (dose, route, regimen and 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 of the RBP for which the proposed biosimilar mAb 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 the WHO Guidelines on evaluation of similar biotherapeutic products (SBPs) also 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 equivalence/non-inferiority 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 a biosimilar mAb against the RBP, equivalence trials are generally preferred. Detailed explanations of the advantages and disadvantages of equivalence and non-inferiority trial designs for SBPs are provided in the WHO Guidelines and in guidance developed by other agencies (1, 9–11). Special considerations for clinical trial design in relation to biosimilar mAbs are explained 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 justifying extrapolation to other indications of the RBP. From a statistical perspective, assay sensitivity is important to provide some confidence that the trial, as planned and designed, will be able 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 end-points may deviate from those which led to approval of the RBP for the specific indication as long as the primary end-points are sensitive to the detection of clinically meaningful differences between the biosimilar mAb and the RBP. Whatever approach is taken, applicants should always justify their selection of end-points, time points for analysis and the predefined margin, irrespective of whether this follows the RBP approach or not. If in doubt, applicants may wish to consult relevant regulatory authorities during the planning and design stage of the trial.
The efficacy of the RBP compared to placebo will have 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 should 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, determination of the required sample size should be based on methods specifically designed for equivalence and 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. When 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 biosimilar mAb and the reference mAb, clinical trials of the biosimilar mAb should be carried out in an appropriately sensitive patient population using end-points that are sensitive to the detection of clinically meaningful differences between the SBP and the reference product for the indication (see section 6.3.3 below). The rationale for the study population selected should be provided. In general, using a homogeneous population of patients (for example, same line(s) of therapy, severity or stage of disease progression) will minimize inter-patient variability and thus increase the likelihood of detecting differences between the biosimilar mAb and the reference mAb, if such differences exist. Patients who have not received previous treatment (for example, 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 biosimilar mAb/reference mAb without interference by other medications, as concomitant medications may affect or mask differences in the PK/PD, efficacy, safety and/or immunogenicity of the tested products. To validate the effect of the reference mAb and the sensitivity of the chosen study population, historical data should be used to justify the selection of the study population and equivalence margin. This could generally be done through a systematic review and/or meta-analysis of the relevant studies.
MAbs can function through various MOAs, including agonist activity or receptor blockade (for example, of vascular endothelial growth factor (VEGF) and EGFR), induction of apoptosis, delivery of a drug or cytotoxic agent, and immune-mediated mechanisms (for example, 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, particularly if functional differences are identified in sensitive assays, and especially where it is known that extrapolation to other indications and uses will be sought.

Clinical studies in an unauthorized population (for example, with respect to line of therapy, combination therapy, disease severity or indication authorized in some but not all jurisdictions) may be acceptable for demonstrating “no clinically meaningful differences” for biosimilar mAbs. However, manufacturers of biosimilar mAbs should consult the relevant regulatory authorities prior to conducting such studies.

6.3.3 **Primary study end-point**

Clinically relevant and sensitive study end-points within a sensitive population should be selected to improve the likelihood of detection of potential differences between the biosimilar mAb and the reference product. In general, clinical outcomes, surrogate outcomes or a combination of both can be used as primary end-points in biosimilar mAb trials. The same study end-points used 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. Alternatively, the study end-points used may be different from those traditionally used or from the end-points recommended by study guidelines for innovator products, as more sensitive end-points and/or time points may exist for detecting clinically meaningful differences in an equivalent trial setting where the objective is assessing similarity of efficacy, safety and immunogenicity, and not re-establishing the clinical benefit already demonstrated by the originator. A surrogate end-point can be used as the primary end-point when surrogacy to the clinical outcome is well established or generally accepted, as is the case, for example, with pathological complete response (pCR) in neoadjuvant treatment of breast cancer. The choice of study end-point should always be scientifically justified. More sensitive clinical end-points could be used as secondary end-points for the innovator product, primary or secondary end-points 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 can be considered as end-points for clinical efficacy studies of biosimilar mAbs in oncology trials because these end-points may be more sensitive and are not time related. However, if progression-free survival (which is one of the end-points frequently used for clinical efficacy testing for innovator products)
is considered more sensitive than ORR, then this may be the preferred option. Likewise, both continuous outcomes (for example, changes in DAS28 from baseline) and dichotomous outcomes (for example, ACR20) are considered in rheumatoid arthritis trials for determining clinical comparability (14).

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

NRAs may not always agree on the choice of study end-points. For an SBP 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 end-points 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 data collection depends on the type and severity of known safety issues for the reference product. The SBP study population should be followed to provide information on safety events of interest according to experiences with the reference mAb. Care should be taken to compare the nature, severity and frequency of adverse events between the biosimilar mAb and the reference product in clinical trials that enrolled a sufficient number of patients treated for an acceptable period. 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 of a biosimilar mAb
Therapeutic mAbs, like other rDNA-derived biotherapeutics, may be recognized by the human immune system leading to an unwanted immune response. As mAbs may often be immunogenic in patients, the goal during development of a biosimilar mAb is to demonstrate similar immunogenicity to the reference product. There are some special considerations regarding the immunogenicity of mAbs as compared to other biotherapeutics. For example, mAbs do not evoke cross-reacting antibodies against the body’s endogenous proteins, as some growth factors and proteins for replacement therapy do. However, developing assays to test for 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 the WHO Guidelines on evaluation of similar biotherapeutic products (SBPs) and the WHO Guidelines on the quality, safety and efficacy of biotherapeutic protein products prepared by recombinant DNA technology (1, 2). These general guidance resources should be taken into account when assessing biosimilar mAbs. In addition, further details regarding the advantages and disadvantages of particular assays, as well as some considerations on the interpretation of the results and on the decision-making process, are provided in several review articles (15–18).

The basic data package contains the incidence, titre, neutralization ability and persistence of antibodies against the product/mAb, 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 form the basis of a risk-based immunogenicity programme. It is recommended that the application for marketing authorization of a mAb includes 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
- risk-based immunogenicity programme
- comparative immunogenicity
- assays and mAb characterization
- clinical immunogenicity assessment.

6.3.4.2.1 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 (for example, 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 immunomodulatory therapy and potential pre-existing immunity, antigenicity and sensitivity.

6.3.4.2.2  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 pretreatment, 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. As with other biotherapeutic products a multi-tiered assessment approach is needed. 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 drug target and the residual drug in the sample. To mitigate this 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 removal of the drug. Inclusion of any of these measures should 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 PK measurements, as well as 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 provide 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.

6.3.4.2.3 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.

6.3.4.2.4 Assays and mAb characterization

ADA assays should ideally be capable of detecting all antibodies against both the reference and biosimilar 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 received by each patient. The challenge is to develop two assays with similar sensitivity. Cross-testing all serum samples by both tests is useful for exploring assay performance and antigen epitopes. The use of a single assay with the active substance of the SBP as the antigen/capturing agent for evaluation of all samples (including those from reference-product-treated patients) will be able to detect all antibodies developed against the biosimilar molecule (that is, the conservative approach). In general, the manufacturer should justify the chosen assay approach and should demonstrate the suitability of the method(s) used to measure similarly the immune response against the product received by each patient, irrespective of whether the patient was treated with the RBP or the SBP.
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 assay (for example, a competitive ligand-binding (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 MOA 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 (for example, occurrence of anaphylaxis or use of certain assay formats). It may also be necessary to locate the antigenic sites (for example, antigen-binding region versus constant region of the antibody molecule). The banking of patient samples is necessary in order to retain the possibility for retesting in case of technical problems in the original assay.

6.3.4.2.5 Clinical immunogenicity assessment

The selected patient population should be sensitive for the detection of differences in immunogenicity. It is also important that the controlled 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 reference products.

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 (for example, when the reference product is known to have serious but rare immune-mediated effects, such as anaphylaxis).
Evaluation of immunogenicity includes antibody incidence, titre, neutralization capacity and persistency, as well as correlations to exposure, safety and efficacy. Currently, there is no generally accepted statistical methodology that could be used to define the limits of comparable immunogenicity. In general, an increase in immunogenicity of an SBP when compared with the RBP is incompatible 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 regulatory and scientific process of extending information and conclusions available from one patient population in order to make inferences for other populations. In the context of SBPs, it refers to granting a clinical indication to an SBP for which the reference product is authorized, without conducting clinical efficacy and safety studies to support that indication. Extrapolation cannot be claimed automatically for all indications of the reference product and requires sound scientific justification based on the totality of evidence. 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 biosimilarity. Current WHO guidance on SBP evaluation (1) sets out a number of recommended principles regarding the extrapolation of clinical data across indications which also apply to biosimilar mAbs. 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 or 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 – for example, ligand blockade, receptor blockade, receptor down-regulation, cell depletion (via ADCC, CDC or 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. In order to support extrapolation, the mechanisms that contribute to the efficacy of the mAb in each indication should ideally be well
understood and clearly defined. In practice, this is often not the case. Therefore, extrapolation may pose additional challenges when a mAb is indicated for a variety of diseases in which the MOAs are not the same or are not well understood for each indication. In this situation, it is important to explore the comparability of in vitro functions of the mAb. In cases where significant functional differences exist, further nonclinical or clinical data are needed to support extrapolation. 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 accordingly (for example, ADCC assays under different conditions). If minor quality differences are found, and the affected mechanism is not considered active in the studied indication, additional steps may be necessary to reach a conclusion on biosimilarity. Additional data, with appropriate supporting scientific rationale, could include quality, preclinical and/or PK/PD data and might impact on the selection of the final clinical, safety and efficacy study. 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 should be put 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 guidelines (1, 2) it is essential to record the product brand name, batch number and manufacturers’ name, and, where it exists, the International Nonproprietary Name (INN). 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 their relatively small sample size, biosimilar mAb clinical trials may have the statistical power to detect only common adverse events. Thus, as for any biological medicine, pharmacovigilance is essential in order to detect potential overt new or rare biosimilar mAb-specific safety issues and to allow for the identification and assessment of potential post-marketing risks.

Authors and acknowledgements

The first draft of these WHO Guidelines was prepared by Dr K. Gao, World Health Organization, Switzerland; Dr E. Griffiths, Consultant, Kingston-upon-Thames, England; Dr H-K. Heim, Federal Institute for Drugs and Medical Devices, Germany; Dr I. Knezevic, World Health Organization, Switzerland; Dr P. Kurki, Finnish Medicines Agency, Finland; Dr N. Ekman, Finnish Medicines
Agency, Finland; Dr C. Njue, Health Canada, Canada; Dr R. Thorpe, Consultant, Welwyn, England; Dr M. Wadhwa, National Institute for Biological Standards and Control, England; Dr J. Wang, Health Canada, Canada; and Dr E. Wolff-Holz, Paul-Ehrlich-Institut, Germany.

The drafting group took into consideration proposals made during a WHO informal consultation on monoclonal antibodies as similar biotherapeutic products (SBPs) held in Geneva, Switzerland, 27–28 April 2015 and attended by: Mrs A. Abas, Ministry of Health Malaysia, Malaysia; Dr A. Abdelaziz, Jordan Food and Drug Administration, Jordan; Dr K. Bangarurajan, Central Drug Standards Control Organization, India; Mrs J. Bernat (International Federation of Pharmaceutical Manufacturers & Associations (IFPMA) representative), IFPMA, Switzerland; Dr M.C. Bielsky, Medicines and Healthcare products Regulatory Agency, England; Dr B. Boonyapiwat, Ministry of Public Health, Thailand; Dr S. Bukoñzer (International Generic Pharmaceutical Alliance (IGPA) representative), Hospira, the USA; Ms J. Dahlan, National Agency of Drug and Food Control, Indonesia; Mrs D. Darko, Food and Drug Authority, Ghana; Mr D. Goryachev, Ministry of Health, Russian Federation; Dr E. Griffiths, Consultant, Kingston-upon-Thames, England; Dr K. Heidenreich (IFPMA representative), Novartis International AG, Switzerland; Dr H-K. Heim, Federal Institute for Drugs and Medical Devices, Germany; Dr C. Ilonze, National Agency for Food and Drug Administration and Control, Nigeria; Mr R. Ivanov, Biocad, Russian Federation; Mrs W. Jariyapan, Ministry of Public Health, Thailand; Dr J. Joung, Ministry of Food and Drug Safety, Republic of Korea; Dr D. Khokal, Health Sciences Authority, Singapore; Mr B. Kim, Celltrion, Republic of Korea; Mrs S. Kox (IGPA representative), Medicines for Europe, Belgium; Dr A. Kudrin, Celltrion, Republic of Korea; Dr P. Kurki (on behalf of the Biosimilar Medicinal Products Working Party, European Medicines Agency), Finnish Medicines Agency, Finland; Ms I. Lyadova, Ministry of Health, Russian Federation; Dr J. Macdonald (IFPMA representative), Pfizer, the United Kingdom; Mrs S. Marlina, National Agency of Drug and Food Control, Indonesia; Dr J. Meriakol; Ministry of Health, Kenya; Mr K. Nam, Ministry of Food and Drug Safety, Republic of Korea; Dr C. Njue, Health Canada, Canada; Mrs Y.H. Nunez, Centro para el Control Estatal de la Calidad de los Medicamentos, Cuba; Ms H.B. Pedersen, WHO Regional Office for Europe, Denmark; Dr M.L. Pomran, Pan American Health Organization, the USA; Dr A. Qu (Developing Countries Vaccine Manufacturers Network (DCVMN) representative), Shanghai Institute of Biological Products Co., Ltd, China; Dr S. Ramanan (IFPMA representative), Amgen Inc., the USA; Dr T. Schreitmueller (IFPMA representative), F. Hoffmann–La Roche Ltd, Switzerland; Dr M. Schiestl (Medicines for Europe representative), Sandoz GmbH, Austria; Dr J. Siegel (IFPMA representative), Johnson & Johnson, the USA; Dr E. Spitzer (Asociación Latinoamericana de Industrias Farmacéuticas (ALIFAR) representative), Argentina;
Ms I. Susanti (DCVMN representative), PT Bio Farma, Indonesia; Dr R. Thorpe, Consultant, Welwyn, England; Dr C.P. Vaca Gonzalez, Ministry of Health, Colombia; Ms B. Valente, Agência Nacional de Vigilância Sanitária, Brazil; Ms N. Vergel, Instituto Nacional de Vigilância de Medicamentos y Alimentos, Colombia; Dr M. Wadhwa, National Institute for Biological Standards and Control, England; Dr J. Wang, Health Canada, Canada; Dr J. Wang, National Institutes for Food and Drug Control, China; Dr L. Wang, National Institutes for Food and Drug Control, China; Dr K. Watson (IFPMA representative), AbbVie Ltd, the United Kingdom; Dr C. Webster (IGPA representative), Baxter, the USA; Dr E. Wolff-Holz, Paul-Ehrlich-Institut, Germany; Dr S. Xie, China Food and Drug Administration, China; Dr T. Yamaguchi, Pharmaceuticals and Medical Devices Agency, Japan; Dr A. Zhang (DCVMN representative), China National Biotec Group Co. Ltd, China; and Dr K. Gao, Dr H. Kang, Dr I. Knezevic and Dr D. Wood, World Health Organization, Switzerland.

The document WHO/BS/2016.2290 was prepared by the drafting group and posted on the WHO Biologicals website for a round of public consultation from 22 July to 16 September 2016, with feedback received from both regulators and industry.

Further changes were subsequently made to document WHO/BS/2016.2290 by the WHO Expert Committee on Biological Standardization.

References


