



EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

30 May 2012
EMA/CHMP/BMWP/403543/2010
Committee for Medicinal Products for Human Use (CHMP)

Guideline on similar biological medicinal products containing monoclonal antibodies – non-clinical and clinical issues

Draft Agreed by Similar Biological Medicinal Products Working Party	October 2010
Adoption by CHMP for release for consultation	18 November 2010
End of consultation (deadline for comments)	31 May 2011
Final agreed by BMWP	March 2012
Adoption by CHMP	30 May 2012
Date for coming into effect	1 December 2012

Keywords	<i>Biosimilars, monoclonal antibodies, similar biological medicinal products, relevant animal model, non-clinical studies, in vitro studies, clinical use, clinical endpoints, extrapolation</i>
-----------------	---



Guideline on Similar Biological Medicinal Products Containing Monoclonal Antibodies – Non-clinical and Clinical Issues

Table of contents

Executive summary	3
1. Introduction	4
2. Scope.....	4
3. Legal basis and relevant guidelines	4
4. Non-clinical studies	5
4.1. In vitro studies = step 1.....	6
4.2. Determination of the need for in vivo studies = step 2	6
4.3. In vivo studies = step 3	7
5. Clinical studies	7
5.1. Pharmacokinetics (PK) = step 1	8
5.2. Pharmacodynamics (PD)	11
5.3. Clinical Efficacy – step 2.....	11
5.4. Clinical Safety	13
6. Extrapolation of indications.....	15
7. Pharmacovigilance	15

Executive summary

This guideline lays down the non-clinical and clinical requirements for monoclonal antibody (mAb) containing medicinal products claiming to be similar to another one already authorised. The non-clinical section addresses the pharmaco-toxicological requirements and the clinical section the requirements for pharmacokinetic, pharmacodynamic, efficacy and safety studies as well as pharmacovigilance aspects.

The overall aim of this guideline is to establish general principles that enable applicants for a development programme that establishes comparability of a biosimilar mAb with a reference mAb, while ensuring that the previously proven safety and efficacy of the drug is conserved. A stepwise approach is normally recommended throughout the development programme, and the extent and nature of the non-clinical and clinical programme depends on the level of evidence obtained in the previous step(s). All studies should be planned with the intention to detect any potential differences between biosimilar and reference medicinal product and to determine the relevance of such differences, should they occur.

As regards non-clinical development, a step-wise approach to evaluate mAbs is recommended to decide on the choice and extent of *in vitro* and *in vivo* studies on a case-by-case basis. Comparative *in vitro* studies to assess differences in binding or functions should be conducted first. In a second step, it should be determined whether additional *in vivo* non-clinical work is warranted. If an *in vivo* study is deemed necessary, the focus of the study depends on the need for additional information, and the availability of a relevant animal model. The conduct of toxicological studies in non-human primates is usually not recommended.

During the clinical development programme, patients are usually enrolled commensurate with the level of evidence obtained from preceding steps which support comparability. A comparative pharmacokinetic study in a sufficiently sensitive and homogeneous study population (healthy volunteers or patients) normally forms an initial step of biosimilar mAb development. Pharmacokinetic data can be helpful to extrapolate data on efficacy and safety between different clinical indications of the reference mAb. It may, on a case-by-case basis, be necessary to undertake multidose pharmacokinetic studies in patients, or even to perform pharmacokinetic assessment as part of the clinical study designed to establish similar efficacy and safety. Pharmacokinetic studies can be combined with pharmacodynamic (PD) endpoints, where available. Normally, similar clinical efficacy should be demonstrated in adequately powered, randomised, parallel group comparative clinical trial(s), preferably double-blind, normally equivalence trials. To establish comparability, deviations from disease-specific guidelines issued by the CHMP may be warranted. The guiding principle is to demonstrate similar clinical efficacy and safety compared to the reference medicinal product, not patient benefit *per se*, which has already been shown for the reference medicinal product. In principle, the most sensitive model and study conditions (pharmacodynamic or clinical) should be used in a homogeneous patient population. In cases where comparative pharmacodynamic studies are claimed to be most suitable to provide the pivotal evidence for similar efficacy, applicants will have to choose clinically relevant markers, justify these markers, and also provide sufficient reassurance of clinical safety, particularly immunogenicity. Extrapolation of clinical efficacy and safety data to other indications of the reference mAb, not specifically studied during the clinical development of the biosimilar mAb, is possible based on the results of the overall evidence provided from the comparability exercise and with adequate justification. As regards post-authorisation follow-up, the concept to be proposed by applicants may have to exceed routine pharmacovigilance, and may have to involve post-authorisation safety studies (PASS).

1. Introduction

Monoclonal antibodies have been established as a major product class of biotechnology-derived medicinal products. Different mAb products share some properties, e.g. being cytotoxic to their target, or neutralizing a cytokine, but differ in aspects like the mechanism of action. They are structurally complex, and may have several functional domains within a single molecule, depending on the isotype (antigen-binding region, complement-binding region, constant part interacting with Fc receptors). Each individual mAb presents a unique profile with respect to the antigen-binding region, the Fc cytotoxic effector function, and binding to Fc receptors. Various assays have been established in the past years that allow for more in-depth characterisation of complex proteins, both on a physicochemical and a functional level, e.g. with potency assays, and there is experience in the assessment of minor quality differences due to changes in manufacturing processes for monoclonal antibodies. However, it may at the current stage of knowledge be difficult to interpret the relevance of minor quality differences in the physicochemical and biological characterization when comparing a biosimilar mAb to a reference mAb.

This guideline lays down the non-clinical and clinical requirements for monoclonal antibody-containing medicinal products claimed to be similar to another one already authorised, i.e. similar biological medicinal products (biosimilars). The studies described here should be planned with the intention to detect any potential differences between biosimilar and reference medicinal product and to determine the relevance of such differences, should they occur. A biosimilar mAb should be similar to the reference mAb in physicochemical and biological terms. Any observed relevant difference would have to be duly justified and could contradict the biosimilar principle. For quality aspects the principles as laid out in the guidelines on biosimilars including the "Guideline on similar biological medicinal products containing biotechnology-derived proteins as active substance: Quality issues" (EMA/CHMP/BWP/49348/2005), and the "Guideline on development, production, characterisation and specifications for monoclonal antibodies and related substances" (EMA/CHMP/BWP/157653/2007) apply.

2. Scope

The "Guideline on similar biological medicinal products containing biotechnology-derived proteins as active substance: non-clinical and clinical issues" (EMA/CHMP/42832/2005/) lays down the general requirements for demonstration of the similar nature of two biological products in terms of safety and efficacy. This guidance specific for monoclonal antibodies complements the above guideline and presents the current view of the CHMP on the application of the guideline for demonstration of comparability of two mAb-containing medicinal products from a Marketing Authorisation Application (MAA) perspective. While this guidance is specifically related to mAbs, the principles discussed can also be applied to related substances like for example fusion proteins based on IgG Fc (-cept molecules).

Next-generation mAbs, defined as mAbs that are structurally and/or functionally altered (for example, glyco-engineered mAbs with higher potency), in comparison to already licensed reference medicinal products, to achieve an improved or different clinical performance, are not biosimilars and therefore beyond the scope of this guideline.

3. Legal basis and relevant guidelines

Directive 2001/83/EC, as amended in particular in Directive 2001/83/EC Art 10(4) and Part II of the Annex I of Directive 2001/83/EC, as amended. In addition, in particular, the following guidelines should be read in conjunction:

- Guideline on similar biological medicinal products (CHMP/437/04)
- Guideline on similar biological medicinal products containing biotechnology-derived proteins as active substance: non-clinical and clinical issues (EMA/CHMP/42832/2005)
- Guideline on similar biological medicinal products containing biotechnology-derived proteins as active substance: Quality issues (EMA/CHMP/BWP/49348/2005)
- Guideline on development, production, characterisation and specifications for monoclonal antibodies and related substances (EMA/CHMP/BWP/157653/07)
- ICH guideline S 6 (R1) Preclinical safety evaluation of biotechnology-derived pharmaceuticals (EMA/CHMP/ICH/731268/1998)
- Guideline on the clinical investigation of the pharmacokinetics of therapeutic proteins (CHMP/EWP/89249/2004)
- Guideline on the investigation of bioequivalence (CPMP/EWP/QWP/1401/98)
- Guideline on bioanalytical method validation (EMA/CHMP/EWP/192217/2009)
- Guideline on the evaluation of anticancer medicinal products in man (CHMP/EWP/205/95)
- ICH E10 Choice of Control Group in Clinical Trials (CPMP/ICH/364/96)
- Guideline on the choice of a non-inferiority margin (EMA/CPMP/EWP/2158/99)
- Reflection paper on the extrapolation of results from clinical studies conducted outside Europe to the EU-population (EMA/CHMP/EWP/692702/08)
- Guideline on Immunogenicity Assessment of Biotechnology-derived Therapeutic Proteins (EMA/CHMP/BMWP/14327/06)
- Guideline on Immunogenicity assessment of monoclonal antibodies intended for in vivo clinical use (EMA/CHMP/BMWP/86289/2010)
- Eudralex Volume 9A of The Rules Governing Medicinal Products in the European Union – Guidelines on Pharmacovigilance for Medicinal Products for Human Use - to be replaced by Good Vigilance Practice after July 2012
- ICH E2A Clinical Safety Data Management: Definitions and Standards for Expedited Reporting (CPMP/ICH/377/95)
- ICH E2E Note for guidance on Planning Pharmacovigilance Activities (CPMP/ICH/5716/03)

4. Non-clinical studies

As regards non-clinical development, a step-wise approach is applied to evaluate the similarity of biosimilar and reference mAb.

Non-clinical studies should be performed before initiating clinical trials. *In vitro* studies should be conducted first and a decision then made as to the extent of what, if any, *in vivo* work will be required.

The approach taken will need to be fully justified in the non-clinical overview.

4.1. *In vitro* studies = step 1

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

In vitro non-clinical studies should be performed with an appropriate number of batches of product representative of that intended to be used in the clinical trial. These studies should include relevant assays on:

- Binding to target antigen(s)
- Binding to representative isoforms of the relevant three Fc gamma receptors (FcγRI, FcγRII and FcγRIII), FcRn and complement (C1q)
- Fab-associated functions (e.g. neutralization of a soluble ligand, receptor activation or blockade)
- Fc-associated functions (e.g. antibody-dependent cell-mediated cytotoxicity, ADCC; complement-dependent cytotoxicity, CDC; complement activation)

These studies should be comparative in nature and should be designed to be sensitive enough to detect differences in the concentration–activity relationship between the similar biological medicinal product and the reference medicinal product, and should not just assess the response per se. It should be noted that an evaluation of ADCC and CDC is generally not needed for mAbs directed against non-membrane bound targets. As indicated in the ICH S6 (R1) guideline, tissue cross-reactivity studies are not suitable to detect subtle changes in critical quality attributes and are thus not recommended for assessing comparability.

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. As the *in vitro* assays may be more specific and sensitive than studies in animals, these assays can be considered paramount in the non-clinical comparability exercise.

If the comparability exercise using the above strategy indicates that the test mAb and the reference mAb cannot be considered biosimilar, it may be more appropriate to consider developing the product as a stand alone.

4.2. *Determination of the need for in vivo studies = step 2*

It is acknowledged that some mAbs may mediate effects that cannot be fully elucidated by *in vitro* studies. Therefore, evaluation in an *in vivo* study may be necessary, provided that a relevant *in vivo* model with regard to species or design is available. Factors to be considered when the need for additional *in vivo* non-clinical studies is evaluated, include but are not restricted to:

- Presence of relevant quality attributes that have not been detected in the reference product (e.g. new post-translational modification structure)
- Presence of quality attributes in significantly different amounts than those measured in the reference product
- 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 issues should be considered together to assess the level of concern and whether there is a need for *in vivo* testing.

If the comparability exercise in the *in vitro* studies in step 1 is considered satisfactory and no factors of concern are identified in step 2, or these factors of concern do not block direct entrance into humans, an *in vivo* animal study may not be considered necessary.

If there is a need for additional information, the availability of a relevant animal species or other relevant models (e.g. transgenic animals or transplant models) should be considered. Due to the specificity of mAbs, the relevant species studied is in most cases a non-human primate. In all cases the limitations of an *in vivo* study (such as sensitivity and variability) should be taken into account.

If a relevant *in vivo* animal model is not available the applicant may choose to proceed to human studies taking into account principles to mitigate any potential risk.

4.3. *In vivo* studies = step 3

If an *in vivo* study is deemed necessary, the focus of the study (PK and/or PD and/or safety¹) depends on the need for additional information. Animal studies should be designed to maximise the information obtained. In addition, depending on the endpoints needed, it may not be necessary to sacrifice the animals at the end of the study. The principles of the 3Rs (replacement, refinement, reduction) should be considered when designing any *in vivo* study. The duration of the study (including observation period) should be justified, taking into consideration the PK behaviour of the mAb and its clinical use.

When the model allows, the PK and PD of the similar biological medicinal product and the reference medicinal product should be quantitatively compared, including concentration-response assessment covering the therapeutic doses in humans.

The conduct of repeated dose toxicity studies in non-human primates is usually not recommended. Also, the conduct of toxicity studies in non-relevant species (i.e. to assess unspecific toxicity only, based on impurities) is not recommended. Due to 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 level should be kept to a minimum, which is the best strategy to minimise any associated risk. Qualitative or quantitative difference(s) of product-related variants (e.g. glycosylation patterns, charge variants) may affect biological functions of the mAb and are expected to be evaluated by appropriate *in vitro* assays. These quality differences may have an effect on immunogenic potential and potential to cause hypersensitivity. It is acknowledged that these effects are difficult to predict from animal studies and should be further assessed in clinical studies. Immunogenicity assessment in animals is generally not predictive for immunogenicity in humans, but may be needed for interpretation of *in vivo* studies in animals. Blood samples should be taken and stored for future evaluations if then needed.

Studies regarding safety pharmacology and reproduction toxicology are not required for non-clinical testing of biosimilar mAbs. Studies on local tolerance are usually not required. If excipients are introduced for which there is no or little experience with the intended clinical route, local tolerance may need to be evaluated. If other *in vivo* studies are performed, evaluation of local tolerance may be part of the design of that study instead of the performance of separate local tolerance studies.

5. Clinical studies

Comparative clinical studies between the biosimilar and reference medicinal product should always be conducted. The number and type of studies might vary according to the reference product and should be justified based on a sound scientific rationale. A stepwise approach is normally recommended

¹ "Safety" in this context does not usually refer to a complete repeated dose toxicity study, but rather an in-life evaluation of safety parameters such as clinical signs, body weight and vital functions.

throughout the development programme, and the extent and nature of the clinical programme depends on the level of evidence obtained in the previous step(s). During the clinical development programme, patients are normally enrolled commensurate with the level of evidence obtained from preceding steps which support comparability.

5.1. Pharmacokinetics (PK) = step 1

The comparison of the pharmacokinetic properties of the biosimilar product and the reference medicinal product forms normally the first step of a biosimilar mAb development. The design of the study depends on various factors, including clinical context, safety, the PK characteristics of the antibody (target-mediated disposition, linear or non-linear PK, time-dependencies, half-life, etc.) and should take into account the recommendations as outlined in the "Guideline on the clinical investigation of the pharmacokinetics of therapeutic proteins" (CHMP/EWP/89249/2004) and the "Guideline on the investigation of bioequivalence" (CPMP/EWP/QWP/1401/98 Rev. 1/ Corr). Furthermore, bioanalytical assays should be appropriate for their intended use and adequately validated as outlined in the "Guideline on bioanalytical method validation" (EMA/CHMP/EWP/192217/2009).

5.1.1. Study design

The primary objective of the pharmacokinetic studies performed to support a MAA for a biosimilar is to show comparability in pharmacokinetics of the biosimilar with the reference medicinal product in a sufficiently sensitive and homogeneous population. This is expected to reduce variability, and thus the sample size needed to prove equivalence, and can simplify interpretation.

Healthy volunteers are likely to have less variability in PK as target-mediated clearance may be less important than in patients. Hence, if feasible, a single dose study in healthy volunteers is recommended, which could provide important information on biosimilarity. From a pharmacokinetic perspective, a single dose cross-over study with full characterisation of the PK profile, including the late elimination phase, is preferable. A parallel group design may be necessary due to the long half-life of mAbs and the potential influence of immunogenicity.

A study in healthy volunteers may not be possible in case of a toxic mechanism of action, or in case the information obtained would not be sufficient to establish biosimilarity. Under these circumstances a study in patients may be a better option. If a single dose study is not feasible in patients, a multiple dose study should be conducted.

It may be necessary to perform the PK study in a different population from that selected to establish similar clinical efficacy, since the most sensitive population where PK characteristics can be compared may not be the same as the most sensitive population where similar efficacy and safety can be demonstrated. In such scenarios, population PK measurements during the clinical efficacy trial are recommended since such data may add relevant data to the overall database to claim comparability.

The choice of the patient population for the PK study should be fully justified, based on a comprehensive survey of scientific literature as regards its sensitivity, and also the possibility to infer PK results to other clinical indications where the reference mAb is licensed.

In case a PK study in healthy volunteers is conducted to support bioequivalence, supportive PK data from clinical studies in patients are encouraged and could provide highly supportive evidence of a similar PK behaviour.

The following factors impact on the strategy of designing PK evaluations:

Disease and patient characteristics

Factors that may influence the choice of the patient population are age of usual manifestation and age range (since lower age may be less prone to presence of concomitant clinical conditions), number of previous treatments, concomitant treatments, or expression of antigen (which may be related to disease stage). For mAbs which are indicated for both monotherapy and in combination with immunosuppressant or chemotherapy, it may be sensible to study the comparative PK in the monotherapy setting in order to minimize the sources for variability. However, a first line setting, where patients are in a better clinical condition, or an adjuvant setting in patients with early cancer, where the tumour burden is low, may be preferable; in these instances, the mAb is typically administered in combination with other therapies.

PK characteristics of the reference mAb

Pharmacokinetics of anticancer mAbs may be time dependent, as the tumour burden may change after multiple dosing (e.g. increased half-life with multiple dosing) and this should be taken into account in the design of the study.

The existence of target-mediated clearance in addition to non target-mediated clearance may affect the number of studies needed. In case target-mediated clearance is not relevant, one comparative PK study may be sufficient. If the reference mAb is eliminated both by target-mediated and non target-mediated mechanisms, comparable PK should be demonstrated where each mechanism of clearance predominates: preferably one study in healthy volunteers for non target-mediated clearance and one supportive study in patients, which can be part of the efficacy trial, to investigate comparability in target-mediated clearance.

For mAb targets that involve receptor shedding, it is advisable to measure shed receptor levels at baseline and, if relevant, during the conduct of the study, in order to verify the baseline comparability of the treatment groups. Stratification by tumour burden or receptor shedding, if possible, may help to ensure baseline comparability. An exploratory statistical analysis on post-baseline comparability at the time point relevant to the conclusion of PK equivalence could be helpful.

For mAbs licensed in several clinical indications, it is not generally required to investigate the pharmacokinetic profile in all of them. However, if distinct therapeutic areas are involved for one particular mAb (e.g. autoimmunity and oncology), separate PK studies may be needed if different target-mediated clearance exists for different therapeutic areas.

Doses

In principle, it is not required to test all therapeutic dosage regimens; the most sensitive dose should be selected to detect potential differences in PK between the biosimilar and the reference products. When limited data are available to know which dose is the most sensitive it is recommended to investigate a low or the lowest recommended therapeutic dose where it is assumed that the target-mediated clearance is not yet saturated and a high or the highest therapeutic dose where it is believed that the non specific clearance mechanism dominates. A single dose study with the lowest therapeutic dose in patients is considered the most adequate design to investigate the differences in target-mediated clearance, if any.

Routes of administration

If the reference product can be administered intravenously and subcutaneously and if both routes are applied for, it is preferable to investigate both routes of administration. However, as the evaluation of subcutaneous administration covers both absorption and elimination, it may be possible to waive the

evaluation of intravenous administration if comparability in both absorption and elimination has been demonstrated for the subcutaneous route using additional PK parameters such as partial AUCs (see 5.1.2).

5.1.2. Sampling times

In single dose studies, the sampling times should be selected to characterise the whole profile, including the late elimination phase. For those products administered as two (or more) consecutive doses useful information can be obtained from both the first and last administrations since the first administration is preferred for comparative purposes and the last one can provide information on the final elimination phase that cannot be observed after the first dose.

If a multiple dose PK study in patients is used to show similarity between the biosimilar and reference medicinal product and if elimination after the last dose cannot be characterised, sampling should normally be undertaken to characterise the concentration-time profile both after the first dose and later, preferably at steady state. Characterisation of the full concentration-time profile at steady state is especially important in case of non-linear PK of the reference mAb (e.g. many anticancer mAbs with cellular targets exhibit dose- or time-dependent PK or immunogenicity-related changes in distribution or elimination kinetics).

5.1.3. PK parameters of interest

In a single dose study, the primary parameter should be the $AUC_{(0-inf)}$. Secondary parameters such as C_{max} , t_{max} , volume of distribution, and half-life, should also be estimated. In case of subcutaneous administration, C_{max} should be a co-primary parameter. In addition, if no data are provided for the intravenous route, partial AUCs should be assessed to ensure comparability of both absorption and elimination.

In a multiple dose study, the primary parameters should be the truncated AUC after the first administration until the second administration (AUC_{0-t}) and AUC over a dosage interval at steady state (AUC_{τ}). Secondary parameters are C_{max} and C_{trough} at steady state.

Anti-drug antibodies should be measured in parallel of PK assessment using the most appropriate sampling time points.

Comparability margins have to be defined a priori and appropriately justified. For some reference mAbs, inter-subject variability for some parameters was reported to be considerable. This may have to be accounted for in the choice of the comparability margin at least for such parameters. As a principle, any widening of the conventional equivalence margin beyond 80-125% for the primary parameters requires thorough justification, including an estimation of potential impact on clinical efficacy and safety. For secondary parameters, confidence intervals (CI) for ratio or differences can be presented together with descriptive statistics but no acceptance range needs to be defined. The clinical relevance of estimated differences and associated confidence intervals should be discussed.

5.1.4. Timing of the PK evaluation

Usually, proof of similar PK profiles should precede clinical efficacy trials. However, in certain scenarios, e.g. for mAbs where PK is inevitably highly variable even within one clinical indication, it may, for feasibility reasons, be necessary to explore PK comparisons as part of a clinical study that is designed to establish similar clinical efficacy (as only this trial will be large enough to demonstrate PK equivalence). To start with a comparative clinical efficacy trial that includes PK evaluation, without a formal preceding comparative PK study, could be problematic with no former human exposure to the

biosimilar mAb, together with potentially limited non-clinical *in vivo* data, depending on the mAb. Therefore, such a plan could only be justified on a case by case basis depending on the product profiles observed in the quality and non-clinical data.

5.2. Pharmacodynamics (PD)

Pharmacodynamic parameters may contribute to the comparability exercise for certain mAbs and in certain indications. Depending on the mAb and availability of PD endpoints, the following scenarios are, theoretically, possible:

5.2.1. PD markers as support to establish comparability

Pharmacokinetic studies can be combined with pharmacodynamic (PD) endpoints, where available. This could add valuable information for the overall comparability exercise. PD markers are especially valuable if they are sensitive enough in order to detect small differences, and if they can be measured with sufficient precision. The use of multiple PD markers, if they exist, is recommended. With regard to pharmacodynamic evaluation, there is often a lack of specific PD endpoints. The emphasis may then have to be on non-clinical PD evaluations, e.g. *in vitro* testing.

5.2.2. PD markers as pivotal proof of comparability

Sponsors should always explore possibilities to study dose-concentration-response relationships or time-response relationships, since this approach, if successful, may provide strong evidence of comparability, provided that the selected doses are within the linear part of the dose-response curve.

The following prerequisites need to be met to accept that PD markers can constitute the pivotal evidence for the efficacy comparability exercise:

- A clear dose-response relationship is shown.
- At least one PD marker is an accepted surrogate marker and can be related to patient outcome to the extent that demonstration of similar effect on the PD marker will ensure a similar effect on the clinical outcome variable.

If that is not the case, then proceed to Step 2 (i.e. clinical efficacy).

When PD markers are planned as pivotal evidence to establish similarity, it is recommended to discuss such approach with regulatory authorities. This should include a proposal of the size of the proposed equivalence margin and its clinical justification as regards lack of a clinically meaningful difference.

A comparative single or repeat dose study in the saturation part of the dose-concentration-response curve is unlikely to discriminate between different activities, should they exist, and a dose in the linear part of the dose-response curve may result in treating a patient with a too low dose. It is also acknowledged that dose-response data may not exist for the reference mAb, and that exposing patients to a relatively low dose of the mAbs, in a worst case scenario, might also sensitize them to develop anti-mAb antibodies, and, consequently, may make them treatment resistant. However, for some reference mAbs clinical conditions may exist where such studies are feasible.

5.3. Clinical Efficacy – step 2

If dose comparative and highly sensitive PD studies cannot be performed convincingly showing comparability in a clinically relevant manner, similar clinical efficacy between the similar and the

reference product should be demonstrated in adequately powered, randomised, parallel group comparative clinical trial(s), preferably double-blind, normally equivalence trials.

For most of the clinical conditions that are licensed for mAbs, specific CHMP guidance on the clinical requirements to demonstrate efficacy exists. However, to establish comparability, deviations from these guidelines (choice of endpoint, timepoint of analysis of endpoint, nature or dose of concomitant therapy, etc) will be warranted in some circumstances. Such deviations need to be scientifically justified on the basis that the proposed clinical concept is designed to establish biosimilarity by employing PD markers, clinical outcomes, or both. The guiding principle is to demonstrate similar efficacy and safety compared to the reference medicinal product, not patient benefit *per se*, which has already been established by the reference medicinal product. Therefore, in general the most sensitive patient population and clinical endpoint is preferred to be able to detect product-related differences, if present and, at the same time, to reduce patient and disease-related factors to a minimum in order to increase precision and to simplify interpretation. For example, patients with different disease severity and with different previous lines of treatment might be expected to respond differently, and thus differences between the study arms may be difficult to interpret, and it may remain uncertain whether such differences would be attributable to patient or disease related factors rather than to differences between the biosimilar mAb and reference mAb.

Comparability should be demonstrated in scientifically appropriately sensitive clinical models and study conditions (whether licensed or not), and the applicant should justify that the model is relevant as regards efficacy and safety, and sensitive to demonstrate comparability in the indication(s) applied for. The safety of patients should not be compromised by a comparability exercise, and patients should only be treated as medically warranted. In case there are no endpoints that are sufficiently sensitive to detect relevant differences, applicants need to implement additional measures to enable sufficient sensitivity of the overall clinical dataset obtained from the clinical study. For example, the study could be combined with a multiple dose study, or applicants could measure pharmacodynamic markers in addition to clinical endpoints in order to further establish comparability.

Clinical studies in special populations like the paediatric population or the elderly are normally not required since the overall objective of the development programme is to establish comparability, and therefore the selection of the primary patient population is driven by the need for homogeneity and sensitivity.

The inclusion of patients from non-European countries is generally possible if there are no intrinsic differences, but it may increase heterogeneity. Knowledge of efficacy and safety of the reference mAb in a particular region may be necessary in order to prospectively define an equivalence margin. Stratification and appropriate subgroup analyses are normally expected if patients from different global regions are included in order to demonstrate consistency with the overall effect. Diagnostic and treatment strategies should be comparable in order to prevent the influence of extrinsic factors.

5.3.1. Additional considerations for mAbs licensed in anticancer indications

Establishing similar clinical efficacy and safety of biosimilar and reference mAb may be particularly challenging in an anticancer setting: According to the "Guideline on the evaluation of anticancer medicinal products in man" (CHMP/EWP/205/95) the preferred endpoint to prove efficacy in cancer indications would be either progression free / disease free survival (PFS / DFS) or overall survival (OS). Such endpoints are important to establish patient benefit for a new anticancer drug, but may not be feasible or sensitive enough for establishing comparability of a biosimilar mAb to a reference mAb, since they may be influenced by various factors not attributable to differences between the biosimilar mAb and the reference mAb, but by factors like tumour burden, performance status, previous lines of

treatments, underlying clinical conditions, subsequent lines of treatment (for OS), etc. They may therefore not be suitable to establish similar efficacy of the biosimilar and the reference mAb.

Again, the focus of the comparability exercise is to demonstrate similar efficacy and safety compared to the reference medicinal product, not patient benefit *per se*, which has already been established by the reference medicinal product. In general the most sensitive patient population and clinical endpoint is preferred to be able to detect product-related differences, if present and, at the same time, to reduce patient and disease-related factors to a minimum in order to increase precision. A clinical trial in a homogeneous patient population with a clinical endpoint that measures activity as primary endpoint may be considered. An example may be Overall Response Rate (ORR, proportion of patients in whom a Complete Response (CR) or Partial Response (PR) was observed). It may also be worthwhile to explore ORR measured at a certain timepoint (i.e., ORR at x months) or percentage change in tumour mass from baseline or pathological Complete Response (pCR) in certain clinical settings. Applicants should engage in efforts for a standardized assessment and clear definitions of endpoints with patients evaluated at appropriate intervals. PFS and OS should be recorded, where feasible. It is acknowledged that data on survival may have to be interpreted with caution due to numerous factors influencing survival beyond the performance of the biosimilar mAb or the reference mAb. However, in case PFS is likely to be more sensitive than ORR as outcome measure, this is the preferred option even though this will prolong the clinical study.

Novel endpoints may be tested on an exploratory basis (e.g. time to response) and may add supportive evidence for biosimilarity.

5.4. Clinical Safety

Clinical safety is important throughout the clinical development programme and is captured during initial PK and/or PD evaluations and also as part of the pivotal clinical study establishing comparability. Care should be given to compare the type, severity and frequency of the adverse reactions between the biosimilar mAb and the reference mAb, particularly those described for the reference product. Where no homogeneous definition exists for safety parameters (e.g., measurement of cardiotoxicity) it is recommended to use the same definitions as that used for the reference mAb in its original development programme (if known) or the definitions used during post-authorisation follow-up. Comparison of pharmacologically mediated adverse reactions (e.g., cardiotoxicity), i.e., safety-related pharmacodynamic markers, could also be used as further supportive evidence for clinical comparability, and could be analysed in a similar way to that discussed for efficacy-related PD markers.

In cases where comparative and highly sensitive PD studies are suitable to provide the pivotal evidence for equivalence in clinical efficacy, applicants will have to provide sufficient reassurance of similar clinical safety, including immunogenicity. Actively controlled safety data should normally be collected pre-authorisation, depending on the mAb and the number of exposed patients, and duration of treatment. The duration of safety follow up pre authorisation should be justified.

It might be decided to collect part of the safety data, or also additional safety data, in the post-authorisation setting as described below. Rare events such as progressive multifocal leukoencephalopathy are unlikely to be detected in a pre-authorisation setting. Therefore, applicants need to propose pharmacovigilance and risk management activities for the post-authorisation phase at the time of the marketing authorisation application (see section 7). Usually, similar pharmacovigilance activities as those of the reference medicinal product would be required, rather than a direct comparison with the reference medicinal product, since comparative data will most likely be difficult to interpret due to their rarity of occurrence and consequent lack of precision for estimated differences.

Applicants should reflect upon how re-treatment of patients would be handled. Concepts should be presented at the time of marketing authorisation application on how to systematically measure safety of repeat exposure of patients, for example in oncological indications where patients undergo several treatment cycles. It is highly encouraged to extend the clinical study as a post-authorisation follow-up study to a full treatment cycle, where relevant and feasible.

As regards immunogenicity assessment, applicants should refer to existing CHMP guidance. Systematic and comparative evaluation and discussion of immunogenicity is important, due to clinical consequences like loss of efficacy and also likely resistance against further treatment with the reference mAb. It may be advisable not to include patients previously treated with the reference mAb where possible, or to pre-specify a subgroup analysis for patient previously treated (in order to explore if pre-treatment impacts immunogenicity), as previous treatment could have resulted in an anti-drug antibody response that could hamper interpretation of the safety data and thus also decrease sensitivity for detecting differences. Comparative assessment of unwanted immune responses against the biosimilar and the reference mAb are normally undertaken as part of the clinical study establishing similar clinical efficacy and safety, using the same validated assay(s) (see relevant CHMP guidelines on immunogenicity assessment). A population PK approach with sparse sampling and determination of drug concentration together with anti-drug antibody detection is acceptable. However, for some mAbs, antibodies can be better detected in healthy volunteers, who develop a strong immune response after a single dose within a few days. The dose of mAb administered is also an important factor to consider when investigating immunogenicity: some mAbs inhibit antibody formation when administered at high doses, and therefore studies conducted with low doses, if medically possible, are more sensitive to compare the immune response of the biosimilar and reference medicinal products.

Investigation of unwanted immunogenicity is especially important when a different expression system is employed for the biosimilar mAb compared to the reference mAb which might, for example, yield in relevant quality attributes that have not been detected in the reference product (e.g. new post-translational modification structure) that could result in a higher immunogenicity. This is particularly important if there is limited experience with this expression system in humans. It is recommended that such approaches are discussed in advance with regulatory authorities.

A higher immunogenicity as compared to the reference mAb may become an issue for the benefit/risk analysis and would question biosimilarity. However, also a lower immunogenicity for the biosimilar mAb is a possible scenario, which would not preclude biosimilarity. Here, the efficacy analysis of the entire patient population could suggest that the biosimilar is more efficacious (since fewer patients developed an immune response and thus more patients may show a treatment effect with the biosimilar mAb). It is therefore recommended to pre-specify an additional exploratory subgroup analysis of efficacy and safety in those patients that did not mount an anti-drug antibody response during the clinical trial. This subgroup analysis could be helpful to establish that the efficacy of the biosimilar and the reference mAb are in principle similar if not impacted by an immune response.

Additional long-term immunogenicity and safety data might be required post-authorisation, e.g. in situations where the study duration for establishing similar clinical efficacy is rather short. The need for additional long-term immunogenicity and safety data should be discussed in the risk management plan and, if considered needed, applicants should go beyond routine pharmacovigilance and perform post-authorisation safety studies. As regards safety and immunogenicity across different indications licensed for the reference mAb and claimed by the biosimilar mAb, a post-authorisation concept for obtaining further indication-specific safety data may be needed as described in section 7.

6. Extrapolation of indications

Extrapolation of clinical efficacy and safety data to other indications of the reference mAb, not specifically studied during the clinical development of the biosimilar mAb, is possible based on the overall evidence of comparability provided from the comparability exercise and with adequate justification. If pivotal evidence for comparability is based on PD and for the claimed indications different mechanisms of action are relevant (or uncertainty exists), then applicants should provide relevant data to support extrapolation to all claimed clinical indications. Applicants should support such extrapolations with a comprehensive discussion of available literature including the involved antigen receptor(s) and mechanism(s) of action.

For example, if a reference mAb is licensed both as an immunomodulator and as an anticancer antibody, the scientific justification as regards extrapolation between the two (or more) indications is more challenging. The basis for such extrapolation forms an extensive quality and non-clinical database, including potency assay(s) and *in vitro* assays that cover the functionality of the molecule, supplemented by relevant clinical data as described further in this document. The possibility of extrapolating safety including immunogenicity data also requires careful consideration, and may have to involve more specific studies (see sections 5 and 7). For the mechanism of action, e.g. the depletion of immune cells, several mechanisms may play a role in the various clinical conditions. For example, ADCC appears to be more important in some indications than in others. To provide further evidence about the mechanism of action, it may also be helpful to perform a literature search to identify what is known, e.g. about potential signalling inhibition by the reference mAb that would not be covered by ADCC/CDC tests, in particular direct induction of apoptosis. This could provide more knowledge on potential read-outs that could be used to support comparability on a molecular level.

7. Pharmacovigilance

For the marketing authorisation procedure the applicant should present a risk management plan/pharmacovigilance plan in accordance with current EU legislation and pharmacovigilance guidelines. Risk minimisation activities in place for the reference medicinal product may have to be implemented as well into the Risk Management Plan of the biosimilar.

Further to safety considerations as discussed above, applicants should provide at the time of MAA a comprehensive concept how to further study safety in a post-authorisation setting including also the following aspects:

- Safety in indications licensed for the reference mAb that are claimed based on extrapolation of efficacy and safety data, including long term safety data unless otherwise justified.
- Occurrence of rare and particularly serious adverse events described and predicted, based on the pharmacology, for the reference mAb. The pharmacovigilance plan should be proportionate to identified and potential risks and should be informed by the safety specification for the reference mAb in addition to relevant knowledge regarding similar biological products as appropriate.
- Detection of novel safety signals, as for any other biological medicinal product.
- Activities to obtain additional immunogenicity data, if considered needed.

The concept is likely to have to exceed routine pharmacovigilance, and may have to involve more proactive pharmacovigilance activities, e.g. where possible, registries or large population based databases, in which data is captured in a standardised way to ensure accurate and consistent data capture/ review. In addition, participation in already existing registries is recommended and should be

presented as part of the Risk Management Plan. The adequacy of such proposals will have to be assessed in the context of the safety data at the time of approval, the overall data from the comparability exercise, and the known safety profile of the reference mAb. The need for additional risk minimisation activities should be clearly evaluated taking into account the requirements for the reference medicinal product.

For suspected adverse reactions relating to biological medicinal products, the definite identification of the concerned product with regard to its manufacturing is of particular importance. Therefore, all appropriate measures should be taken to identify clearly any biological medicinal product which is the subject of a suspected adverse reaction report, with due regard to the name of the medicinal product and the batch number.

Depending on the handling of biosimilars and reference medicinal products in clinical practice at national level, 'switching' and 'interchanging' of medicines that contain a given mAb might occur. Thus, applicants are recommended to follow further development in the field and consider these aspects as part of the risk management plan.