EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

6 October 2025
EMA/CHMP/55697/2025
European Medicines Agency

Reflection paper on non-human primates in safety testing of human medicinal products and opportunities for 3Rs implementation

Draft agreed by NcWP and 3RsWP	15 July 2025
Adopted by CHMP for release for consultation	6 October 2025
Start of public consultation	23 October 2025
End of consultation (deadline for comments)	31 January 2026

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

Keywords	**Non-human primates, non-clinical, safety, in vivo, drug regulation, regulatory science, toxicology, animal testing, weight of evidence approach, 3Rs, NAMs**

Official address Domenico Scarlattilaan 6 ● 1083 HS Amsterdam ● The Netherlands
Address for visits and deliveries Refer to www.ema.europa.eu/how-to-find-us
Send us a question Go to www.ema.europa.eu/contact **Telephone** +31 (0)88 781 6000 An agency of the European Union

Reflection paper on non-human primates in safety testing of human medicinal products and opportunities for 3Rs implementation

List of Abbreviations

3Rs	Replacement, Reduction, Refinement
3RsWP	3Rs Working Party
ADA	Anti-Drug-Antibody
ADCs	Antibody-Drug Conjugates
AI	Artificial Intelligence
ATMPs	Advanced Therapy Medicinal Products
CCG	Concurrent Control Group
DART	Developmental and Reproductive Toxicity
EFD	Embryo-Foetal Development
EMA	European Medicines Agency
ePPND	Enhanced Pre- and Postnatal Development
EU	European Union
GTMPs	Gene Therapy Medicinal Products
IMP	Investigational Medicinal Product
ITF	Innovative Task Force
MAA	Marketing Authorization Application
mAbs	Monoclonal Anti-bodies
MEFL	Malformations and Embryo-Foetal Lethality
MPS	Micro Physiological Systems
NAMs	New Approach Methodologies
NcWP	Non-Clinical Working Party
NHP	Non-Human Primates
OoC	Organ-on-Chip
RDT	Repeat Dose Toxicity
SDLT	Severely Debilitating and Life Threatening
US-FDA	United States Food and Drug Administration
VCG	Virtual Control Group
VDS	Voluntary Data Submission
WOCBP	Women of Childbearing Potential
WoE	Weight of Evidence

1. Executive Summary

This Reflection Paper aims to provide an overview of the scientific and regulatory considerations for non-human primate use in safety testing of human medicinal products. It highlights the existing flexibility within published guidelines to incorporate 3Rs approaches and describes novel alternative approaches which may become available in the future. Notwithstanding the detailed conditions outlined herein, some important examples include; use of rodent species (including transgenics) only to evaluate repeat dose toxicity, the waiving of long-term (6 month) studies to evaluate the safety risk associated with monoclonal antibodies, the use of alternative assays to predict malformations or embryo-foetal lethality in developmental and reproductive toxicity.

2. Introduction

Regulatory considerations for the safety testing of novel medicines are based on global and regional guidelines. In addition, in accordance with the provisions of Directive 2010/63/EU on protection of animals used for scientific purposes, the 3Rs principles (replacement, reduction and refinement) are applicable to regulatory safety testing of medicinal products. Specifically, the Directive states that non-human primates (NHP) **shall not be used in procedures** except in translation and applied research or regulatory testing aimed at the avoidance, prevention, diagnosis or treatment of debilitating or potentially life-threatening clinical conditions in human beings, and where there is scientific justification to the effect that the purpose of the procedure cannot be achieved by the use of species other than NHP.

The requirement for safety studies in non-rodent species is driven by the need to generate data that provide sufficient evidence that the investigational medicinal product (IMP) has an acceptable safety profile. This may not be achievable using only rodent species. In general, non-rodent studies are recommended for repeated dose toxicity (RDT), safety pharmacology, and developmental and reproductive toxicity (DART) testing, and, when scientifically justified, to address other specific concerns (e.g., juvenile toxicity, immunotoxicity or mechanistic studies to elucidate potential human relevance of observed toxicity).

In accordance with Directive 2010/63/EU, the use of NHP as the default non-rodent species for non-clinical safety testing is increasingly challenged. In some cases, use of NHP may not be necessary to provide reliable hazard or risk identification for humans, while in other cases it might yield poorly translatable or scientifically unreliable results. Therefore, the use of NHPs in non-clinical safety testing should be a last resort based on a sound justification, only when alternative species or testing approaches are not available, and the number of NHPs used should be limited to the minimum necessary.

During the COVID-19 pandemic, the pre-existing shortage of NHP specimens was further exacerbated. Constraints on the availability of sexually mature monkeys have the potential to; delay access to innovative medicines for patients (specifically where NHP testing is required), increase illicit trading of NHP, result in rising costs of research and development, and hence of medicines, and raise public health issues (through zoonosis of illicit animals). The acute shortage during the pandemic prompted the United States Food and Drug Administration (US-FDA) to issue guidance for industry to highlight the already existing regulatory opportunities to minimize NHP use for medicine safety testing (1). In the European Union (EU), the use of NHP for regulatory safety studies has been critically monitored by regulatory authorities including European Medicines Agency (EMA) through scientific and regulatory activities, in line with the 3Rs principles and the objectives of Directive 2010/63/EU. Novel 3Rs approaches are accepted if it can be demonstrated that they can provide at least an equivalent level of information on safety (see also Section 8).

3. Scope

The scope of this paper is to reflect on the current use of NHPs for non-clinical safety assessment of human medicinal products and on current and future opportunities for reduction or avoidance of NHP use. In this regard, the paper provides scientific and regulatory considerations for NHP use, as well as highlighting the potential to leverage existing flexibility in current guidelines and future 3Rs opportunities.

3.1. Legal basis

This reflection paper should be read together with, and with reference to, the following regulations and guidance:

3.2. EU Regulations

Directive 2001/83/EC, on the Community code relating to medicinal products for human use.

Directive 2010/63/EU, on the protection of animals used for scientific purposes.

3.3. Global guidelines

ICH M3(R2) and Q&A Nonclinical Safety Studies for the Conduct of Human Clinical Trials and Marketing Authorization for Pharmaceuticals (EMA/CPMP/ICH/286/1995).

ICH S5(R3) Detection of Reproductive and Developmental Toxicity for Human Pharmaceuticals (EMA/CHMP/ICH/544278/1998).

ICH S6(R1) Preclinical Safety Evaluation of Biotechnology-Derived Pharmaceuticals (EMA/CHMP/ICH/731268/1998).

ICH S7A Safety Pharmacology Studies for Human Pharmaceuticals (CPMP/ICH/539/00).

ICH S7B/E14 Q&A The Non-clinical Evaluation of the Potential for Delayed Ventricular Repolarization (QT Interval Prolongation) by Human Pharmaceuticals (EMA/CHMP/ICH/415588/2020).

ICH S8 Note for guidance on immunotoxicity studies for human pharmaceuticals (CHMP/167235/2004).

ICH S9 and Q&A Non-Clinical Evaluation of Anticancer Pharmaceuticals (EMA/CHMP/ICH/646107/2008) (EMA/CHMP/ICH/453684/2016).

ICH guideline S11 on nonclinical safety testing in support of development of paediatric pharmaceuticals (EMA/CHMP/ICH/616110/2018).

ICH S12 Guideline on nonclinical biodistribution considerations for gene therapy products (EMA/CHMP/ICH/318372/2021).

WHO guideline on non-clinical evaluation of vaccines (WHO Technical Report Series No 927).

WHO guideline on the non-clinical evaluation of vaccine adjuvants and adjuvanted vaccines (Annex II, WHO Technical Report Series No. 987, 2014).

3.4. EMA guidance

Guideline on quality, non-clinical and clinical requirements for investigational advanced therapy medicinal products in clinical trials (EMA/CAT/22473/2025).

Guideline on the principles of regulatory acceptance of 3Rs (replacement, reduction, refinement) testing approaches (EMA/CHMP/CVMP/JEG-3Rs/450091/2012).

Draft reflection paper on the current regulatory testing requirements for medicinal products for human use and opportunities for implementation of the 3Rs - Revision 1 (EMA/CHMP/CVMP/3Rs/742466/2015 Rev.1).

Guideline on the risk-based approach according to annex I, part IV of Directive 2001/83/EC applied to advanced-therapy medicinal products (EMA/CAT/CPWP/686637/2011).

Guideline on the non-clinical investigation of the dependence potential of medicinal products (EMEA/CHMP/SWP/94227/2004).

4. General Considerations on NHP use and 3Rs opportunities

NHPs should be used for regulatory safety testing only as a last resort. It is acknowledged that some variability exists in relation to the interpretation of regulatory guidelines. Here, the potential to exploit existing flexibility in guidance to minimise NHP use where possible, is reflected on. Moreover, opportunities for further reduction of NHP use based on regulatory precedence or scientific literature are also considered.

4.1. Flexibility and Weight of Evidence Approaches

In general, regulatory guidelines provide flexibility as to the design of a non-clinical safety testing programme and/or specific studies therein. Deviations from the recommendations are possible based on a robust scientific justification and data (new or from published literature), where appropriate, to demonstrate that an alternative approach provides at least an equivalent level of information on safety. This is relevant for implementation of 3Rs testing approaches, including those aiming to replace, reduce or refine NHP use. Non-animal approaches and other available information on a drug candidate may also be used to compile a weight of evidence (WoE) to justify deviation from general regulatory guidance. WoE considerations for safety assessment can integrate for example, pharmacological properties, the role of the pharmacological target, drug specificity and pharmacokinetic (PK) data and available non-clinical or clinical data. Where relevant, additional factors can also be considered, such as clinical risk mitigation strategies or the feasibility of performing a study in the selected species. For pharmaceuticals with high selectivity and specificity, existing data from products with the same target should be considered, while read-across approaches can be considered for substances with similar structural characteristics.

Some WoE approaches (not exclusive to NHPs) have already been implemented in international guidelines including ICH S5 (R3), ICH S6(R1), ICH S8, ICH S9 and ICH S11. For example, as per ICH S5(R3), WoE approaches can be accepted in the context of demonstrating that a molecule will elicit malformations and embryo-foetal lethality (MEFL) in DART studies. Based on such WoE approaches, in vivo DART studies, including those using NHPs, can be avoided. In ICH S11, a standardised WoE model has been implemented to justify whether a juvenile animal toxicity study is needed. Outside of these internationally harmonised guidelines, published literature also supports the use of WoE approaches. Although less common, regulators have accepted cases where absence of human risk was demonstrated, based on arguments around specific patient populations, dosing (posology) or (lack of) exposure, to justify waiving in vivo DART studies (2). Other non-standard WoE approaches can also be considered on a case-by-case basis.

Taken together, 3Rs-based WoE strategies may result in the deferral or waiving of an NHP study, reductions in study duration, reduction in the number of study groups or animals used, the replacement of specific NHP studies by non-animal alternatives, or a combination of these. Advice on the acceptability of the proposed approach can be requested through formal or informal discussions with EMA (see also Section 8).

4.2. Non-animal approaches

NHPs are primarily used for RDT and DART testing and use of non-animal approaches within these contexts should therefore be considered when available. Suggestions for use of 3Rs testing approaches are provided in the Reflection Paper on the current regulatory testing requirements for medicinal products for human use and opportunities for implementation of the 3Rs (EMA/CHMP/CVMP/3Rs/742466/2015 Rev. 1). The regulatory acceptance of non-animal approaches, either stand-alone, combined in a testing battery or as part of a WoE-approach, should generally follow the principles described in the Guideline on the principles of regulatory acceptance of 3Rs testing approaches (EMA/CHMP/CVMP/JEG-3Rs/450091/2012). Non-animal approaches encompass in vitro, in silico, or in chemico methods. These include both simple and complex human cell-based assays, microphysiological systems (MPS) such as organ-on-chip (OoC), as well as in silico and computational models incorporating machine learning and artificial intelligence (AI). They have the potential not only to reduce NHP use, but also to increase the translational relevance of non-clinical studies to humans, by utilising human-derived cell systems or data.

For any non-animal approach, a description of the circumstances under which it will be used in the assessment of human medicinal products, and the limitations within which the available data adequately support its use, will be needed to allow regulatory acceptance. Further information about the procedures to reach regulatory acceptance and ways that developers can interact with regulators are described in Section 8. In specific cases, such approaches have been successfully used to enable First in Human (FiH) trials without the need for NHP studies (3). Non-conventional alternative approaches to minimise the need for NHP studies should be explored with regulatory authorities as early as possible to allow meaningful integration into a regulatory acceptable non-clinical strategy.

5. Non-rodent species selection and the need for a study in NHPs

To ensure reliable non-clinical risk assessment of pharmaceuticals, the species selected for non-clinical safety testing should be pharmacologically or toxicologically relevant and studies should be feasible. A scientific justification for the species selected for the non-clinical safety studies should be provided in any regulatory application for a medicinal product (e.g., clinical trial application, marketing authorisation application (MAA)). Thorough consideration should be given to the use of all potential non-rodent experimental species (e.g., rabbit, dog, minipig) and NHPs should only be considered if the aims of the study cannot be achieved using any other non-rodent species. Practical aspects, such as animal size (i.e. volume of test material), availability of relevant historical control data and analytical tools could also form part of the justification for species selection. For DART studies, information on the advantages and disadvantages of using various non-rodent species, including NHP, have been provided in Annex 1 of ICH S5(R3). Selection criteria should be based on similarity to humans in terms of physiological characteristics, (functional) target binding, tolerability, formation of human-relevant metabolites and the pharmacokinetic profile. For biotechnology-derived medicinal products, the identification of

pharmacologically relevant models should ideally be based on sequence homology of the binding epitope, binding affinity (including binding kinetics, baseline target expression levels and receptor occupancy) and functional activity.

6. NHP use considerations by modality

6.1. Considerations for small molecules

For small molecules, in cases where NHP is the only pharmacologically relevant species, the need and feasibility of a specific study should be scientifically justified. To this end, all relevant data for the product and biological target, including data emerging during initial non-clinical and clinical studies, should be taken into account.

In the context of DART, rodent data is generally considered sufficient for assessment of fertility and NHP do not appear to provide more human-relevant data. For embryo-foetal developmental (EFD) toxicity studies, rabbits are considered the routine non-rodent species in line with ICH S5(R3).

No class -specific guidance exists for small synthetic peptides and non-clinical testing strategies often align with ICH M3(R2). However, unless significantly modified (e.g., to alter physicochemical properties), synthetic peptides often behave as their endogenous or recombinant biotechnology derived counterparts do. Therefore, with accumulating experience, a non-clinical strategy in line with ICH M3(R2) may not always be needed. The opportunities to follow ICH S6(R1) should be discussed with EMA or national competent authorities (NCAs) at an early stage and a WoE approach should be used to justify the nonclinical strategy.

6.2. Consideration for biotechnology products

For molecules produced through recombinant biotechnological processes, and monoclonal antibodies (mAbs) in particular, NHP is often the only available test species that is pharmacologically responsive or cross-reactive with the IMP. This has been acknowledged in ICH S6(R1), which proposes a case-by-case approach in the design of the non-clinical testing program. With the increasing interest in the development of multi-specific mAbs, fusion proteins and other complex biologics, target specificity often results in NHP not expressing the target or NHP showing limited recapitulation of the pharmacology. In such cases, alternative approaches are more important as the translatability of in vivo data to clinical practice decreases (see also the section on lack of a relevant species).

For RDT studies when both rodent and non-rodent species are pharmacologically responsive and the outcomes of short-term studies in rodent and non-rodent are similar, it is recommended to conduct chronic RDT testing in the rodent species only. In certain cases, it may even be sufficient to conduct all RDT studies in the rodent species only (e.g. where the biological activity of the biopharmaceutical is well understood and the rodent species is pharmacologically relevant). If a surrogate molecule (homologous protein in a species expressing an ortholog of the human target) that is active in rodent is available or can be generated, it is recommended to use this molecule for hazard identification to reduce or replace use of NHP in a non-clinical safety testing program. Similarly, if a transgenic rodent model expressing the human target exists that responds to the drug candidate or that can be used to evaluate effects of alteration of the intended target, the use of such a model should be considered. Both of these testing strategies need scientific justification.

In scientific literature, a WoE approach has been proposed to evaluate whether a conventional 6-month RDT study in NHP is necessary to establish the safety profile of mAbs (4). The main

rationale for this approach is based on the intrinsic properties of mAbs, with adverse effects largely related to their pharmacological action. Other WoE approaches have successfully led to regulatory acceptance of reduced NHP testing (5). Therefore, WoE approaches can be acceptable on a case-by-case basis and should be discussed via formal or informal EMA or NCA procedures (see also section 8).

In line with ICH S6(R1), dedicated fertility studies are not required for biotechnology products when NHP is the only pharmacologically relevant species. For these products, effects on fertility can be assessed based on evaluations included in RDT studies. Similarly, if a product is developed for an indication that does not include women of child-bearing potential (WOCBP), or in cases where other scientific justifications can be made, DART studies may not be needed at the time of MAA. Several publications have suggested that a standardised WoE, based on common risk factors, can be developed for DART studies with biotechnology products which commonly use NHP as a test species (6). An evaluation of the model developed by Rocca et al. shows its utility in demonstrating the need for an enhanced pre- and postnatal development (ePPND) study (7). Interestingly, despite the model being considered conservative (i.e. it suggested that a DART study was needed despite a negative in vivo outcome), it identified all products with a risk of developmental toxicity and it was concluded that it would have resulted in a 42% reduction of NHP DART studies without loss of important patient information in the label (7). Similar to other non-conventional WoE approaches, a WoE to support clinical trials and a MAA without an ePPND study can be accepted on a case-by-case basis and should be discussed via formal or informal EMA or NCA procedures (see section 8).

Lack of relevant species / exogenous target

For certain products — such as pharmaceuticals designed for highly human-specific targets, or targets not expressed in humans or animals — regulators may accept a non-clinical safety assessment based on in vitro or in silico data, integrated in a WoE approach that considers knowledge on the target and relevant non-clinical and clinical data from similar medicinal products. While ICH S6(R1) suggests the need for short duration animal studies in a non-pharmacologically responsive species, it is now acknowledged that these do not reliably inform on clinically relevant hazard. If a surrogate molecule or transgenic animal model that expresses the human target is available, and a totality of evidence (via in vitro, in silico and WoE) is insufficient in for characterising risk, such models are preferred.

Immunogenicity

Many biotechnology products intended for human use are immunogenic in NHP. A review of EMA scientific advice procedures has shown that when study-limiting immunogenicity was observed in NHP studies, subsequent studies (e.g., long term RDT or DART studies) were generally waived. These advice procedures included products in development for chronic as well as advanced cancer indications. The products were generally first in class or complex (e.g. bi-specific mAbs). In line with this, when immunogenicity significantly limits systemic exposure, leads to a substantial or complete loss of pharmacological activity, or is associated with limiting anti-drug-antibody (ADA)-mediated toxicity, subsequent RDT studies of longer duration in NHP are generally not recommended. This should always be evaluated on a case-by-case basis.

Considerations for antibody-drug conjugates (ADCs)

For ADCs developed for advanced cancer indications, the ICH S9 guidance and ICH S9 Q&A recommend that non-clinical safety testing can be limited to the evaluation of the complete ADC molecule if the payload is well-characterised. In a series of retrospective analyses conducted by the US-FDA, animal toxicity data of ADCs in investigational medicinal product (IMP) applications

were compared to clinical outcomes from phase I studies (8,9). A conclusion of these studies was that dose-limiting toxicity is almost exclusively driven by the cytotoxic payload, as most clinical findings were predicted both by studies with the ADC and payload alone. This indicates that the value of safety studies of ADCs in pharmacologically relevant species, including NHPs, may be limited and provides further justification for using rodents instead of NHP in safety evaluation of ADCs. Based on this experience, this approach could also be considered when a novel linker or cytotoxic payload is developed for the ADC on a case-by-case basis and should be supported by a WoE approach.

Therefore, it is possible to cross-refer to payload-related toxicity studies and clinical data from other relevant ADC developmental programs, complemented with a WoE -based assessment of potential target-related toxicity, to further reduce NHP use. If the payload (or payload + linker) is new, it can be evaluated in a separate arm in the toxicological studies investigating the ADC.

6.3. Considerations for Vaccines

When selecting species for safety testing of vaccines, the ability to mount an immune-response to the vaccine (and in specific cases of live-attenuated vaccines, the sensitivity to the pathogen/toxin), should be considered as recommended in WHO guidelines on non-clinical evaluation of vaccines (WHO Technical Report Series No 9279, 2025) and on non-clinical evaluation of vaccine adjuvants and adjuvanted vaccines (WHO Technical Report Series No. 987, 2014). In principle, NHPs should only be used if no other relevant animal species is available.

6.4. Considerations for ATMPs

Advanced therapy medicinal products (ATMPs) are a heterogenous group of products where complexity and species-specificity can differ greatly between different types of products. The study requirements are thus specific to the type of product. In addition, a risk-based approach following the Guideline on the risk-based approach according to annex I, part IV of Directive 2001/83/EC applied to Advanced therapy medicinal products (EMA/CAT/CPWP/686637/2011) is generally recommended to identify which non-clinical data is needed. Current regulatory guidelines on ATMPs do not provide specific recommendations regarding the use of NHPs. However, they do provide general guidance on 3Rs opportunities that can be applied on a case-by-case basis to replace or reduce the use of NHPs as outlined in EMA/CAT/22473/2025. Where appropriate, in silico, in vitro and/or ex vivo data can be used to substitute or supplement in vivo animal data (directive 2010/63/EU)). Alternative in vivo models can also be considered, for example immunodeficient animals or homologous animal models using a species-specific vector/transgene or the respective cells from the same animal species.

Use of NHP in safety testing of ATMPs is mainly restricted to gene therapy medicinal products (GTMPs) using novel viral vectors as a delivery system, where NHP selection is driven by perceived pharmacological and anatomical relevance, sufficient sequence homology with humans and technical feasibility of the administration procedure. However, other test species, such as dogs and rats, have also been accepted by regulatory authorities when sufficient homology (e.g. regarding the gene product and tropism of the vector) between the chosen animal species and humans has been demonstrated. For GTMPs that use a well-characterised vector system, applying proprietary or publicly available data supported by a WoE can be used to waive certain or all in vivo safety studies, such as biodistribution and shedding studies (see ICH S12).

Regarding cell-based GTMPs (e.g. CAR-T cells), cell therapy medicinal products and tissue-engineered products, immunogenicity of the human-specific drug candidate (or the lack of

necessary factors to support survival of human cells) often hinders conventional safety evaluation of the clinical drug candidate in animals, including NHPs. Moreover, for highly human-specific immune-cell based ATMPs (e.g. CAR-T cells) in vivo safety studies are not always appropriate or even feasible and the use of NHP is not recommended. (See ICH S12 and Guideline on quality, non-clinical and clinical requirements for investigational advanced therapy medicinal products in clinical trials).

6.5. *Other modalities*

Regarding new modalities, the appropriate use of NHP should be outlined and scientifically justified. Use of alternative models is recommended over animal studies when such models can provide appropriate characterisation of risk. Non-clinical testing strategies using novel approaches should be cross -referenced with any relevant newly developed guidelines, such as the ICH guideline on non-clinical safety evaluation of oligonucleotide-based therapeutics (ICHS13 Concept paper: Non-clinical Safety Evaluation of Oligonucleotide-based Therapeutics).

7. NHP study design and testing strategy considerations

Opportunities exist in medicinal product development programs to reduce the number of NHPs used through holistic consideration of the overall non-clinical program and individual study design. An informed animal study design can be beneficial to the scientific value of a study as well as the number of animals used. Aside from study duration, ICH M3(R2) does not define details of study design, in this reflection paper the term 'short-term studies' is used to describe studies shorter than 1 month, 'sub-chronic studies' describe studies above 1 month and below 6 months, and chronic studies are used to define study durations from 6 months onwards.

7.1. Number of animals per group

In general, the size of the treatment group should be sufficient to allow for meaningful scientific interpretation of the data generated with consideration for ethical as well as practical aspects (EMA Guideline on repeated dose toxicity (CPMP/SWP/1042/99 Rev 1 Corr*)).

ICH M3(R2) does not provide specifics on group size except for the extended single dose toxicity study where the usual design for non-rodents consists of 3/sex /group for all groups on day 2 and 2 animals/sex/group for the dose levels assessed on day 14.

For biotechnologically derived pharmaceuticals, ICH S6(R1) indicates that the number of animals used per dose has a direct bearing on the ability to detect toxicity. Limitations imposed by sample size may be in part compensated by increasing the frequency and duration of monitoring.

With regards to RDT studies in non-rodent studies, dose groups usually consist of at least 3 animals/sex/group, with an additional 2/sex/group for recovery, if appropriate (see ICH S9). Both sexes should generally be used, or justification should be given for specific omissions. These numbers have been confirmed in retrospective analyses with 3(+2) animals/sex/group being used for 1-month repeat dose toxicity studies, and 4 animals/sex/group in studies of longer duration (e.g. 3- and 6- or 9 -months) (10,11,12).

While it is agreed that the lowest possible number of animals should be used, there is no consensus on an ideal approach. RDT studies are designed to characterise potential hazard and are not powered to detect statistical significance beyond basic descriptive statistics. Increasing the number of animals from a common 3 animals/sex/group (main study group animals) design to 4 or 5 animals/sex/group design for 3- to 6/9-month repeated dose toxicity studies, does not

meaningfully increase confidence in a reported biological effect. Therefore, if an NHP study is considered necessary, a 3 animals/sex/group (main study group animals) design is considered sufficient. Reductions in control group size are considered acceptable provided sufficiently large historical control datasets are available.

Specific considerations apply with regards to the use of non-rodents for in vivo assessment of potential delayed ventricular repolarization (QT prolongation) (see ICH S7B and ICH E14/S7B Q&A). If study results are to be used to support an integrated non-clinical and clinical risk assessment as described in ICH E14 Q&A 6.1, the study should have sensitivity to detect a QTc prolongation effect of a magnitude similar to dedicated clinical QT prolongation studies, taking into consideration interspecies differences in the normal range of values for the QTc interval. The overall sensitivity of the non-clinical assay in comparison to clinical QT prolongation studies depends on both the electrocardiographic assessment and the exposure achieved in the in vivo assay relative to high clinical exposure. These principles also apply when in vivo QT assessment is integrated in RDT studies (as recommended in ICH M3(R2) and ICH S9). ICH S7A relates group size to the detection potential, stating that the number of animals should be adequate to demonstrate or rule out the presence of biologically relevant effects.

NHP group sizes in ePPND studies (see ICH S5(R3) Annex 1 and ICH S6(R1)) should yield a sufficient number of infants (6-8 per group at postnatal day 7) to assess an effect on postnatal development and provide the opportunity for specialist evaluation, if warranted (e.g., for effects on immune system). Recent experience has shown that fewer pregnant NHP can be used to achieve this compared to the recommendations in ICH S6(R1) (13).

7.2. Number of dose groups (including control groups)

The application of approaches to minimise the number NHPs used in studies through reducing dose groups requires global regulatory acceptance within the framework of existing guidelines, taking into account 3Rs principles and statistical considerations. Scientific justification of the dose levels should be supported by data (e.g., based on clinical exposure data, saturation of target binding or from experience with similar molecules). Control data should be available from animals treated with the same vehicle and route of administration, with ongoing monitoring data to confirm the absence of drift of key parameters.

Under ICH S6(R1), ePPND studies in NHPs are considered hazard identification studies and, as such, the use of a control group and only one dose group can be acceptable provided that the dose level selected is scientifically justified. This scientific justification needs to take into consideration aspects such as clinical exposure and saturation of target binding.

Virtual control groups (VCGs) are a promising tool to reduce the number of animals, including NHP, needed for safety testing. The use of concurrent control groups (CCG) in a study could be partially or fully replaced by VCGs through the use of matching criteria to select historical controls from a virtual historical database of NHP controls. Promising initiatives to develop such virtual control databases are ongoing, but further development is needed for regulatory acceptance. Key requirements for VCG approaches to succeed are wide access to VCG data and positive consensus from (global) regulatory agencies, culminating into (global) guidance.

To facilitate regulatory acceptance for VCG approaches, it is encouraged that applicants share use of VCG data in regulatory applications. Where possible, VCG data (full or hybrid) could be submitted for non-GLP studies (e.g. dose range finding studies) to generate confidence and experience with this system.

7.3. Need for both sexes

While sex is considered an important biological variable, there may be some instances where conduct of safety studies in only one sex is deemed appropriate based on scientific justification. Factors that can be considered include clinical prevalence of the disease or absence of sex differences in target expression levels. For biotechnology derived pharmaceuticals, ICH S6 (R1) states that both sexes should generally be used or justification given for specific omissions.

ICH M3(R2) Q&A indicates that exploratory clinical studies do not represent a commitment to full development. Therefore, when the intention is to conduct the exploratory clinical study in one sex only, the single-dose toxicity studies can be restricted to that sex, provided that animal group sizes for the day 2 termination are increased, as it is normal to combine effects from both sexes with respect to identifying and characterizing toxicities that are not sex-specific.

7.4. Need for recovery groups

ICHM3(R2) Q&A, section 2 (1) provides practical recommendations to assist in the identification of the need for reversibility assessment in non-clinical toxicity studies.

In general, evaluation of reversibility should only be provided when severe toxicity findings with potential clinical impact are described in a non-clinical study. This evaluation of reversibility can be performed by scientific assessment or by including recovery groups in a toxicity study. Therefore, the inclusion of recovery arms is not always considered critical to determine whether an adverse event is reversible.

The timing and need for recovery groups should be scientifically justified and duly considered in line with 3Rs considerations (see ICH M3(R2) Q&A and ICH S9 Q&A)). Inclusion of recovery animals is not needed when toxicity (e.g., identified in dose range finding or pharmacology studies):

- can be readily monitored in humans at an early stage before the toxicity becomes severe or is known to be irrelevant to humans or
- is only observed at high exposures that are not considered clinically relevant (>10-fold clinical exposure multiple) or
- is similar to that induced by related agents, and the toxicity based on prior clinical experience with these related agents is considered a manageable risk.

For biotechnology derived pharmaceuticals, ICH S6(R1) also indicates that recovery from pharmacological and toxicological effects with potential adverse clinical impact should be understood when these occur at clinically relevant exposure. In this case, assessment of reversibility in one study at one dose level – both scientifically justified – is considered appropriate.

A toxicity study that includes a recovery group is generally warranted if a scientific assessment cannot predict whether the toxicity will be reversible and if:

- there is severe toxicity at clinically relevant exposures (e.g., ≤10-fold the clinical exposure) or
- the toxicity is only detectable at an advanced stage of the pathophysiology in humans and significant reduction in organ function is expected. The assessment of reversibility in this case should be considered even at >10-fold exposure multiples.

Where a study of reversibility is deemed necessary, it should be available to support clinical studies of a duration similar to that at which the adverse effects were seen non-clinically. However, a reversibility study is generally not warranted to support clinical trials of a duration equivalent to that at which the adverse effect was not observed non-clinically.

The duration of the recovery period should be sufficient to allow an evaluation of the reversibility of effects (e.g., based on the nature of known or predicted end of dose findings). However, demonstration of complete recovery of all effects is not essential in general. For biotechnology products, the typically long half-life should be taken into account to determine an adequate recovery period. It should be noted, in line with ICH S6(R1), that the addition of a recovery period just to assess potential for immunogenicity, is not required. In the absence of more specific data suggesting when to assess recovery, a recovery period of 5–7 half-lives has been suggested for molecules with half-lives of one week or longer, as approximately 95% of the drug is cleared after the fourth half-life (>99% by the seventh) (14).

In addition to these general recommendations, additional refinement of recovery animal strategies are possible and are extensively described in the literature. From a regulatory point of view, these approaches can be acceptable on a case-by-case basis. It is recommended to carefully evaluate the suitability of these approaches for each planned study (14, 15).

8. Regulatory acceptance of alternative approaches and future perspectives

The overarching principles of regulatory acceptance of 3Rs testing approaches are described in EMA guidance (EMA/CHMP/CVMP/JEG-3Rs/450091/2012). In addition, EMA has published a reflection paper on testing requirements and 3Rs opportunities. Though not restricted to NHP studies, this paper describes 'implemented' and 'newly identified' opportunities for 3Rs implementation. The latter generally refers to approaches that are currently under investigation and will necessitate data review and further discussion before a definitive impact on 3Rs can be achieved.

Regulatory assessment of non-clinical safety data is performed on a product-specific basis. While existing guidance provides a framework for the conduct of safety studies, scientific and technological advances have enabled, and will enable, development of alternative approaches not yet covered by regulatory guidance. Alternative methods and 3Rs-based adaptations to NHP testing which are not described in current available guidance can be discussed with regulators at various stages of the development of a novel method and/or medicinal product. Mechanisms for these discussions include the Innovation Task Force (ITF) and national or EMA Scientific Advice. Novel methodologies, particularly those with a 3Rs application, can also be considered for Qualification. Finally, EMA also accepts voluntary data submission (VDS) obtained by using a new 3Rs testing approach in parallel with data generated using 'existing' (typically in vivo animal) methods. Data generated with the new 3Rs testing approaches will not be used as part of the regulatory decision-making process and should be evaluated independently and solely for the purpose of evaluation of the novel approaches for possible future regulatory acceptance of testing approaches. VDS therefore allows regulators to gain experience with new approach methodology data and provides the opportunity to achieve early regulatory alignment.

Future Perspectives

Regulatory science and innovation continue to develop at a rapid pace, including the development of novel approaches and sophisticated analysis tools using AI and machine learning. These have significant potential to reduce the number of NHPs required for regulatory testing. A single method

will not replace, in a one-to-one context, the use of an NHP for the assessment of risk. Therefore, there is a focus on WoE approaches that may or may not incorporate innovative non-animal methodologies.

Using WoE approaches as an example, strategies to minimise NHP use that are currently outside of guidance (see Section 3) could be further developed and used more generally to justify the decision to conduct or omit specific studies. Immediate use cases already exist and have been accepted by regulatory authorities in the form of follow-on biologics and/or products that bind to highly characterised targets since risk has often been largely characterised. In addition, alternative approaches can be accepted by regulators for products where the target is either not present in humans and animals, is not replicated in animals, is not expressed at human-relevant levels or has a different pharmacological function in NHP. A WoE model could also be considered to minimise the need for specific studies with NHP for severely debilitating and life-threatening indications, in line with the ICH S9 guideline on advanced cancer therapeutics, where residual risk can be accepted given the severity of the indication. WoE approaches can also be readily applied to product classes such as ATMPs, which are highly complex and species specific and, to a lesser extent or on a case-by-case basis, to oligonucleotides and small molecules.

Generating regulatory experience with WoE approaches is essential and can be achieved both within and outside regulatory procedures: using the following approaches:

- ***Retrospective WoE exercises to evaluate the need for animal studies with NHP***

  Retrospective WoE analysis can be performed by regulatory authorities or companies with marketed products and/or non-marketed products within their own portfolio. Such analysis allows for increased experience without any regulatory consequence.

- ***Prospective WoE exercises to evaluate the need for animal studies with NHP***

  Prospective exercises benefit greatly from VDS by pharmaceutical companies to regulatory authorities. These regulatory sandboxes are essential to foster innovation and allow scientific advancements to gradually enter into regulatory practice.

9. Recommendations and Conclusion

This reflection paper outlines the currently available opportunities to reduce, refine or replace the use of NHPs in pharmaceutical development, and identifies promising advances that, with accumulating evidence, may become more broadly implemented and acceptable in the near future. Given the rapidly evolving landscape in non-clinical safety testing, additional opportunities are anticipated to arise over the coming years.

Based on the reflections in this paper, the following recommendations are made in relation to the use of NHPs for regulatory testing:

- Use of NHPs in non-clinical safety testing should be scientifically justified and considered only as a last resort. The number of animals used can and must be limited to the absolute minimum to allow characterisation of risk for clinical trial participants.
- Applicants should make full use of existing guidance documents, which generally offer multiple opportunities for 3Rs implementation that are currently underutilised, to reduce reliance on NHPs in drug development.
- Innovative, non-standard, WoE strategies that may scientifically justify avoiding NHP studies in non-clinical development are encouraged.

- For novel methodologies not yet addressed in regulatory guidance, pharmaceutical and assay developers are urged to engage with the EMA ITF briefing meetings, VDS, and scientific or qualification advice procedures.
- When the use of NHPs is being considered for non-clinical safety studies of pharmaceuticals, applicants are encouraged to engage with regulatory authorities early in the development process. This includes seeking regulatory alignment, particularly when exploring 3Rs approaches that are in- or out of scope of ICH guidance.

While complete replacement of NHP use may remain challenging, accumulating evidence indicates that replacement is often achievable and can be aligned with current regulatory acceptance. Applicants are encouraged to use opportunities for alternative testing strategies already outlined in existing guidelines. The approaches described in this reflection paper offer substantial steps towards considerable reduction in the short- to medium term. In addition, advancing novel approaches including non-animal methods is critical from a scientific and ethical point of view and will be achievable in the long term through innovation, collaborative efforts and proactive regulatory engagement.

10. References

1. Brown PC, Wange RL. Considerations regarding the use of nonhuman primates in assessing safety endpoints for pharmaceuticals. Regul Toxicol Pharmacol. 2023 Sep;143:105449. doi: 10.1016/j.yrtph.2023.105449. PMID: 37453554.

2. David Ackley, Joanne Birkebak, Jorg Blumel, Todd Bourcier, Christina de Zafra, Andrew Goodwin, Wendy Halpern, Danuta Herzyk, Sven Kronenberg, Robert Mauthe, Jacintha Shenton, Dana Shuey, Ronald L. Wange. FDA and industry collaboration: Identifying opportunities to further reduce reliance on nonhuman primates for nonclinical safety evaluations. Regul Toxicol Pharmacol. 2023,105327,ISSN 0273-2300, doi: 10.1016/j.yrtph.2022.105327.

3. Beilmann, M., Adkins, K., Boonen, H.C.M. et al. Application of new approach methodologies for nonclinical safety assessment of drug candidates. Nat Rev Drug Discov (2025). doi: 10.1038/s41573-025-01182-9.

4. Chien HT, Prior H, Andrews L, van Aerts L, Cauvin A, Clarke DO, Datta K, Dempster M, Dybdal N, Freebern W, de Haan L, Herzyk D, Hey A, Kissner T, Kronenberg S, Leach MW, Lee D, Schutte K, Sewell F, Trouba K, Ulrich P, Weir L, van Meer P. Re-evaluating the need for chronic toxicity studies with therapeutic monoclonal antibodies, using a weight of evidence approach. Regul Toxicol Pharmacol. 2023 Feb;138:105329. doi: 10.1016/j.yrtph.2022.105329. PMID: 36592682.

5. Hoberman AM, Maki K, Mikashima F, Naota M, Wange RL, Lansita JA, Weis SL. Alternatives to Monkey Reproductive Toxicology Testing for Biotherapeutics. Int J Toxicol. 2023 Dec;42(6):467-479. doi: 10.1177/10915818231200859. PMID: 37714565.

6. Rocca M, Morford LL, Blanset DL, Halpern WG, Cavagnaro J, Bowman CJ. Applying a weight of evidence approach to the evaluation of developmental toxicity of biopharmaceuticals. Regul Toxicol Pharmacol. 2018 Oct;98:69-79. doi: 10.1016/j.yrtph.2018.07.006. PMID: 30009863.

7. Chien HT, Roos P, Russel FGM, Theunissen PT, van Meer PJK. The use of weight-of-evidence approaches to characterize developmental toxicity risk for therapeutic monoclonal antibodies in humans without in vivo studies. Regul Toxicol Pharmacol. 2024 Sep;152:105682. doi: 10.1016/j.yrtph.2024.105682. PMID: 39094967.

8. Saber H, Leighton JK. An FDA oncology analysis of antibody-drug conjugates. Regul Toxicol Pharmacol. 2015 Apr;71(3):444-52. doi: 10.1016/j.yrtph.2015.01.014. Epub 2015 Feb 7. PMID: 25661711.

9. Saber H, Simpson N, Ricks TK, Leighton JK. An FDA oncology analysis of toxicities associated with PBD-containing antibody-drug conjugates. Regul Toxicol Pharmacol. 2019 Oct;107:104429. doi: 10.1016/j.yrtph.2019.104429. Epub 2019 Jul 17. PMID: 31325532.

10. Sparrow SS, Robinson S, Bolam S, Bruce C, Danks A, Everett D, Fulcher S, Hill RE, Palmer H, Scott EW, Chapman KL. Opportunities to minimise animal use in pharmaceutical regulatory general toxicology: a cross-company review. Regul Toxicol Pharmacol. 2011 Nov;61(2):222-9. doi: 10.1016/j.yrtph.2011.08.001. PMID: 21855593.

11. Prior H, Baldrick P, Clarke DO, Passini E, Sewell F, van Meer P. Exploring Greater Flexibility for Chronic Toxicity Study Designs to Support Human Safety Assessment While Balancing 3Rs Considerations. Int J Toxicol. 2024 Sep-Oct;43(5):456-463. doi: 10.1177/10915818241255885. PMID: 38822557; PMCID: PMC11408957.

12. Leung, C., Jessen, B., Rattel, B., Ackley, D., Dybdal, N., Hopfer, U. and Hart, T., 2025. Considerations of non-human primate use in nonclinical toxicity study package for oncology therapeutics with well-characterized target in supports of 3Rs–an IQ DruSafe industry survey. Regul Toxicol Pharmacol. 2025 Nov;162:105902. doi: 10.1016/j.yrtph.2025.105902. Epub 2025 Jul 7. PMID: 40633898.

13. Luetjens CM, Fuchs A, Baker A, Weinbauer GF. Group size experiences with enhanced pre- and postnatal development studies in the long-tailed macaque (Macaca fascicularis). Primate Biol. 2020 Mar 11;7(1):1-4. doi: 10.5194/pb-7-1-2020. PMID: 32232119.

14. Pandher K, Leach MW, Burns-Naas LA. Appropriate use of recovery groups in nonclinical toxicity studies: value in a science-driven case-by-case approach. Vet Pathol. 2012 Mar;49(2):357-61. doi: 10.1177/0300985811415701. PMID: 21810619.

15. Prior H, Andrews L, Cauvin A, Chien H, Clarke DO, Datta K, Dempster M, Dybdal N, Freebern W, de Haan L, Herzyk D, Hey A, Kissner T, Kronenberg S, Leach MW, Lee D, Reid K, Schutte K, Sewell F, Trouba K, Ulrich P, van Aerts L, van Meer P, Weir L. The use of recovery animals in nonclinical safety assessment studies with monoclonal antibodies: further 3Rs opportunities remain. Regul Toxicol Pharmacol. 2023 Feb;138:105339. doi: 10.1016/j.yrtph.2023.105339. PMID: 36649820.