June 2011
EMA/CHMP/ICH/731268/1998
Committee for medicinal products for human use (CHMP)

ICH guideline S6 (R1) – preclinical safety evaluation of biotechnology-derived pharmaceuticals

Step 5

Part I (Parent guideline)	
Transmission to CHMP	November 1996
Release for consultation	November 1996
Deadline for comments	May 1997
Final approval by CHMP	September 1997
Date for coming into operation	March 1998
Part II (Addendum)	
Transmission to CHMP	November 2009
Release for consultation	November 2009
Deadline for comments	February 2010
Final approval by CHMP	July 2011
Date for coming into operation	December 2011

7 Westferry Circus • Canary Wharf • London E14 4HB • United Kingdom
Telephone +44 (0)20 7418 8400 **Facsimile** +44 (0)20 75 23 70 40
E-mail ich@ema.europa.eu **Website** www.ema.europa.eu

Preclinical safety evaluation of biotechnology-derived pharmaceuticals

Table of contents

Part I

1. Introduction

1.1. Background

Biotechnology-derived pharmaceuticals (biopharmaceuticals) were initially developed in the early 1980s. The first marketing authorisations were granted later in the decade. Several guidelines and points-to-consider documents have been issued by various regulatory agencies regarding safety assessment of these products. Review of such documents, which are available from regulatory authorities, may provide useful background in developing new biopharmaceuticals.

Considerable experience has now been gathered with submission of applications for biopharmaceuticals. Critical review of this experience has been the basis for development of this guidance that is intended to provide general principles for designing scientifically acceptable preclinical safety evaluation programs.

1.2. Objectives

Regulatory standards for biotechnology-derived pharmaceuticals have generally been comparable among the European Union, Japan and United States. All regions have adopted a flexible, case-by-case, science-based approach to preclinical safety evaluation needed to support clinical development and marketing authorisation. In this rapidly evolving scientific area, there is a need for common understanding and continuing dialogue among the regions.

The primary goals of preclinical safety evaluation are: 1) to identify an initial safe dose and subsequent dose escalation schemes in humans; 2) to identify potential target organs for toxicity and for the study of whether such toxicity is reversible; and 3) to identify safety parameters for clinical monitoring. Adherence to the principles presented in this document is intended to improve the quality and consistency of the preclinical safety data supporting the development of biopharmaceuticals.

1.3. Scope

This guidance is intended primarily to recommend a basic framework for the preclinical safety evaluation of biotechnology-derived pharmaceuticals. It applies to products derived from characterised cells through the use of a variety of expression systems including bacteria, yeast, insect, plant, and mammalian cells. The intended indications may include *in vivo* diagnostic, therapeutic, or prophylactic uses. The active substances include proteins and peptides, their derivatives and products of which they are components; they could be derived from cell cultures or produced using recombinant DNA technology including production by transgenic plants and animals. Examples include but are not limited to: cytokines, plasminogen activators, recombinant plasma factors, growth factors, fusion proteins, enzymes, receptors, hormones, and monoclonal antibodies.

The principles outlined in this guidance may also be applicable to recombinant DNA protein vaccines, chemically synthesised peptides, plasma derived products, endogenous proteins extracted from human tissue, and oligonucleotide drugs.

This document does not cover antibiotics, allergenic extracts, heparin, vitamins, cellular blood components, conventional bacterial or viral vaccines, DNA vaccines, or cellular and gene therapies.

2. Specification of the test material

Safety concerns may arise from the presence of impurities or contaminants. It is preferable to rely on purification processes to remove impurities and contaminants rather than to establish a preclinical testing program for their qualification. In all cases, the product should be sufficiently characterised to allow an appropriate design of preclinical safety studies.

There are potential risks associated with host cell contaminants derived from bacteria, yeast, insect, plants, and mammalian cells. The presence of cellular host contaminants can result in allergic reactions and other immunopathological effects. The adverse effects associated with nucleic acid contaminants are theoretical but include potential integration into the host genome. For products derived from insect, plant and mammalian cells, or transgenic plants and animals there may be an additional risk of viral infections.

In general, the product that is used in the definitive pharmacology and toxicology studies should be comparable to the product proposed for the initial clinical studies. However, it is appreciated that during the course of development programs, changes normally occur in the manufacturing process in order to improve product quality and yields. The potential impact of such changes for extrapolation of the animal findings to humans should be considered.

The comparability of the test material during a development program should be demonstrated when a new or modified manufacturing process or other significant changes in the product or formulation are made in an ongoing development program. Comparability can be evaluated on the basis of biochemical and biological characterisation (i.e., identity, purity, stability, and potency). In some cases additional studies may be needed (i.e., pharmacokinetics, pharmacodynamics and/or safety). The scientific rationale for the approach taken should be provided.

3. Preclinical safety testing

3.1. General principles

The objectives of the preclinical safety studies are to define pharmacological and toxicological effects not only prior to initiation of human studies but throughout clinical development. Both *in vitro* and *in vivo* studies can contribute to this characterisation. Biopharmaceuticals that are structurally and pharmacologically comparable to a product for which there is wide experience in clinical practice may need less extensive toxicity testing.

Preclinical safety testing should consider: 1) selection of the relevant animal species; 2) age; 3) physiological state; 4) the manner of delivery, including dose, route of administration, and treatment regimen; and 5) stability of the test material under the conditions of use.

Toxicity studies are expected to be performed in compliance with Good Laboratory Practice (GLP); however, it is recognised that some studies employing specialised test systems which are often needed for biopharmaceuticals, may not be able to comply fully with GLP. Areas of non-compliance should be identified and their significance evaluated relative to the overall safety assessment. In some cases, lack of full GLP compliance does not necessarily mean that the data from these studies cannot be used to support clinical trials and marketing authorisations.

Conventional approaches to toxicity testing of pharmaceuticals may not be appropriate for biopharmaceuticals due to the unique and diverse structural and biological properties of the latter that may include species specificity, immunogenicity, and unpredicted pleiotropic activities.

3.2. Biological activity / pharmacodynamics

Biological activity may be evaluated using *in vitro* assays to determine which effects of the product may be related to clinical activity. The use of cell lines and/or primary cell cultures can be useful to examine the direct effects on cellular phenotype and proliferation. Due to the species specificity of many biotechnology-derived pharmaceuticals, it is important to select relevant animal species for toxicity testing. *In vitro* cell lines derived from mammalian cells can be used to predict specific aspects of *in vivo* activity and to assess quantitatively the relative sensitivity of various species (including human) to the biopharmaceutical. Such studies may be designed to determine, for example, receptor occupancy, receptor affinity, and/or pharmacological effects, and to assist in the selection of an appropriate animal species for further *in vivo* pharmacology and toxicology studies. The combined results from *in vitro* and *in vivo* studies assist in the extrapolation of the findings to humans. *In vivo* studies to assess pharmacological activity, including defining mechanism(s) of action, are often used to support the rationale of the proposed use of the product in clinical studies.

For monoclonal antibodies, the immunological properties of the antibody should be described in detail, including its antigenic specificity, complement binding, and any unintentional reactivity and/or cytotoxicity towards human tissues distinct from the intended target. Such cross-reactivity studies should be carried out by appropriate immunohistochemical procedures using a range of human tissues.

3.3. Animal species / model selection

The biological activity together with species and/or tissue specificity of many biotechnology-derived pharmaceuticals often preclude standard toxicity testing designs in commonly used species (e.g., rats and dogs). Safety evaluation programs should include the use of relevant species. A relevant species is one in which the test material is pharmacologically active due to the expression of the receptor or an epitope (in the case of monoclonal antibodies). A variety of techniques (e.g., immunochemical or functional tests) can be used to identify a relevant species. Knowledge of receptor/epitope distribution can provide greater understanding of potential *in vivo* toxicity.

Relevant animal species for testing of monoclonal antibodies are those that express the desired epitope and demonstrate a similar tissue cross-reactivity profile as for human tissues. This would optimise the ability to evaluate toxicity arising from the binding to the epitope and any unintentional tissue cross-reactivity. An animal species which does not express the desired epitope may still be of some relevance for assessing toxicity if comparable unintentional tissue cross-reactivity to humans is demonstrated.

Safety evaluation programs should normally include two relevant species. However, in certain justified cases one relevant species may suffice (e.g., when only one relevant species can be identified or where the biological activity of the biopharmaceutical is well understood). In addition even where two species may be necessary to characterise toxicity in short term studies, it may be possible to justify the use of only one species for subsequent long term toxicity studies (e.g., if the toxicity profile in the two species is comparable in the short term).

Toxicity studies in non-relevant species may be misleading and are discouraged. When no relevant species exists, the use of relevant transgenic animals expressing the human receptor or the use of homologous proteins should be considered. The information gained from use of a transgenic animal model expressing the human receptor is optimised when the interaction of the product and the humanised receptor has similar physiological consequences to those expected in humans. While useful information may also be gained from the use of homologous proteins, it should be noted that the production process, range of impurities/contaminants, pharmacokinetics, and exact

pharmacological mechanism(s) may differ between the homologous form and the product intended for clinical use. Where it is not possible to use transgenic animal models or homologous proteins, it may still be prudent to assess some aspects of potential toxicity in a limited toxicity evaluation in a single species, e.g., a repeated dose toxicity study of ≤ 14 days duration that includes an evaluation of important functional endpoints (e.g., cardiovascular and respiratory).

In recent years, there has been much progress in the development of animal models that are thought to be similar to the human disease. These animal models include induced and spontaneous models of disease, gene knockout(s), and transgenic animals. These models may provide further insight, not only in determining the pharmacological action of the product, pharmacokinetics, and dosimetry, but may also be useful in the determination of safety (e.g., evaluation of undesirable promotion of disease progression). In certain cases, studies performed in animal models of disease may be used as an acceptable alternative to toxicity studies in normal animals (Note 1). The scientific justification for the use of these animal models of disease to support safety should be provided.

3.4. Number/gender of animals

The number of animals used per dose has a direct bearing on the ability to detect toxicity. A small sample size may lead to failure to observe toxic events due to observed frequency alone regardless of severity. The limitations that are imposed by sample size, as often is the case for non-human primate studies, may be in part compensated by increasing the frequency and duration of monitoring. Both genders should generally be used or justification given for specific omissions.

3.5. Administration/dose selection

The route and frequency of administration should be as close as possible to that proposed for clinical use. Consideration should be given to pharmacokinetics and bioavailability of the product in the species being used, and the volume which can be safely and humanely administered to the test animals. For example, the frequency of administration in laboratory animals may be increased compared to the proposed schedule for the human clinical studies in order to compensate for faster clearance rates or low solubility of the active ingredient. In these cases, the level of exposure of the test animal relative to the clinical exposure should be defined. Consideration should also be given to the effects of volume, concentration, formulation, and site of administration. The use of routes of administration other than those used clinically may be acceptable if the route must be modified due to limited bioavailability, limitations due to the route of administration, or to size/physiology of the animal species.

Dosage levels should be selected to provide information on a dose-response relationship, including a toxic dose and a no observed adverse effect level (NOAEL). For some classes of products with little to no toxicity it may not be possible to define a specific maximum dose. In these cases, a scientific justification of the rationale for the dose selection and projected multiples of human exposure should be provided. To justify high dose selection, consideration should be given to the expected pharmacological/physiological effects, availability of suitable test material, and the intended clinical use. Where a product has a lower affinity to or potency in the cells of the selected species than in human cells, testing of higher doses may be important. The multiples of the human dose that are needed to determine adequate safety margins may vary with each class of biotechnology-derived pharmaceutical and its clinical indication(s).

3.6. Immunogenicity

Many biotechnology-derived pharmaceuticals intended for human are immunogenic in animals. Therefore, measurement of antibodies associated with administration of these types of products should be performed when conducting repeated dose toxicity studies in order to aid in the interpretation of these studies. Antibody responses should be characterised (e.g., titer, number of responding animals, neutralising or non-neutralising), and their appearance should be correlated with any pharmacological and/or toxicological changes. Specifically, the effects of antibody formation on pharmacokinetic/pharmacodynamic parameters, incidence and/or severity of adverse effects, complement activation, or the emergence of new toxic effects should be considered when interpreting the data. Attention should also be paid to the evaluation of possible pathological changes related to immune complex formation and deposition.

The detection of antibodies should not be the sole criterion for the early termination of a preclinical safety study or modification in the duration of the study design unless the immune response neutralises the pharmacological and/or toxicological effects of the biopharmaceutical in a large proportion of the animals. In most cases, the immune response to biopharmaceuticals is variable, like that observed in humans. If the interpretation of the data from the safety study is not compromised by these issues, then no special significance should be ascribed to the antibody response.

The induction of antibody formation in animals is not predictive of a potential for antibody formation in humans. Humans may develop serum antibodies against humanised proteins, and frequently the therapeutic response persists in their presence. The occurrence of severe anaphylactic responses to recombinant proteins is rare in humans. In this regard, the results of guinea pig anaphylaxis tests, which are generally positive for protein products, are not predictive for reactions in humans; therefore, such studies are considered of little value for the routine evaluation of these types of products.

4. Specific considerations

4.1. Safety pharmacology

It is important to investigate the potential for undesirable pharmacological activity in appropriate animal models and, where necessary, to incorporate particular monitoring for these activities in the toxicity studies and/or clinical studies. Safety pharmacology studies measure functional indices of potential toxicity. These functional indices may be investigated in separate studies or incorporated in the design of toxicity studies. The aim of the safety pharmacology studies should be to reveal any functional effects on the major physiological systems (e.g., cardiovascular, respiratory, renal, and central nervous systems). Investigations may also include the use of isolated organs or other test systems not involving intact animals. All of these studies may allow for a mechanistically-based explanation of specific organ toxicities, which should be considered carefully with respect to human use and indication(s).

4.2. Exposure assessment

4.2.1. Pharmacokinetics and toxicokinetics

It is difficult to establish uniform guidelines for pharmacokinetic studies for biotechnology-derived pharmaceuticals. Single and multiple dose pharmacokinetics, toxicokinetics, and tissue distribution studies in relevant species are useful; however, routine studies that attempt to assess mass

balance are not useful. Differences in pharmacokinetics among animal species may have a significant impact on the predictiveness of animal studies or on the assessment of dose response relationships in toxicity studies. Alterations in the pharmacokinetic profile due to immune-mediated clearance mechanisms may affect the kinetic profiles and the interpretation of the toxicity data. For some products there may also be inherent, significant delays in the expression of pharmacodynamic effects relative to the pharmacokinetic profile (e.g., cytokines) or there may be prolonged expression of pharmacodynamic effects relative to plasma levels.

Pharmacokinetic studies should, whenever possible, utilise preparations that are representative of that intended for toxicity testing and clinical use, and employ a route of administration that is relevant to the anticipated clinical studies. Patterns of absorption may be influenced by formulation, concentration, site, and/or volume. Whenever possible, systemic exposure should be monitored during the toxicity studies.

When using radiolabeled proteins, it is important to show that the radiolabeled test material maintains activity and biological properties equivalent to that of the unlabeled material. Tissue concentrations of radioactivity and/or autoradiography data using radiolabeled proteins may be difficult to interpret due to rapid *in vivo* metabolism or unstable radiolabeled linkage. Care should be taken in the interpretation of studies using radioactive tracers incorporated into specific amino acids because of recycling of amino acids into non-drug related proteins/peptides.

Some information on absorption, disposition and clearance in relevant animal models should be available prior to clinical studies in order to predict margins of safety based upon exposure and dose.

4.2.2. Assays

The use of one or more assay methods should be addressed on a case-by-case basis and the scientific rationale should be provided. One validated method is usually considered sufficient. For example, quantitation of TCA-precipitable radioactivity following administration of a radiolabeled protein may provide adequate information, but a specific assay for the analyte is preferred. Ideally the assay methods should be the same for animals and humans. The possible influence of plasma binding proteins and/or antibodies in plasma/serum on the assay performance should be determined.

4.2.3. Metabolism

The expected consequence of metabolism of biotechnology-derived pharmaceuticals is the degradation to small peptides and individual amino acids. Therefore, the metabolic pathways are generally understood. Classical biotransformation studies as performed for pharmaceuticals are not needed.

Understanding the behaviour of the biopharmaceutical in the biologic matrix, (e.g., plasma, serum, cerebral spinal fluid) and the possible influence of binding proteins is important for understanding the pharmacodynamic effect.

4.3. Single dose toxicity studies

Single dose studies may generate useful data to describe the relationship of dose to systemic and/or local toxicity. These data can be used to select doses for repeated dose toxicity studies. Information on dose- response relationships may be gathered through the conduct of a single dose

toxicity study, as a component of pharmacology or animal model efficacy studies. The incorporation of safety pharmacology parameters in the design of these studies should be considered.

4.4. Repeated dose toxicity studies

For consideration of the selection of animal species for repeated dose studies see section 3.3. The route and dosing regimen (e.g., daily versus intermittent dosing) should reflect the intended clinical use or exposure. When feasible, these studies should include toxicokinetics.

A recovery period should generally be included in study designs to determine the reversal or potential worsening of pharmacological/toxicological effects, and/or potential delayed toxic effects. For biopharmaceuticals that induce prolonged pharmacological/toxicological effects, recovery group animals should be monitored until reversibility is demonstrated. The duration of repeated dose studies should be based on the intended duration of clinical exposure and disease indication. This duration of animal dosing has generally been 1-3 months for most biotechnology-derived pharmaceuticals. For biopharmaceuticals intended for short-term use (e.g., ≤ to 7 days) and for acute life-threatening diseases, repeated dose studies up to two weeks duration have been considered adequate to support clinical studies as well as marketing authorisation. For those biopharmaceuticals intended for chronic indications, studies of 6 months duration have generally been appropriate although in some cases shorter or longer durations have supported marketing authorisations. For biopharmaceuticals intended for chronic use, the duration of long term toxicity studies should be scientifically justified.

4.5. Immunotoxicity studies

One aspect of immunotoxicological evaluation includes assessment of potential immunogenicity (see section 3.6). Many biotechnology-derived pharmaceuticals are intended to stimulate or suppress the immune system and therefore may affect not only humoral but also cell-mediated immunity. Inflammatory reactions at the injection site may be indicative of a stimulatory response. It is important, however, to recognise that simple injection trauma and/or specific toxic effects caused by the formulation vehicle may also result in toxic changes at the injection site. In addition, the expression of surface antigens on target cells may be altered, which has implications for autoimmune potential. Immunotoxicological testing strategies may require screening studies followed by mechanistic studies to clarify such issues. Routine tiered testing approaches or standard testing batteries, however, are not recommended for biotechnology-derived pharmaceuticals.

4.6. Reproductive performance and developmental toxicity studies

The need for reproductive/developmental toxicity studies is dependent upon the product, clinical indication and intended patient population (Note 2). The specific study design and dosing schedule may be modified based on issues related to species specificity, immunogenicity, biological activity and/or a long elimination half-life. For example, concerns regarding potential developmental immunotoxicity, which may apply particularly to certain monoclonal antibodies with prolonged immunological effects, could be addressed in a study design modified to assess immune function of the neonate.

4.7. Genotoxicity studies

The range and type of genotoxicity studies routinely conducted for pharmaceuticals are not applicable to biotechnology-derived pharmaceuticals and therefore are not needed. Moreover, the administration of large quantities of peptides/proteins may yield uninterpretable results. It is not

expected that these substances would interact directly with DNA or other chromosomal material (Note 3).

Studies in available and relevant systems, including newly developed systems, should be performed in those cases where there is cause for concern about the product (e.g., because of the presence of an organic linker molecule in a conjugated protein product). The use of standard genotoxicity studies for assessing the genotoxic potential of process contaminants is not considered appropriate. If performed for this purpose, however, the rationale should be provided.

4.8. Carcinogenicity studies

Standard carcinogenicity bioassays are generally inappropriate for biotechnology-derived pharmaceuticals. However, product-specific assessment of carcinogenic potential may still be needed depending upon duration of clinical dosing, patient population and/or biological activity of the product (e.g., growth factors, immunosuppressive agents, etc.) When there is a concern about carcinogenic potential a variety of approaches may be considered to evaluate risk.

Products that may have the potential to support or induce proliferation of transformed cells and clonal expansion possibly leading to neoplasia should be evaluated with respect to receptor expression in various malignant and normal human cells that are potentially relevant to the patient population under study. The ability of the product to stimulate growth of normal or malignant cells expressing the receptor should be determined. When *in vitro* data give cause for concern about carcinogenic potential, further studies in relevant animal models may be needed. Incorporation of sensitive indices of cellular proliferation in long term repeated dose toxicity studies may provide useful information.

In those cases where the product is biologically active and non-immunogenic in rodents and other studies have not provided sufficient information to allow an assessment of carcinogenic potential then the utility of a single rodent species should be considered. Careful consideration should be given to the selection of doses. The use of a combination of pharmacokinetic and pharmacodynamic endpoints with consideration of comparative receptor characteristics and intended human exposures represents the most scientifically based approach for defining the appropriate doses. The rationale for the selection of doses should be provided.

4.9. Local tolerance studies

Local tolerance should be evaluated. The formulation intended for marketing should be tested; however, in certain justified cases, the testing of representative formulations may be acceptable. In some cases, the potential adverse effects of the product can be evaluated in single or repeated dose toxicity studies thus obviating the need for separate local tolerance studies.

Notes

Note 1

Animal models of disease may be useful in defining toxicity endpoints, selection of clinical indications, and determination of appropriate formulations, route of administration, and treatment regimen. It should be noted that with these models of disease there is often a paucity of historical data for use as a reference when evaluating study results. Therefore, the collection of concurrent control and baseline data is critical to optimise study design.

Note 2

There may be extensive public information available regarding potential reproductive and/or developmental effects of a particular class of compounds (e.g., interferons) where the only relevant species is the non-human primate. In such cases, mechanistic studies indicating that similar effects are likely to be caused by a new but related molecule may obviate the need for formal reproductive/developmental toxicity studies. In each case, the scientific basis for assessing the potential for possible effects on reproduction/development should be provided.

Note 3

With some biopharmaceuticals there is a potential concern about accumulation of spontaneously mutated cells (e.g., via facilitating a selective advantage of proliferation) leading to carcinogenicity. The standard battery of genotoxicity tests is not designed to detect these conditions. Alternative *in vitro* or *in vivo* models to address such concerns may have to be developed and evaluated.

PART II: Addendum

Preamble:

This addendum should be read in close conjunction with the original ICH S6 Guideline. In general the addendum is complementary to the guideline, and where the addendum differs from the original guideline, the guidance in the addendum prevails.

1. Introduction

1.1. Purpose of the addendum

The purpose of the addendum is to complement and provide clarification on and update the following topics discussed in the original ICH S6 Guideline: species selection, study design, immunogenicity, reproductive and developmental toxicity and assessment of carcinogenic potential. Scientific advances and experience gained since publication of the original ICH S6 Guideline call for this addendum. This harmonised addendum will help to define the current recommendations and reduce the likelihood that substantial differences will exist among regions.

This guidance should facilitate the timely conduct of clinical trials, reduce the use of animals in accordance with the 3Rs (reduce/refine/replace) principles and reduce the use of other drug development resources. Although not discussed in this guidance, consideration should be given to the use of appropriate in vitro alternative methods for safety evaluation. These methods, if accepted by all ICH regulatory authorities, can be used to replace current standard methods.

This guidance promotes safe and ethical development and availability of new pharmaceuticals.

1.2. Background

The recommendations of this addendum further harmonise the nonclinical safety studies to support the various stages of clinical development among the regions of European Union (EU), Japan, and the United States. The present addendum represents the consensus that exists regarding the safety evaluation of biotechnology-derived pharmaceuticals.

1.3. Scope of the guideline

This addendum does not alter the scope of the original ICH S6 Guideline. For biotechnology-derived products intended to be used in oncology the Guidance on Nonclinical Evaluation for Anticancer Pharmaceuticals (ICH S9 Guideline) should be consulted.

2. Species selection

2.1. General principles

A number of factors should be taken into account when determining species relevancy. Comparisons of target sequence homology between species can be an appropriate starting point, followed by in vitro assays to make qualitative and quantitative cross-species comparisons of relative target binding affinities and receptor/ligand occupancy and kinetics.

Assessments of functional activity are also recommended. Functional activity can be demonstrated in species-specific cell-based systems and/or in vivo pharmacology or toxicology studies. Modulation

of a known biologic response or of a pharmacodynamic (PD) marker can provide evidence for functional activity to support species relevance.

Consideration of species differences in target binding and functional activity in the context of the intended dosing regime should provide confidence that a model is capable of demonstrating potentially adverse consequences of target modulation. When the target is expressed at very low levels in typical healthy preclinical species (e.g., inflammatory cytokines or tumour antigens), binding affinity and activity in cell-based systems can be sufficient to guide species selection.

Assessment of tissue cross reactivity (see Note 1) in animal tissues is of limited value for species selection. However, in specific cases (i.e. where the approaches described above cannot be used to demonstrate a pharmacologically relevant species) tissue cross-reactivity (TCR) studies can be used to guide selection of toxicology species by comparison of tissue binding profiles in human and those animal tissues where target binding is expected.

As described in ICH S6 Guideline, when no relevant species can be identified because the biopharmaceutical does not interact with the orthologous target in any species, use of homologous molecules or transgenic models can be considered.

For monoclonal antibodies and other related antibody products directed at foreign targets (i.e., bacterial, viral targets etc.), a short-term safety study (see ICH S6 Guideline) in one species (choice of species to be justified by the sponsor) can be considered; no additional toxicity studies, including reproductive toxicity studies, are appropriate. Alternatively, when animal models of disease are used to evaluate proof of principle, a safety assessment can be included to provide information on potential target-associated safety aspects. Where this is not feasible, appropriate risk mitigation strategies should be adopted for clinical trials.

Species selection for an antibody-drug/toxin conjugate (ADC) incorporating a novel toxin/toxicant should follow the same general principles as an unconjugated antibody (see above). (See Note 2)

2.2. One or two species

If there are two pharmacologically relevant species for the clinical candidate (one rodent and one non-rodent), then both species should be used for short-term (up to 1 month duration) general toxicology studies. If the toxicological findings from these studies are similar or the findings are understood from the mechanism of action of the product, then longer-term general toxicity studies in one species are usually considered sufficient. The rodent species should be considered unless there is a scientific rationale for using non-rodents. Studies in two non-rodent species are not appropriate.

The use of one species for all general toxicity studies is justified when the clinical candidate is pharmacologically active in only one species. Studies in a second species with a homologous product are not considered to add further value for risk assessment and are not recommended.

2.3. Use of homologous proteins

Use of homologous proteins is one of the alternative approaches described under ICH S6 Guideline Section 3.3. Studies with homologous proteins can be used for hazard detection and understanding the potential for adverse effects due to exaggerated pharmacology, but are generally not useful for quantitative risk assessment. Therefore, for the purposes of hazard identification it can be possible to conduct safety evaluation studies using a control group and one treatment group provided there is a scientific justification for the study design and dose selected (e.g., maximum pharmacological dose).

3. Study design

3.1. Dose selection and application of PK/PD principles

The toxicity of most biopharmaceuticals is related to their targeted mechanism of action; therefore, relatively high doses can elicit adverse effects which are apparent as exaggerated pharmacology.

A rationale should be provided for dose selection taking into account the characteristics of the dose-response relationship. Pharmacokinetic-pharmacodynamic (PK-PD) approaches (e.g., simple exposure-response relationships or more complex modeling and simulation approaches) can assist in high dose selection by identifying 1) a dose which provides the maximum intended pharmacological effect in the preclinical species and 2) a dose which provides an approximately 10-fold exposure multiple over the maximum exposure to be achieved in the clinic. The higher of these two doses should be chosen for the high dose group in preclinical toxicity studies unless there is a justification for using a lower dose (e.g., maximum feasible dose).

Where in vivo/ex vivo PD endpoints are not available, the high dose selection can be based on PK data and available in vitro binding and/or pharmacology data. Corrections for differences in target binding and in vitro pharmacological activity between the nonclinical species and humans should be taken into account to adjust the exposure margin over the highest anticipated clinical exposure. For example, a large relative difference in binding affinity and/or in vitro potency might suggest that testing higher doses in the nonclinical studies is appropriate. In the event that toxicity cannot be demonstrated at the doses selected using this approach, then additional toxicity studies at higher multiples of human dosing are unlikely to provide additional useful information.

3.2. Duration of studies

For chronic use products, repeat dose toxicity studies of 6 months duration in rodents or non-rodents are considered sufficient, providing the high dose is selected in accordance with the principles above in section 3.1. Studies of longer duration have not generally provided useful information that changed the clinical course of development.

For chronic use of biopharmaceutical products developed for patients with advanced cancer, the principles for duration of toxicology studies are outlined in ICH S9 Guideline.

3.3. Recovery

Recovery from pharmacological and toxicological effects with potential adverse clinical impact should be understood when they occur at clinically relevant exposures. This information can be obtained by an understanding that the particular effect observed is generally reversible / non-reversible or by including a non-dosing period in at least one study, at at least one dose level, to be justified by the sponsor. The purpose of the non-dosing period is to examine reversibility of these effects, not to assess delayed toxicity. The demonstration of complete recovery is not considered essential. The addition of a recovery period just to assess potential for immunogenicity is not required.

3.4. Exploratory clinical trials

The flexible approaches to support exploratory clinical trials as outlined in ICH M3(R2) Guideline can be applicable to biopharmaceuticals. It is recommended that these approaches be discussed and agreed upon with the appropriate regulatory authority.

4. Immunogenicity

Immunogenicity assessments are conducted to assist in the interpretation of the study results and design of subsequent studies. Such analyses in nonclinical animal studies are not relevant in terms of predicting potential immunogenicity of human or humanized proteins in humans.

Measurement of anti-drug antibodies (ADA) in nonclinical studies should be evaluated when there is 1) evidence of altered PD activity; 2) unexpected changes in exposure in the absence of a PD marker; or 3) evidence of immune-mediated reactions (immune complex disease, vasculitis, anaphylaxis, etc.). Since, it is difficult to predict whether such analysis will be called for prior to completion of the in-life phase of the study, it is often useful to obtain appropriate samples during the course of the study, which can subsequently be analyzed when warranted to aid in interpretation of the study results. When ADAs are detected, their impact on the interpretation of the study results should be assessed (see also section 3.6, paragraph 2 in ICH S6 Guideline for further guidance on the impact of immunogenicity).

Characterization of neutralizing potential is warranted when ADAs are detected and there is no PD marker to demonstrate sustained activity in the in vivo toxicology studies. Neutralizing antibody activity can be assessed indirectly with ex-vivo bioactivity assay or an appropriate combination of assay formats for PK-PD, or directly in a specific neutralizing antibody assay.

5. Reproductive and developmental toxicity

5.1. General comments

Reproductive toxicity studies should be conducted in accordance with the principles outlined in ICH S5(R2) Guideline; however, the specific study design and dosing schedule can be modified based on an understanding of species specificity, the nature of the product and mechanism of action, immunogenicity and/or pharmacokinetic behaviour and embryo-fetal exposure.

An assessment of reproductive toxicity with the clinical candidate in a relevant species is generally preferred. The evaluation of toxicity to reproduction should be conducted only in pharmacologically relevant species. When the clinical candidate is pharmacologically active in rodents and rabbits, both species should be used for embryo-fetal development (EFD) studies, unless embryo-fetal lethality or teratogenicity has been identified in one species.

Developmental toxicity studies should only be conducted in non-human primates (NHPs) when they are the only relevant species.

When the clinical candidate is pharmacologically active only in NHPs, there is still a preference to test the clinical candidate. However an alternative model can be used in place of NHPs if appropriate scientific justification is provided

When no relevant animal species exists for testing the clinical candidate, the use of transgenic mice expressing the human target or homologous protein in a species expressing an ortholog of the human target can be considered, assuming that sufficient background knowledge exists for the model (e.g., historical background data) (see Note 1 of ICH S6 Guideline). For products that are directed at a foreign target such as bacteria and viruses, in general no reproductive toxicity studies would be expected (See Section 2.1).

When the weight of evidence (e.g., mechanism of action, phenotypic data from genetically modified animals, class effects) suggests that there will be an adverse effect on fertility or pregnancy

outcome, these data can provide adequate information to communicate risk to reproduction, and under appropriate circumstances additional nonclinical studies might not be warranted.

5.2. Fertility

For products where mice and rats are pharmacologically relevant species, an assessment of fertility can be made in one of these rodent species (see ICH S5 Guideline). ICH S5 Guideline study designs can be adapted for other species provided they are pharmacologically relevant; in addition, the design of the study should be amended as appropriate, for example to address the nature of the product and potential for immunogenicity.

It is recognized that mating studies are not practical for NHPs. However, when the NHP is the only relevant species, the potential for effects on male and female fertility can be assessed by evaluation of the reproductive tract (organ weights and histopathological evaluation) in repeat dose toxicity studies of at least 3 months duration using sexually mature NHPs. If there is a specific cause for concern based on pharmacological activity or previous findings, specialized assessments such as menstrual cyclicity, sperm count, sperm morphology/motility, and male or female reproductive hormone levels can be evaluated in a repeat dose toxicity study.

If there is a specific concern from the pharmacological activity about potential effects on conception / implantation and the NHP is the only relevant species, the concern should be addressed experimentally. A homologous product or transgenic model could be the only practical means to assess potential effects on conception or implantation when those are of specific concern. However, it is not recommended to produce a homologous product or transgenic model solely to conduct mating studies in rodents. In absence of nonclinical information, the risk to patients should be mitigated through clinical trial management procedures, informed consent and appropriate product labeling.

5.3. Embryo-fetal (EFD) and pre/post-natal development (PPND)

Potential differences in placental transfer of biopharmaceuticals should be considered in the design and interpretation of developmental toxicity studies – see Note 3.

For products pharmacologically active only in NHPs, several study designs can be considered based on intended clinical use and expected pharmacology. Separate EFD and/or PPND studies, or other study designs (justified by the sponsor) can be appropriate, particularly when there is some concern that the mechanism of action might lead to an adverse effect on embryo-fetal development or pregnancy loss. However, one well-designed study in NHPs which includes dosing from day 20 of gestation to birth (enhanced PPND; ePPND) can be considered, rather than separate EFD and/or PPND studies.

For the single ePPND study design described above, no Caesarean section group is warranted, but assessment of pregnancy outcome at natural delivery should be performed. This study should also evaluate offspring viability, external malformations, skeletal effects (e.g., by X-ray) and, ultimately, visceral morphology at necropsy. Ultrasound is useful to track maintenance of pregnancy but is not appropriate for detecting malformations. These latter data are derived from post-partum observations. Because of confounding effects on maternal care of offspring, dosing of the mother post-partum is generally not recommended. Other endpoints in the offspring can also be evaluated if relevant for the pharmacological activity (see Note 3). The duration of the postnatal phase will be dependent on which additional endpoints are considered relevant based on mechanism of action (see Note 4).

Developmental toxicity studies in NHPs can only provide hazard identification. The number of animals per group should be sufficient to allow meaningful interpretation of the data (see Note 5).

The sponsor should justify the study design if other NHP species are used. The developmental toxicity studies in NHP as outlined above are just hazard identification studies; therefore it might be possible to conduct these studies using a control group and one dose group, provided there is a scientific justification for the dose level selected. An example of an appropriate scientific justification would be a monoclonal antibody which binds a soluble target with a clinical dosing regimen intended to saturate target binding. If such a saturation of target binding can be demonstrated in the animal species selected and there is an up to 10-fold exposure multiple over therapeutic drug levels, a single dose level and control group would provide adequate evidence of hazard to embryo-fetal development.

5.4. Timing of studies

If women of child-bearing potential are included in clinical trials prior to acquiring information on effects on embryo-fetal development, appropriate clinical risk management is appropriate, such as use of highly effective methods of contraception (see ICH M3(R2) Guideline).

For biopharmaceuticals pharmacologically active only in NHPs, where there are sufficient precautions to prevent pregnancy (see ICH M3(R2) Guideline), Section 11.3 paragraph 2), an EFD or ePPND study can be conducted during Phase III, and the report submitted at the time of marketing application. When a sponsor cannot take sufficient precaution to prevent pregnancy in clinical trials, either a complete report of an EFD study or an interim report of an ePPND study (see note 6) should be submitted before initiation of Phase III. Where the product is pharmacologically active only in NHPs and its mechanism of action raises serious concern for embryo-fetal development, the label should reflect the concern without warranting a developmental toxicity study in NHPs and therefore administration to women of child-bearing potential should be avoided.

If the rodent or rabbit is a relevant species, see ICH M3(R2) Guideline for timing of reproductive toxicity studies. ICH M3(R2) Guideline should also be followed for the timing of data on fertility for products where rodents are relevant species.

For oncology products which fall within the scope of ICH S9 Guideline, see that guidance for aspects relating to timing of study conduct.

6. Carcinogenicity

The need for a product-specific assessment of the carcinogenic potential for biopharmaceutical should be determined with regard to the intended clinical population and treatment duration (see ICH S1A Guideline). When an assessment is warranted, the sponsor should design a strategy to address the potential hazard.

This strategy could be based on a weight of evidence approach, including a review of relevant data from a variety of sources. The data sources can include published data (e.g., information from transgenic, knock-out or animal disease models, human genetic diseases), information on class effects, detailed information on target biology and mechanism of action, in vitro data, data from chronic toxicity studies and clinical data. In some cases, the available information can be sufficient to address carcinogenic potential and inform clinical risk without additional nonclinical studies.

The mechanism of action of some biopharmaceuticals might raise concern regarding potential for carcinogenicity (e.g., immunosuppressives and growth factors). If the weight of evidence (see above) supports the concern regarding carcinogenic potential, rodent bioassays are not warranted.

In this case potential hazard can be best addressed by product labeling and risk management practices. However, when the weight of evidence is unclear, the sponsor can propose additional studies that could mitigate the mechanism-based concern (see ICH S6 Guideline, section 4.8).

For products where there is insufficient knowledge about specific product characteristics and mode of action in relation to carcinogenic potential, a more extensive assessment might be appropriate (e.g., understanding of target biology related to potential carcinogenic concern, inclusion of additional endpoints in toxicity studies).

If the weight of evidence from this more extensive assessment does not suggest carcinogenic potential, no additional nonclinical testing is recommended. Alternatively, if the weight of evidence suggests a concern about carcinogenic potential, then the sponsor can propose additional nonclinical studies that could mitigate the concern, or the label should reflect the concern.

The product-specific assessment of carcinogenic potential is used to communicate risk and provide input to the risk management plan along with labeling proposals, clinical monitoring, post-marketing surveillance, or a combination of these approaches.

Rodent bioassays (or short-term carcinogenicity studies) with homologous products are generally of limited value to assess carcinogenic potential of the clinical candidate.

Alternative approaches can be considered as new strategies / assays are developed.

7. Endnotes

Note 1

Tissue cross-reactivity (TCR) studies are in vitro tissue-binding assays employing immunohistochemical (IHC) techniques conducted to characterize binding of monoclonal antibodies and related antibody-like products to antigenic determinants in tissues. Other technologies can be employed in place of IHC techniques to demonstrate target /binding site distribution.

A TCR study with a panel of human tissues is a recommended component of the safety assessment package supporting initial clinical dosing of these products. However, in some cases the clinical candidate is not a good IHC reagent and a TCR study might not be technically feasible.

TCR studies can provide useful information to supplement knowledge of target distribution and can provide information on potential unexpected binding. Tissue binding per se does not indicate biological activity in vivo. In addition, binding to areas not typically accessible to the antibody in vivo i.e. cytoplasm is generally not relevant. Findings should be evaluated and interpreted in the context of the overall pharmacology and safety assessment data package.

When there is unexpected binding in human tissues an evaluation of selected animal tissues can provide supplemental information regarding potential correlations or lack thereof with preclinical toxicity. TCR using a full panel of animal tissues is not recommended.

Since a bi-specific antibody product will be evaluated in a TCR study using a panel of human tissues, there is no need to study the individual binding components.

Evaluating the tissue binding of homologous products does not provide additional value when TCR studies have been conducted with the clinical candidate in a human tissue panel, and is not recommended.

TCR studies cannot detect subtle changes in critical quality attributes. Therefore TCR studies are not recommended for assessing comparability of the test article as a result of process changes over the course of a development program,

Note 2

If two species have been used to assess the safety of the ADC, an additional short-term study or arm in a short-term study should be conducted in at least one species with the unconjugated toxin. In these cases a rodent is preferred unless the toxin is not active in the rodent. If only one pharmacologically relevant species is available, then the ADC should be tested in this species. A novel toxicant calls for an approach to species selection similar to that used for a new chemical entity on a case-by case approach (e.g., for anticancer products in accordance with ICH-S9 Guideline). For toxins or toxicants which are not novel and for which there is a sufficient body of scientific information available, separate evaluation of the unconjugated toxin is not warranted. Data should be provided to compare the metabolic stability of the ADC in animals with human.

Note 3

The species-specific profile of embryo-fetal exposure during gestation should be considered in interpreting studies. High molecular weight proteins (>5,000 D) do not cross the placenta by simple diffusion. For monoclonal antibodies with molecular weight as high as 150,000 D, there exists a specific transport mechanism, the neonatal Fc receptor (FcRn) which determines fetal exposure and varies across species.

In the NHP and human, IgG placental transfer is low in the period of organogenesis and begins to increase in early second trimester, reaching highest levels late in the third trimester (Pentsuk and van der Laan, 2009). Therefore, standard embryo-fetal studies in NHPs, which are dosed from early pregnancy up to Gestation Day 50, might not be of value to assess direct embryo-fetal effects in the period of organogenesis, although effects on embryo-fetal development as an indirect result of maternal effects can be evaluated. Furthermore, maternal dosing in NHP after delivery is generally without relevance as IgG is only excreted in the milk initially (i.e., in the colostrum), and not later during the lactation and nursing phase.

Rodents differ from the NHPs and humans, as IgG crosses the yolk sac in rodents by FcRn transport mechanisms and exposure can occur relatively earlier in gestation than with NHP's and humans. In addition, delivery of rodents occurs at a stage of development when the pups are not as mature as the NHP or the human neonate. Therefore, rat/mouse dams should be dosed during lactation in order to expose pups via the milk up to at least day 9 of lactation when the offspring are at an equivalent stage of development as human neonates.

Note 4

The minimum duration of postnatal follow-up should be one month to cover early functional testing (e.g., growth and behaviour).

In general, if there is evidence for adverse effects on the immune system (or immune function) in the general toxicology studies, immune function testing in the offspring during the post-partum phase of the enhanced Pre/Post-Natal Development (ePPND) study is warranted. When appropriate, immunophenotyping can be obtained as early as post-natal day 28. The duration of postnatal follow-up for assessment of immune function can be 3- 6 months depending on the functional test used.

Neurobehavioural assessment can be limited to clinical behavioural observations. Instrumental learning calls for a training period, which would result in a postnatal duration of at least 9 months and is not recommended.

Note 5

A detailed discussion of the approach to determine group sizes in cynomolgus monkey ePPND studies can be found in Jarvis et al, 2010. Group sizes in ePPND studies should yield a sufficient number of infants (6-8 per group at postnatal day 7) in order to assess post-natal development and provide the opportunity for specialist evaluation if necessary (e.g., immune system).

Most ePPND studies accrue pregnant animals over weeks and months. Consideration should be given to terminating further accrual of pregnant animals into the study, and adapting the study design (e.g., by Caesarean section) when prenatal losses in a test item group indicate a treatment-related effect.

Reuse of vehicle-control treated maternal animals is encouraged.

If there is some cause for concern that the mechanism of action might lead to an effect on EFD or pregnancy loss, studies can be conducted in a limited number of animals in order to confirm the hazard.

Note 6

Endpoints to be included in an interim report of an ePPND study in NHPs:

Dam data: survival, clinical observations, bodyweight, gestational exposure data (if available), any specific PD endpoints;

Pregnancy data: number of pregnant animals started on study, pregnancy status at both the end of organogenesis (gestation day (GD) 50) and at GD100, occurrence of abortions and timing of abortions.

There is no need for ultrasound determinations of fetal size in the interim report; these are not considered essential since actual birth weight will be available;

Pregnancy outcome data: number of live births / still births, infant birth weight, infant survival and bodyweight at day 7 post-partum, qualitative external morphological assessment (i.e., confirming appearance is within normal limits), infant exposure data (if available), any specific PD endpoints in the infant if appropriate.

8. References

1. ICH S6 Guideline: Safety Studies for Biotechnological Products; July 1997.
2. ICH S5(R2) Guideline: Detection of Toxicity to Reproduction for Medicinal Products and Toxicity to Male Fertility; June 1993.
3. ICH S1A Guideline: Guideline on the Need for Carcinogenicity Studies for Pharmaceuticals; November 1995.
4. ICH M3(R2) Guideline: Guidance on Nonclinical Safety Studies for the Conduct of Human Clinical Trials and Marketing Authorisation for Pharmaceuticals; June 2009.
5. ICH S9 Guideline: Nonclinical Evaluation for Anticancer Pharmaceuticals; November 2008.

6. Pentsuk N, van der Laan JW. An interspecies comparison of placental antibody transfer: new insights into developmental toxicity testing of monoclonal antibodies. Birth defects research (Part B) 2009; 86: 328-344.

7. Jarvis P, Srivastav S, Vogelwedde E, Stewart J, Mitchard T, Weinbauer G. The Cynomolgous Monkey as a model for Developmental Toxicity Studies: Variability of Pregnancy losses, Statistical power estimates, and Group Size considerations. Birth Defects Research (part B), 2010, 89: 175-187.