Exploring the Definition of "Similar Toxicities": Case Studies Illustrating Industry and Regulatory Interpretation of ICH S6(R1) for Long-Term Toxicity Studies in One or Two Species

Helen Prior[1] ⬤, David O Clarke[2], David Jones[3], Eleni Salicru[4] ⬤, Melissa M Schutten[5], and Fiona Sewell[1] ⬤

Abstract

ICH S6 (R1) states that safety evaluation of biotherapeutics should normally include 2 relevant species when available (i.e., a rodent and non-rodent species in which the test material is pharmacologically active), at least for short-term toxicology studies (generally supporting Phase I trials). For subsequent long-term toxicology studies (e.g., chronic studies up to 6 months dosing duration), there are options to reduce to only one species when justified, including when the mechanism of action of the biologic is well-understood or the toxicity findings in the short-term studies are "similar" in both the rodent and non-rodent species. Across the industry, around 25 to 33% of biologics assess multiple species within short-term toxicity studies but it is often unclear how different companies and regulators are applying the ICH S6 (R1) principles of "similar toxicity profiles" to progress with either 1 or 2 species in the long-term studies, in particular whether the absence of toxicities is considered within this definition. Sponsors may potentially continue to use 2 species to avoid regulatory risk and potential delays in development timelines, representing missed opportunities for reducing animal use, particularly of non-human primates, during drug development.

This article summarizes presentations from a symposium at the 41st Annual meeting of the American College of Toxicology (ACT) in November 2020, in which industry case studies and regulatory perspectives addressed considerations and decisions for using 1 or 2 species for long-term toxicity studies, highlighting any common themes or experience that could be applicable for use in future decision-making.

Keywords

biologics, ICH S6(R1), long-term studies, non-rodent, rodent

Introduction

Biotechnology-derived therapeutics (hereafter referred to as biologics) are continuing to expand within company pipelines and the marketplace, with 12 new antibody-based therapeutics licensed by the US Food and Drug Administration (FDA) in 2020.[1] The design of human clinical trials and the safety of participating volunteers and patients is usually supported by non-clinical data from toxicity studies in animals, conducted in compliance with specific regulatory guidance, namely ICH S6 (R1) Preclinical safety evaluation of biotechnology-derived pharmaceuticals[2] in conjunction with general principles as outlined within ICH M3 (R2) Non-clinical safety studies for the conduct of human clinical trials for pharmaceuticals.[3] A key determinant in species selection for toxicity testing is pharmacological relevance, based on species expression of the target (receptor or epitope) and appropriate functional engagement by the biological product that evokes a similar pharmacological response as that expected in humans. As many biological products

[1]National Centre for the Replacement Refinement and Reduction of Animals in Research, London, UK
[2]Lilly Research Laboratories, Lilly Corporate Center, Indianapolis, USA
[3]Medicines Healthcare products Regulatory Agency (MHRA), London, UK
[4]Food and Drug Administration (FDA), Silver Spring, ML, USA
[5]Genentech Inc, South San Francisco, CF, USA

Corresponding Author:
Helen Prior, National Centre for the Replacement Refinement and Reduction of Animals in Research, Gibbs Building, 215 Euston Rd, London NW1 2BE, UK.
Email: helen.prior@nc3rs.org.uk

are highly selective, often there is only 1 relevant species identified and this is frequently a non-human primate (NHP), owing to higher genome sequence identity to humans and similarity in physiological systems. Nevertheless, when a biologic does allow consideration of multiple pharmacologically relevant species, toxicity data are expected in 2 species (a rodent and non-rodent) in a similar manner as new chemical entities such as small molecules, at least for short-term studies. Although the ICH S6 (R1) guidelines state these as "studies up to 1 month dosing duration," longer studies (e.g., up to 3 months dosing duration) are sometimes performed and the guidance is generally taken to refer to any first-in-human (FIH)-enabling toxicity studies (often referred to as IND-enabling toxicity studies in the US) supporting Phase I trials. If the toxicological findings from these short-term studies are similar between species or the findings are understood from the mechanism of action of the product, then a long-term toxicity study (chronic study up to 6 months dosing duration) in 1 species is usually considered sufficient, with the rodent progressed unless there is a scientific rationale for using non-rodents.

A key decision for progression to long-term studies in 1 or 2 species can therefore revolve around the "similarity" of toxicities observed in the short-term studies in rodent and non-rodent species. However, definitions of "similar toxicity profile in 2 species" remain vague and there is a lack of clarity on how to apply this in practice. Additionally, because high target specificity and selectivity is typical for biologics, these molecules often, but not always, exhibit non-adverse effects related to their primary pharmacology and an absence of off-target toxicity; hypersensitivity-related reactions due to im-munogenicity are also often observed. Whether these scenarios constitute a "similar" toxicity profile and how then to decide which species to progress, are uncertainties often encountered. As a consequence, Sponsors may continue to use 2 species to avoid regulatory risk and potential delays in development timelines. Such decisions can represent missed opportunities for applying the 3Rs (replacement, reduction, and refinement) and reducing animal use, particularly of NHPs, during drug development.

A Symposium was held in November 2020 as part of the 41st annual meeting of the American College of Toxicology to discuss this topic and to share experience for promoting ap-proaches where reducing to a single species for long-term toxicity studies is scientifically justifiable. The speakers and audience explored case-studies from both industry and reg-ulatory agencies to highlight practical examples of how and why decisions were made to reduce to a single species (or to retain 2 species) for long-term toxicity studies. The presen-tations and resulting discussions are summarized herein.

How often are 2 species used for short and/or long-term toxicity studies with biologics? (Helen Prior, NC3Rs)

Since 2016, the UK National Centre for the replacement, refinement, and reduction of animals in research (NC3Rs)

has led a large international working group data sharing activity to review the use of 2 species within regulatory toxicology packages. The data collected up to 2017 (172 drug candidates, from 18 companies) consisted of 92 small molecules and 80 biologics that were following ICH S6 (R1) guidelines (various drug modalities, as per next sentence). Short-term toxicology studies in 2 species were conducted for 41 biologics (24% of the data set), with recombinant proteins, synthetic peptides and antibody drug conjugates (ADCs) more likely to use 2 species (87%, 100 and 83% of each drug modality) than the monoclonal antibodies (mAbs) within this data set (30% of all mAbs)[4]. Five of the biologics (3 mAbs and 2 ADCs) reduced to a single species (the NHP) after initial, non-GLP toxicity studies in 2 species, generally due to immunogenicity or other project decisions. For the 36 biologics that used 2 species for FIH-enabling studies, no or similar toxicities were observed for 21 (58%), providing an opportunity to reduce to 1 species for subsequent long-term studies. However, only 11 biologics reported both short and long-term studies within the data set (the others having either stopped within development or not yet completed the later phase). For these, 5 biologics (2 mAbs, 1 recombinant protein, and 2 synthetic peptides) had different toxicities in-enabling studies and retained both species for the long-term study. The other 6 biologics (4 mAbs, 1 recombinant protein, and 1 synthetic peptide) had no or similar tox-icities FIH-enabling studies, yet only 2 mAbs reduced to 1 species for the long-term study, progressing the rat in both cases.

Other data regarding 1 or 2 species use for biologics has been obtained via reviews of publicly available regulatory submission documentation. A review of 39 mAbs submitted in Japan up to 2016[5] indicated that repeat-dose toxicity studies were conducted in 2 species for 36% of the mAbs. A review of 23 mAbs authorized by the European Medicines Agency between 2016 and 2019[6] indicated that short-term toxicity studies were conducted in 2 species for 35% of the mAbs, with the majority of these reducing to 1 species (the NHP in all cases) for the long-term studies. Other publications contain case study examples[7,8] or reports for individual mAbs[9,10,11] that describe the use of 2 species for short and/or long-term studies.

More recently, in conjunction with the Netherlands Medicines Evaluation Board (MEB), the NC3Rs collected data on 142 mAbs from 11 different pharmaceutical companies within a European Partnership for Alternative Approaches to Animal Testing (EPAA)-funded project, to investigate the optimal duration of non-clinical studies[12]. Included within this dataset were details of the species used in each short-term FIH-enabling study and in longer-term studies supporting later development (for studies conducted up to 2019), along with reasons for selection of the species. This provides another set of data on the in-cidence of multi-species use within mAb packages (23%

Table 1. Reasons for Maintaining Two Species for Long-Term Toxicity Studies (Data From EPAA-Funded MEB/NC3Rs Project).

Blinded mAb ID (Start date)[1]	Species Used (FIH Enabling Studies)	Reasons for Retaining the Two Species in Long-Term Studies	
		Primary reason(s)	Additional reason(s)
1 (2008)	NHP and rat	Different toxicities/sensitivities	—
2 (2008)	NHP and rat	Standard practice	Similar findings/Different sensitivities
3 (2009)	NHP and rat	Standard practice	Similar (no) findings
4 (2010)	NHP and rat	Regulatory request	—
5 (2010)	NHP and rat	Different toxicities/sensitivities	—
6 (2010)	NHP and rat	Standard practice	Different toxicities/sensitivities
7 (2011)	NHP and rat	Standard practice	—
8 (2011)	NHP and rat	Different toxicities/sensitivities	—
9 (2011)	NHP and rat	Standard practice	Similar (no) findings
10 (2012)	NHP and rat	Different toxicities/sensitivities	—
11 (2012)	NHP and rat	Regulatory request	—
12 (2012)[2]	NHP and rat	Regulatory request	—
13 (2013)	NHP and rat	Standard practice	—
14 (2014)	NHP and rat	Regulatory request	Similar (no) findings
15 (2015)	NHP and TG mouse	Different toxicities/sensitivities	
16 (2015)	NHP, rat and mouse[3]	Similar (no) findings	Timelines
17 (2017)	NHP and mouse	Different toxicities/sensitivities	
18 (2018)	NHP and mouse	Regulatory request	Similar (no) findings

Upper section groups the mAbs with studies in 2011 or earlier and lower section groups the mAbs with studies in 2012 or later (i.e., timings relative to ICH S6 (R1) revision). The reasons for species use in the long-term study were categorized by a sub-team of expert toxicologists, from answers provided within the main European Partnership for Alternative Approaches to Animal (EPAA)-Medicines Evaluation Board (MEB)/NC3Rs project survey. For some mAbs, more information was available allowing multiple categories to be reported. Categories were defined as follows: Similar or no findings: similar toxicities or expected pharmacology findings (non-adverse), or no findings between the species in short-term studies. Different toxicities or sensitivities: different toxicities or sensitivities between the species in short-term studies. Timelines: either limited time for consultation or in-licensed and studies ongoing. Regulatory request: single species proposed by Sponsor, but Health Authority requested the other species progress too. Standard practice: company policy, or older package pre-ICH S6 (R1); not common to reduce to 1 species/limited experience or feedback for the ICH S6 (R1) approach.
[1]Long-term study start date (year).
[2]This mAb performed short-term studies in non-rodent only and had added a rodent long-term study.
[3]This mAb used NHP and rat for the long-term studies. NHP was cynomolgus monkey in all cases. FIH: first-in-human (short-term studies).

of this dataset) and insights into decision-making for species used for biologics toxicity studies. Two species were tested in long-term toxicity studies for 18 mAbs (Table 1). The main reasons for these decisions included: 1) this was the standard practice at the time (around half of studies were performed prior to ICH S6 (R1) release, when two species data were generally expected unless scientific rationale precluded this; companies who performed studies soon after the guidance update may not yet have gained enough experience or regulatory feedback with a single species approach to confidently change standard practices); 2) different toxicities or sensitivities between responses in the 2 species, requiring further evaluation in both species; and 3) regulatory requests for 2 species, often when no toxicities had been identified in the short-term studies. There were 14 mAbs that reduced to a single species (Table 2) for the long-term toxicity studies. The main reasons for these decisions included 1) Immunogenicity in rodents, precluding their use in long-term toxicity studies; 2) progressing the NHP as the more pharmacologically relevant species, or the most sensitive species; and 3) similar or no toxicities in the 2 species.

Across the 32 mAbs overall, 10 mAbs had identified no toxicities in either species in the short-term studies; interestingly half of these mAbs retained both species for the long-term toxicity studies whilst the other half progressed only 1 species, highlighting the uncertainties this scenario often presents. We also looked for any trend for change in company practice since adoption of the ICH S6 (R1) guidance in 2011. Although the dataset is rather limited to draw clear conclusions, 2 out of 5 mAbs selected the rodent to progress to the long-term studies conducted after ICH S6 (R1) revision, compared to only 1 out of 9 mAbs for studies conducted before this time, leading to additional use of NHPs for the majority of mAbs.

The collective data clearly indicate that 2 species are used in initial, short-term toxicology studies for around a quarter to a third of all mAbs. For these molecules, the data demonstrate that reduction to 1 species for longer-term toxicology studies has been acceptable in a number of cases and for a variety of reasons; however, the data also identify potential missed opportunities where it might have been possible to reduce to a single species. By highlighting the considerations for or against reducing to 1 species based on the results of short-term studies and by illustrating case studies where a single species

Table 2. Reasons for Reducing to One Species for Long-Term Toxicity Studies (Data From EPAA-Funded MEB/NC3Rs Project).

Blinded mAb ID (Start date)[1]	Species Used		Reasons for Reducing to One Species for Long-Term Studies	
	FIH Enabling Studies	Long-Term Studies	Primary Reason	Additional reason(s)
1 (1994)	NHP and mouse	NHP	Immunogenicity in rodent	NHP more relevant
2 (2005)	NHP and rat	NHP	Similar (no) findings	—
3 (2006)	NHP and rat	NHP	NHP more relevant	Similar (no) findings
4 (2009)	NHP and rat	NHP	Immunogenicity in rodent	—
5 (2010)	NHP and rat	NHP	Similar findings	—
6 (2010)	NHP and rat	NHP	NHP more relevant	—
7 (2010)	NHP and guinea-pig	NHP	Most-sensitive species	—
8 (2010)	NHP and rat	NHP	Immunogenicity in rodent	NHP more relevant
9 (2011)	NHP and rat	Rat	Similar (no) findings	—
10 (2012)	NHP and mouse	NHP	Immunogenicity in rodent	—
11 (2016)	NHP and rat	Rat	Rodent (ICH S6)	Similar findings
12 (2017)	Minipig and rat	Rat	Rodent (ICH S6)	Similar (no) findings
13 (2018)	NHP and mouse	NHP	Similar (no) findings	NHP more relevant/Most-sensitive species
14 (2019)	NHP and rat	NHP	NHP more relevant	Most-sensitive species

Upper section groups the mAbs with studies in 2011 or earlier and lower section groups the mAbs with studies in 2012 or later (i.e., timings relative to ICH S6(R1) revision). The reasons for species use in the long-term study were categorized by a sub-team of expert toxicologists, from answers provided within the main European Partnership for Alternative Approaches to Animal (EPAA)-Medicines Evaluation Board (MEB)/NC3Rs project survey. For some mAbs, more information was available allowing multiple categories to be reported. Categories were defined as follows: Immunogenicity: immunogenicity prevented use of 1 species—other 1 progressed. Most-sensitive species: progressed the species where most toxicities or pharmacologies were observed. Non-rodent more relevant: progressed non-rodent as more pharmacologically relevant than rodent. Rodent (ICH S6): progressed rodent for ethical reasons, or to allow collection of additional data (e.g., Developmental and Reproductive Toxicity), or mentioned the ICH S6(R1) recommendation to use rodent. Similar or no findings: similar toxicities or expected pharmacology findings, or no findings in short-term studies.
[1]Long-term study start date (year); NHP was cynomolgus monkey in all cases, except mAb(1) where it was Rhesus macaque. FIH: first-in-human (short-term studies).

was appropriately progressed, this symposium aimed to promote best practices by companies and regulators for sound decision-making.

Industry Perspective 1 (Melissa Schutten, Genentech)

During the course of the drug development process, there are numerous opportunities to implement 3Rs principles with the ultimate goal of reducing animal usage, replacing animals with novel in vitro tools, and refining study designs. The accepted approach within ICH S6 (R1) to potentially reduce to a single preclinical animal species is a strategy that Genentech uses regularly to reduce overall animal use as part of internal 3Rs considerations. The following case examples illustrate the ways this has been achieved across multiple projects and therapy areas.

Case example 1: Molecule-specific example, where pharmacology and immunogenicity were key to decision-making for species progression. This example is illustrated for a humanized bispecific antibody (IgG, effector-less Fc), currently in clinical trials for the treatment of non-alcoholic steatohepatitis (NASH) and type 2 diabetes. The intended target is expressed in white and brown adipose tissue, the CNS and pancreatic β-cells. In light of the previously described biology[13], the

expected pharmacologic effects were significant weight loss and the improvement or reversal of obesity-associated metabolic derangements, including insulin resistance, hyperglycemia, dyslipidemia, and hepatosteatosis. The binding affinity for the presumed critical sequences of the epitope were tested in cynomolgus and rhesus monkeys, mouse, rat, rabbit, dog, minipig, and marmoset, with only cynomolgus monkey and mouse displaying similar affinity to humans. FIH-enabling toxicity studies were performed in these 2 species. One unique aspect of the non-clinical development program was the inclusion of Diet-Induced Obesity (DIO) mice in addition to standard CD-1 mice; DIO mice were included to demonstrate expected differences in diseased vs normal mice strains and to identify potential safety liabilities in the context of target engagement in a simulated disease state. Animals received 5 weekly doses, with weight loss and changes in metabolic parameters (cholesterol, triglycerides, and glucose) observed more significantly in DIO than CD-1 mice, but well-tolerated overall. In cynomolgus monkeys, a robust anti-drug antibody (ADA) response limited the ability to sustain adequate exposures over the 4 weeks, although some reductions in bodyweight were observed in the presence of sustained exposures. Due to the prevalence of ADAs and significant body weight loss in the animals with demonstrated pharmacology, a 12 week Phase 1b-enabling study in mice only (both DIO and CD-1 strains) was proposed to support the single ascending

dose trial (12 week) in patients with type 2 diabetes mellitus. However, the FDA disagreed with this proposal and suggested multiple strategies to potentially "dose through the ADA" (administration of B cell depleting antibody prior to dosing the test item, using a loading dose of the test item, etc.). As such, a 12-week study was performed in cynomolgus monkey, with dosing modifications, food supplementation, and objective assessment of maximum tolerated dose (MTD) based on weight loss and body conditioning assessments by the study veterinarian. Significant body weight loss (expected pharmacology) and ADA-related decreases in exposures limited the duration of the study, and the FDA subsequently agreed that chronic toxicity could be assessed in mouse models only (6 month duration using both CD-1 and DIO mice).

Case example 2: A platform approach to assess ADC-related toxicities. For ADCs intended for hematologic and solid malignancies, a "1 species approach" has been widely applied to support INDs and registration, alongside additional guidance within ICH S9 and ICH S9 Q&A[14] for advanced cancer products. Per these ICH guidances, the antibody portion of the ADC would normally be expected to be tested in at least 1 pharmacologically relevant binding species, where toxicity is generally considered as exaggerated pharmacology. The small molecule (payload) portion of the ADC would normally be expected to be tested in 1 rodent and/or non-rodent species. Safety assessment of the small molecule payload is essential to understand given that ADC-related toxicities are generally payload-driven (i.e., antigen-independent). In addition, the conjugate (antibody and small molecule) should be tested in 1 or 2 relevant species. At Genentech, the non-clinical development strategy is to assess the toxicity of the small molecule payload(s) and the ADC bearing the same payloads in a small number of rats early in the project; the objectives of these studies are to identify antigen-independent toxicities, to facilitate payload candidate selection, and to inform dose selection for a NHP pilot toxicity study. Often, the candidate antibodies in the ADC studies are not cross reactive in rodents, so these rodent studies facilitate the understanding of antigen-independent toxicities. After candidate selection, a pilot (non-GLP) repeat-dose (often two doses, every 3 weeks) study in cynomolgus monkeys (1 sex only if appropriate, with n = 2 or 3/group) is used to further characterize antigen-independent and antigen-dependent toxicities, in a single pilot study in NHPs using a non-binding ADC. Finally, the antigen-dependent toxicities are characterized in a 3 month IND-enabling/registration study, the longest duration required for oncology products and ICH S9[15], using the cross reactive, clinical candidate molecule. Importantly, given that the same payloads are often conjugated to different antibodies, and that the MTD is often quite consistent across different ADCs (whether cross reactive or not), the number of NHP studies for lead candidate molecules are limited by forgoing short-term toxicity studies and moving directly to a 3 month registration enabling study.

Case example 3: A therapeutic area/platform approach for mAbs/ recombinant proteins for chronic eye conditions. For multiple types of biologics (mAbs, F (ab')$_2$, bispecifics) intended for various ophthalmology indications, a "1 species approach" has been widely applied due to difficulties in relation to the intravitreal route of administration required for these studies. Although rabbits have been commonly used as a toxicology species if pharmacologically relevant as their large eye size facilitates dosing, shared industry experience indicates that the rabbit seems to be a more sensitive species to mounting a significant inflammatory reaction to humanized biologics, with high incidence and severity of intraocular inflammation that ultimately compromises the objectives of a toxicology study and impacts overall animal welfare. Immunogenicity (rapid development of ADA) is also common[15]. The cynomolgus monkey or minipig are often considered the more relevant and suitable species for ocular toxicity assessment, an approach accepted for multiple Sponsor's submissions to FDA.

Industry Perspective 2 (David Clarke, Eli Lilly and Company)

Of the mAbs in Lilly's past and present development portfolio, approximately one-third cross-react with the intended target in both NHP and rodent species, providing the basis for pharmacological relevancy of each species and their dual use in short-term FIH-enabling toxicology studies. Consideration for reducing to a single species for longer-term studies for these molecules has anchored on the comparison of toxicity profiles between the NHP and rodent species in initial toxicology studies, as originally defined by ICH S6, and subsequently, as modified within the ICH S6 (R1) addendum of 2011, on the similarity of toxicity profiles and/or whether the findings are understood from the mAb's mechanism of action. Additional factors, such as recapitulation of target signaling pathways and/or pharmacological effects, tolerability, species-specific off-target toxicity, exposure to drug, and immunogenicity in each species, as well as timing and study logistics, were also considered on a case-by-case basis (Table 3). As such, a holistic approach was taken to determine whether long-term toxicology studies in 1 or 2 species was warranted or possible to inform human risk assessment.

For this discussion, 18 mAbs that cross-react with multiple species and represent a variety of therapeutic areas were reviewed. Notably, for 16 mAbs (almost 90%), no toxicity was observed in the initial GLP toxicity studies in either rodent or NHP (cynomolgus monkey in each case; study durations were between 2 and 13 weeks), suggesting significant opportunities to reduce to a single species for longer-term toxicity studies.

Table 3. Lilly mAbs With Two Pharmacologically Relevant Species: Considerations for Species Use in Long-Term Studies.

Therapeutic Area	Short-Term GLP Tox Studies (weeks)		Additional Considerations	Longer-Term GLP Tox Studies (weeks)		Agency Input
	Species Used	Species Toxicity		Species Used	Species Toxicity	
—	Species Used	Species Toxicity	—	Species Used	Species Toxicity	—
Metabolic	NHP + rat (4)	Different	Pharmacology	NHP (39) + rat (26)	Different	Yes
CNS	NHP + Tg mouse (6)	None	Species relevance	Tg mouse (26)	None	Yes
Renal	NHP (5) + rat (12)	None	Immunogenicity	NHP (39)	New	Yes
Metabolic	NHP + rat (6)	None	Pharmacology	NHP (39) + rat (26)	None	—
CNS	NHP + rat (6)	None	Agency expectation	NHP + rat (12 then 26)	None	Yes
Metabolic	NHP + rat (2)	None	Study duration	NHP + rat (13)*	None	—
Inflammation[1]	NHP + rat (8)	None	Pharmacology	NHP (26)	None	Yes
CNS[2]	NHP + rat (13)	None	—	* (Rat)	—	Yes
CNS	NHP + Tg mouse (5)	None	Timing	NHP + Tg mouse (26)	None	—
Renal	NHP + rat (4)	None	Pharmacology	NHP + rat (13)*	None	—

Upper section groups the mAbs with studies in 2011 or earlier and lower section groups the mAbs with studies in 2012 or later (i.e., timings relative to ICH S6(R1) revision). Species Toxicity "None" refers to no toxicities identified in either species. CNS: Central Nervous System. GLP: Good Laboratory Practice. Tg: Transgenic.
*mAb stopped development.
[1]case example 1 in Industry Perspective 2 main text.
[2]case example 2 in Industry Perspective 2 main text. Of note, there were 8 additional mAbs (3 for oncology indications, 2 for CNS, and 1 for cardiovascular, metabolic or renal indications, respectively) that had used NHP+rat for short-term studies; these studies identified no toxicities in either species for 6 mAbs and different toxicities for 1 mAb; however, development was ceased or had not sufficiently progressed before making species decisions for the longer-term GLP toxicology strategy.

Several molecules were discontinued, however, and discussion was focused on 10 mAbs for which species decisions had been made for the longer-term studies: 6 prior to and 4 following the release of ICH S6 (R1); there was no clear distinction in decision-making between the 2 timeframes. Brief details for these 10 mAbs are presented in Table 3, indicating the 2 species used for initial GLP toxicity studies, comparison of species toxicity profiles and any additional considerations regarding species selection for the subsequent long-term studies.

Of the 10 mAbs, regulatory agency input was sought and received in making the decision to reduce to a single species for 6 mAbs; all 4 mAbs where this was done and for 2 of 6 mAbs where the decision was made to progress both NHP and rodent species. For the 6 mAbs where both species were retained, this was due to differences in toxicities identified in the initial studies for just 1 mAb (4 week studies; different toxicities were again observed in chronic studies) and was for other considerations for the other 5 mAbs for which no toxicities were identified in the initial studies. These considerations were: species differences in the pharmacological effect profile observed in initial studies; initial single-dose "4 week observational" studies or repeat-dose 2 week studies were deemed insufficient to make a single-species decision for 2 particular molecules; agency-expressed opinion and expectation (1 instance, prior to ICH S6 (R1)); and, an aggressive development timeline that required the conduct of initial and chronic toxicity studies in close succession with insufficient time to obtain regulatory agency feedback. Despite these considerations, the longer-term studies (13- or 26 weeks duration) for these 5 mAbs still did not reveal toxicity in either species. For the

4 mAbs where a single species was selected for long-term studies, the NHP was progressed for 2 mAbs, either because immunogenicity prevented use of the rat or because the rat was not fully pharmacologically relevant (case example 1). A rodent was progressed for the other 2 mAbs, either a transgenic mouse to evaluate a surrogate antibody given this was the more relevant model compared to NHP, or the rat because of similar toxicity profiles in rat and NHP (case example 2).

Case example 1: Partial-pharmacological relevance of rodent led to NHP-only chronic toxicity assessment. A CXCR1/2 ligands mAb being developed to treat autoimmune/inflammatory diseases exhibited high affinity (pM range) binding to, and neutralization of, all 7 human and cynomolgus monkey ELR+ chemokines and 3 out of 5 Sprague-Dawley (SD) rat ELR+ chemokines and produced demonstrable pharmacodynamic responses vis-à-vis increased plasma CXCL1 (rat) or CXCL8 (monkey). Consequently, both species were selected for FIH-enabling toxicity studies. There were no adverse effects identified in either species during these initial 8-week studies which also provided similar exposures with no significant immunogenicity. However, different clinical pathology profiles were noted for neutrophil counts which were thought to be consistent with incomplete (rat) vs. complete (monkey and human) inhibition of CXCR2 pathways. Based on the ELR+ chemokine binding/neutralization profiles and neutrophil data, the monkey was considered the most relevant species. The proposal to FDA to conduct a 6-month chronic toxicology study in the monkey only, due to these pharmacological differences, was accepted.

Case example 2: Non-adverse (similar) toxicities led to rodent-only chronic toxicology assessment. An anti-soluble amyloid-β Fab developed for Alzheimer's disease (linked to ~20 kDa polyethylene glycol (PEG) for increased half-life) exhibited binding to the same epitope across species. Tissue cross-reactivity in the cynomolgus monkey and SD rat was generally similar to human, with observed binding primarily to neural elements consistent with the expected distribution of soluble Aβ in normal tissues. Expected pharmacology of plasma Aβ accumulation was demonstrated in monkey. Consequently, both rat and monkey were selected for IND-enabling toxicity studies. No Fab-related toxicities were observed in either species during these initial 13-week studies, that also demonstrated robust exposures and increased plasma Aβ in both species. Both species also exhibited non-adverse cytoplasmic vacuolation of small magnitude in macrophages of various tissues, considered adaptive changes secondary to phagocytic uptake of PEG. However, in the rat only, adaptive changes extended to vacuolation of renal cortical tubular epithelium. The proposal to FDA to conduct a 6-month chronic toxicology study in the rat only, based on toxicological findings in both species and to a greater extent in rats, and in-line with ICH S6 (R1) recommendations, was accepted.

Regulatory Perspective 1 (Eleni Salicru, FDA)

The Pharmacology/Toxicology Coordinating Committee (PTCC) Non-clinical Biologics Subcommittee (NBSC) consists of pharmacologists, toxicologists, and microbiologists from the FDA Center for Drug Evaluation and Research (CDER) Office of New Drugs (OND) Pharmacology/Toxicology Review Divisions, the Office of Therapeutic Biologics and Biosimilars (OTBB), and the Center for Biologics Evaluation and Research (CBER). In 2020, following the NC3Rs publication evaluating the use of 2 species for longer duration toxicology testing during product development[4], a "2 vs. 1 Species Working Group" consisting of members of the PTCC NBSC was formed, with the objective to determine best practices/points to consider for deciding when to conduct longer duration safety assessment of biologics in 1 species (when 2 are relevant) or when 2 species should be maintained. This provides an opportunity to better understand when it might be feasible to reduce from 2 to 1 species during product development, in line with 3Rs principles.

Review of FDA electronic Common Technical Document (eCTD) database

An internal software tool was developed to evaluate eCTD submissions received in the FDA CDER from ~2004 to December 2019, searching the study report titles for terms including species and study duration. As many CDER IND submissions were paper-based before 2010, and all biological products were submitted to CBER prior to 2005, the CDER database may not reflect all submissions received for these years and this work should be considered a limited retrospective evaluation to illustrate representative trends for consideration and discussion. The initial eCTD search, and subsequent curations, identified 93 INDs for biologics (limited to mAbs, recombinant proteins, ADCs and "others" such as nanobodies, trivalent or bispecific mAb) that used 2 species (a rodent and a non-rodent) for initial toxicity studies and one species for longer duration toxicity studies. The 93 INDs represented 62 unique biologic products consisting of 32 recombinant proteins, 24 mAbs, 3 ADCs and 3 "other" products. For these 62 biologics, preliminary database analyses relied on information contained within FDA non-clinical reviews, Module 2 of the eCTD submissions, and the Investigator's Brochures. Based on these preliminary analyses, the primary reason for dropping to 1 species for longer duration toxicology studies were as follows:

(1) Immunogenicity (19% [12/62]), generally included hypersensitivity, anaphylaxis, and/or decreased product exposure due to ADA.

(2) No or similar toxicities (15% [9/62]), were generally based on conclusions from the Sponsor or the FDA non-clinical review that suggested that there were no adverse treatment-related toxicity or target organs of toxicity or that the treatment-related toxicity or target organs of toxicity were similar between species. Further evaluation is currently being explored.

(3) Non-relevant species (13% [8/62]), generally indicated that after further evaluation the biologic had no or minimal binding affinity and/or activity in the species.

(4) Functional Activity/Target Binding Differences (8% [5/62]), was based on either the Sponsor or the FDA non-clinical review considering that 1 species was significantly more relevant based on functional activity/target binding differences.

(5) Other (11% [7/62]), including unique route of administration (e.g., ophthalmic), low exposure observed in 1 species (not due to ADA) and physiological differences between species.

(6) Undetermined (34% [21/62]), if a primary reason for dropping from 2 to 1 species could not be determined based on preliminary review of available information. Further evaluation is needed to attempt to recategorize these biologics.

Of the 62 biologics noted above, 67% used the non-rodent species as the single species for the longer duration toxicology studies when no or similar toxicities were described. The reason for the choice of the non-rodent species in these cases is not entirely clear and therefore is an area of interest for future evaluation.

Case example 1: Immunogenicity in rodent leading to NHP-only long duration studies. This example highlights an Fc fusion protein to treat autoimmune disease for a FIH Phase 1a/b clinical study in healthy volunteers by the subcutaneous (s.c.) route of administration. Pharmacological activity was confirmed in cynomolgus monkey and SD rats and both species were used for the Sponsor's initial non-GLP 2 week pharmacokinetic and pharmacodynamic studies (intravenous [i.v.] and s.c. dose routes). As ADA developed in rat for most i.v. and all s.c. dose groups, which significantly diminished product exposures, the Sponsor submitted detailed rationale to the FDA to support the use of monkey only for the GLP studies. The Division agreed with the Sponsor that it was not useful to conduct future GLP-compliant repeat-dose toxicology studies in rats due to loss of systemic exposure of the product/pharmacodynamic effects and it was appropriate to drop from 2 to 1 species for subsequent toxicity studies due to immunogenicity.

Case example 2: Similar toxicities leading to most sensitive species (non-rodent) only longer duration studies. This example highlights a humanized mAb to treat autoimmune disease for a FIH Phase 1a/b clinical study in healthy volunteers by i.v. or s.c. route of administration. The target is conserved across humans, mice, and cynomolgus monkey (100% amino acid sequence identity) and pharmacological activity of the product was demonstrated in a mouse model of disease. Both CD-1 mice and cynomolgus monkey were used for 13 week GLP toxicity studies, where 1 non-adverse finding of undetermined relationship to the product was observed in both species (likely to be PD-related). Additionally, minor non-adverse reversible injection site inflammation and 1 adverse finding of unclear relationship to the product (most consistent with a viral etiology) were observed in the monkey study only. Exposure was maintained in all treated mice and monkeys throughout the dosing period, despite evidence of ADA formation in monkeys. The Sponsor did not formally consult the FDA but indicated in their IND submission that longer duration toxicology studies would be conducted in monkey only as there were no product-related findings unique to mice. The OND Review Division noted the Sponsor's rationale for only conducting longer duration studies in the monkey as reasonable, due to monkey demonstrating findings not seen in mouse.

The Agency encourages Sponsors to contact the appropriate OND Review Division to obtain feedback regarding decisions related to selecting 2 vs. one species for biologic products. When proposing to either maintain 2 species or drop from 2 to 1 species for longer duration toxicity studies, it is important to receive Agency feedback on the decision as well as the species selected (rodent vs. non-rodent). Communication between the Sponsor and the Agency is important to help identify opportunities to reduce animal use, based on sound scientific judgment.

Regulatory Perspective 2 (David Jones, MHRA)

Regulatory guidelines describe recommendations for studies to be performed within drug development. However, a scientific rationale should be used to explore study designs and use of methods appropriate for each individual product. Animal studies should only be conducted to evaluate safety concerns that cannot be adequately addressed in other studies by non-animal methods. The purpose of the non-clinical studies is to identify parameters for clinical monitoring of potential adverse effects and to provide adequate characterization of risk under the conditions of the clinical trial to be supported[3]. The non-clinical assessment begins in discovery with assessment of risks inherent in the pharmacological target, continues with selection of the candidate with lowest risk (already a trade off with other properties associated with efficacy) and incorporates all available information, including prior knowledge (any information on the pharmacological or chemical class available from the literature, which may include clinical information). Once data are available from initial non-clinical toxicology studies with any Investigational Medicinal Product (IMP), a Weight of Evidence approach should be conducted to decide whether longer-term studies need to be conducted in 2 species or whether 1 will suffice. The decision to drop to 1 species could be considered as low impact on non-clinical assessment of risk to humans if no effects were identified in either species or only in the species chosen to progress into the longer-term studies. Limited safety data would also generally be available from the initial human clinical trials around this time. When 2 pharmacologically relevant species have been used within initial toxicity assessments for biologics, the approach of using a single species (preferably the rodent) is recommended for progression to the longer-term toxicity studies. For biological products beyond the scope of ICH S6 (R1) such as advanced therapy medicinal products (ATMPs) and vaccines, Sponsors are very used to justifying the use of a single species. For other products including those for oncology or other life-threatening or rare diseases, oligonucleotides and chemically synthesized peptides, a single species should also be sufficient to support development. Additionally, new techniques are being developed, including human tissues/organ chips, in silico methods, new biomarkers etc. that may replace some animal data in the future, such that under certain conditions, later stage toxicology programs using single species could be considered more widely, without being detrimental to human safety.

Concluding Remarks

It is estimated that between a quarter and a third of mAbs across the industry exhibit pharmacological activity in rodent and non-rodent species, warranting the use of both species in initial and/or short-term FIH-enabling toxicology studies. Given the expected continued expansion of mAbs and other biological modalities and their use for multiple therapeutic

Table 4. Considerations for Reducing to One Species vs Maintaining Two Species in Long-Term Studies.

Results of Initial Toxicology Studies	Toxicity Profile in the Two Species	Are There Any Toxicities? Are Toxicities Different or Same/Similar Between Species? Are Any Toxicities Species-specific and Determined Not Relevant to Human?
	Nature of other findings in the 2 species (particularly when toxicities are absent)	What non-adverse findings are present? Are the non-adverse findings considered to be on- or off-target, or considered to progress to toxicity in the longer-term study? Are the non-adverse findings different or same/similar between species? Are any findings species-specific and determined not relevant to human?
	Exposure and immunogenicity in the 2 species	Is the molecule tolerated at doses that provide adequate exposure to characterize toxicity? Is immunogenicity manageable or considered unduly challenging to manage in the longer-term study?
Species biology and pharmacology	Biological relevance of each species	How well does or do the target(s) expression and distribution compare to humans? Has downstream signaling been demonstrated?
	Pharmacological relevance of each species	Is the molecule-to-target binding affinity adequate? Has pharmacological activity been demonstrated?
Target and Indication	Mechanism of action	Is it novel or not? Are there unrealized yet still expected or theoretical safety concerns?
	Clinical population	What is the overall benefit : risk for the diseased population(s)?
Logistics	Integrated toxicology strategy and development timeline	Was the design or duration of initial toxicology studies adequate to provide a sufficiently robust assessment of species comparison? Does the clinical development strategy allow the opportunity to seek regulatory agency input on species use for longer-term studies?
	Regulatory agency advice	Has an agency advised appropriately on species use for longer-term studies?

indications, this represents a significant proportion of biologics within development where there may be opportunities to reduce animal use for longer-term studies. In addition, modalities such as oligonucleotides and synthetic peptides feature properties of both new chemical entities and biologicals. Many of these products currently follow the small molecule (ICH M3 (R2)) approach and maintain use of 2 species for the long-term studies. Oligonucleotide and synthetic peptide products may therefore offer an opportunity for future study by industry and regulators. As described in this manuscript for biologics, areas of assessment and discussion for these modalities might include rationale for species selection, pharmacological relevance, cross-species comparison of toxicity profiles, and potential identification of opportunities to reduce animal use in long-term toxicity studies.

Although it is outlined within ICH S6 (R1) that the rodent should be selected as the single species for long-term studies where possible and scientifically justified, the data from industry does not reflect that this is happening in practice and a high proportion of the long-term studies are conducted using NHPs. The scientific rationale for this is often immunogenicity in rodent, or "partial" pharmacological relevance leading to initial investigation of rodents in short-term studies but selection of the "more" pharmacologically relevant NHP for the longer-term studies. Given that the majority of mAb and other biological programs run entirely as single-species (NHP)

packages when there is only 1 pharmacologically relevant species, there may also be an aspect of "standard practice" within some of the decisions to continue using NHP even if rodent was an option. It is almost a decade since the ICH S6 (R1) guidance was issued and 3Rs principles, including minimizing NHP use, are high considerations for species selection for both regulators and industry. As more biologic products are licensed using long-term rodent data, and as more case-studies and discussions are shared, it should highlight rodent as equally applicable for longer-term toxicity risk assessments and may encourage more rodent-only long-term studies in the future.

As the experiences presented in this article illustrate, a similar toxicity profile (or absence of toxicity) in 2 species in the initial toxicology studies is not the only factor that can determine whether or not to proceed with 1 or 2 species for longer-term studies. Table 4 attempts to capture the various considerations, which may be multifactorial for any given molecule. Evaluation of species responses in the initial studies (to indicate similar or different responses) should consider the profile and nature of both adverse and non-adverse findings and assessment of whether the effects are on-vs. off-target or species-specific. The oftentimes related outcomes of immunogenicity, exposure to active drug and resulting margins of safety from the 2 species are also considerations. As previously noted, even though selection of 2 species for early

toxicity testing is with the expectation that both the rodent and non-rodent are pharmacologically relevant, there can be different degrees of pharmacological relevance; this could be the extent and distribution of target expression (exemplified by Lilly case 1) or evidence or not of downstream physiological response through target signaling in each species, as well as the affinity of the antibody to the target and its ability to produce an intended pharmacodynamic response or pharmacologic activity. Target considerations can also include the mechanism of action, especially how well-understood this is and any expected or theoretical concerns that need to be investigated and warrant 1 or both species to do so. The disease state, whether it is life-threatening or not, and overall benefit-risk to the clinical population could be another potential consideration. Finally, "logistical" considerations may include feedback sought from Regulatory Agencies or absent opportunity to do so given the timing of longer-term toxicology studies to support clinical development plans. Taking this holistic approach to decide on species use for longer term studies, it is ultimately the level of confidence that one has in the data from the initial toxicology studies from each species to either enable a single species as being *sufficient* to provide

information for human risk assessment or indicate that 2 species are still warranted.

It is clear that there are often more opportunities to reduce the number of species used in long-term studies than are currently applied in practice. Definition and harmonization around the criteria for "similarity" of toxicity profiles (including when no toxicities are identified) may support wider adoption of the flexibility within the ICH S6 (R1) guidelines but is not the only consideration when selecting the species to progress. Sharing of experience to promote the scenarios when this approach can be applicable provides useful information to encourage the wider adoption of single species testing for long-term toxicity studies for biologics, and potentially other drug modalities in the future if additional evidence were collected. This provides efficiencies within drug development by reducing animal use as well as reducing costs (1 fewer study, less compound required, etc.), with greater gains if the rodent is selected to progress in preference to NHPs. With the additional current shortage of NHPs available for research and toxicity testing, arguably this could be the optimal time to adopt the approach and to encourage use of a single species, *preferably* the rodent, where scientifically justified for long-term toxicity testing.

APPENDIX

Abbreviations

(ADCs);	Antibody drug conjugates
(ADA);	Anti-drug antibody
(ICH);	International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use
(mAb);	Monoclonal antibody
(NHP)	Non-human primate

Author's Contributions

Prior, H. contributed to conception and design, contributed to acquisition, analysis, and interpretation, and drafted manuscript; Clarke, D. contributed to conception and design and drafted manuscript; Jones, D. contributed to conception and critically revised manuscript; Salicru, E. contributed to acquisition, analysis, and interpretation and drafted manuscript; Schutten, M. contributed to analysis and interpretation and drafted manuscript; Sewell, F. contributed to interpretation and critically revised manuscript. All authors gave final approval and agree to be accountable for all aspects of work ensuring integrity and accuracy.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

ORCID iDs

Helen Prior https://orcid.org/0000-0002-8700-7226
Eleni Salicru https://orcid.org/0000-0002-6498-0824
Fiona Sewell https://orcid.org/0000-0001-5589-9453

References

1. Mullard A. 2020 FDA drug approvals. *Nat Rev Drug Discov.* 2021;20:85-90.

2. ICH S6(R1). Preclinical safety evaluation of biotechnology-derived pharmaceuticals. In: *International Conference on Harmonisation (ICH)*; 2011. Topic S6(R1).

3. ICH M3(R2). Nonclinical safety studies for the conduct of human clinical trials and marketing authorization for pharmaceuticals. In: *International Conference on Harmonisation (ICH)*; 2009. Topic M3(R2).

4. Prior H, Baldrick P, Beken S, et al. Opportunities for use of one species for longer-term toxicology testing during drug development: A cross-industry evaluation. *Regul Toxicol Pharmacol.* 2020;113:104624.

5. Iwasaki K, Uno Y, Utoh M, Yamazaki H. Importance of cynomolgus monkeys in development of monoclonal antibody drugs. *Drug Metabol Pharmacokinet.* 2019;34:55-63.

6. Prior H, Haworth R, Labram B, Roberts R, Wolfreys A, Sewell F. Justification for species selection for pharmaceutical toxicity studies. *ToxicoRes.* 2020b;9:758-770.

7. Sewell F, Chapman K, Couch J, et al. Challenges and opportunities for the future of monoclonal antibody development: Improving safety assessment and reducing animal use. *mAbs.* 2017;9:742-755.

8. Brennan FR, Cavagnaro J, McKeever K, et al. Safety testing of monoclonal antibodies in non-human primates: Case studies highlighting their impact on human risk assessment. *mAbs.* 2018;10:1-17.

9. Lee DW, Zhong S, Pai R, et al. Nonclinical safety assessment of a human interleukin-22 FC IG fusion protein demonstrates in vitro to in vivo and cross-species translatability. *Pharmaco Res Perspectiv.* 2018;6:e00434.

10. Welsh BT, Faucette R, Bilic S, et al. Nonclinical development of SRK-181: An Anti-Latent TGFβ1 monoclonal antibody for the treatment of locally advanced or metastatic solid tumors. *Int J Toxicol.* 2021;40:226-241.

11. Lansita JA, Mease KM, Qiu H, Yednock T, Sankaranarayanan S, Kramer S. Nonclinical development of ANX005: A humanized anti-C1q antibody for treatment of autoimmune and neurodegenerative diseases. *Int J Toxicol.* 2017; 36:449-462.

12. van Meer P, van Aerts L, Weir L, et al. Optimal duration of safety studies with monoclonal antibodies. *ALTEX Proceedings.* 2021; 9(1):88.

13. Sonoda J, Chen MZ, Baruch A. FGF21-receptor agonists: An emerging therapeutic class for obesity-related diseases. *Horm Mol Biol Clin Invest.* 2017;30:20170002.

14. ICH S9. Nonclinical evaluation for anticancer pharmaceuticals. In: *International Conference on Harmonisation (ICH)*. Topic S9; 2010.

15. de Zafra CLZ, Sasseville VG, Matsumoto S, et al. Inflammation and immunogenicity limit the utility of the rabbit as a nonclinical species for ocular biologic therapeutics. *Regul Toxicol Pharmacol.* 2017;86:221-230.