Regulatory Toxicology and Pharmacology 113 (2020) 104624

Contents lists available at ScienceDirect

Regulatory Toxicology and Pharmacology

journal homepage: www.elsevier.com/locate/yrtph

Opportunities for use of one species for longer-term toxicology testing during drug development: A cross-industry evaluation

Helen Prior[a,*], Paul Baldrick[b], Sonja Beken[c], Helen Booler[d], Nancy Bower[e], Paul Brooker[f], Paul Brown[g], Brian Burlinson[h], Leigh Ann Burns-Naas[i], Warren Casey[j], Melissa Chapman[k], David Clarke[l], Lolke de Haan[m], Olaf Doehr[n], Noel Downes[o], Meghan Flaherty[p], Nichola Gellatly[a], Sophia Gry Moesgaard[q], Jennifer Harris[r], Mark Holbrook[s], Julia Hui[t], David Jones[u], Keith Jones[v], Hilla Kedar[w], Andreas Mahl[x], Alli Manninen[y], Aidan McGuire[z], Elisabeth Mortimer-Cassen[aa], Marjorie Peraza[ab], Michael K. Pugsley[ac], Jacques Richard[ad], Ruth Roberts[ae], Wendy Roosen[af], Andreas Rothfuss[ag], Ankie Schoenmakers[ah], Fiona Sewell[a], Richard Weaver[ai], Lucinda Weir[aj], Alison Wolfreys[ak], Ian Kimber[al]

[a] National Centre for the Replacement, Refinement and Reduction of Animals in Research (NC3Rs), Gibbs Building, 215 Euston Road, London, NW1 2BE, UK
[b] Covance Laboratories Ltd, Otley Road, Harrogate, HG3 1PY, UK
[c] Federal Agency for Medicines and Health Products (FAMHP), Victor Hortaplace 40/40, Brussels, 1060, Belgium
[d] Genentech, Inc, 1 DNA Way, South San Francisco, CA, 94080, USA
[e] Eisai Inc, 155 Tice Blvd, Woodcliff Lake, NJ, 07677, USA
[f] Board member, NC3Rs, Gibbs Building, 215 Euston Road, London, NW1 2BE, UK
[g] Food and Drug Administration (FDA), 10903 New Hampshire Avenue, Silver Spring, MD, 20993, USA
[h] Covance, Huntingdon, Cambridgeshire, PE28 4HS, UK
[i] Gilead Sciences, Inc, 333 Lakeside Drive, Foster City, CA, 94404, USA
[j] National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM), National Institute of Environmental Health Sciences, P.O. Box 12233, Research Triangle Park, NC, 27709, USA
[k] Oncology Safety, Clinical Pharmacology and Safety Sciences,R&D, AstraZeneca, Cambridge, UK
[l] Lilly Research Laboratories, Indianapolis, IN, 46285, USA
[m] Biologics and Advanced Therapeutics Safety, Clinical Pharmacology and Safety Sciences, R&D, AstraZeneca, Cambridge, UK
[n] Bayer Pharma AG, Müllerstrasse 170, 13353, Berlin, Germany
[o] Sequani Limited, Bromyard Rd, Ledbury, Herefordshire, HR8 1LH, UK
[p] Takeda Pharmaceuticals, 300 Massachusetts Ave, Cambridge, MA, 02139, USA
[q] Novo Nordisk A/S, Novo Nordisk Park, Maaloev, Denmark
[r] Association of British Pharmaceutical Industry (ABPI), 105 Victoria Street, London, SW1E 6QT, UK
[s] VAST Pharma Solutions Ltd, Harrogate, UK
[t] Celgene, 86 Morris Avenue, Summit, NJ, 07901, USA
[u] Medicines Healthcare products Regulatory Agency (MHRA) 10 South Colonnade, Canary Wharf, London, E14 4PU, UK
[v] CFMD Ltd, Peterborough, UK
[w] Teva Pharmaceuticals, Netanya Area, Israel
[x] Novartis Institutes for BioMedical Research (NIBR), Basel, Switzerland
[y] Prekliniska Byran Sweden AB, Stockholm, Sweden
[z] Charles River Laboratories, Preclinical Services, Tranent, Edinburgh, EH33 2NE, UK
[aa] Regulatory Safety Centre of Excellence, Clinical Pharmacology and Safety Sciences, R&D, AstraZeneca, Cambridge, UK
[ab] Pfizer Drug Safety Research and Development, 300 Technology Square, Cambridge, MA, 02139, USA
[ac] Cytokinetics, South San Francisco, CA, 94080, USA
[ad] Sanofi, 371 Rue du Professeur Blayac, Montpellier, 34000, France
[ae] ApconiX, Alderley Park, Alderley Edge, Cheshire, SK10 4TG, UK
[af] Janssen Research & Development, Turnhoutseweg 30, 2340, Beerse, Belgium
[ag] Roche Pharmaceutical Research and Early Development, Roche Innovation Center Basel, CH - 4070, Basel, Switzerland
[ah] Charles River Laboratories, Den Bosch B.V, the Netherlands
[ai] Institut de Recherches Internationales Servier, Biopharmacy, 92284, Suresnes, Cedex, France

Abbreviations: ADC, Antibody-drug conjugate; ABPI, Association of the British Pharmaceutical Industry; CRO, contract research organisation; EU, European Union; FIH, first-in-human; GLP, Good Laboratory Practice; ICH, International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use; IND, Investigational New Drug; mAbs, monoclonal antibodies; NC3Rs, National Centre for the Replacement, Refinement and Reduction of Animals in Research; NHP, non-human primate

* Corresponding author. Gibbs Building, 215 Euston Rd, London, NW1 2BE, UK.
E-mail address: helen.prior@nc3rs.org.uk (H. Prior).

https://doi.org/10.1016/j.yrtph.2020.104624

Available online 29 February 2020

[aj] GlaxoSmithKline, Park Road, Ware, Hertfordshire, SG12 0DP, UK
[ak] UCB Biopharma, Bath Rd, Slough, SL1 3WE, UK
[al] University of Manchester, Faculty of Biology, Medicine and Health, Oxford Rd, Manchester, M13 9PL, UK

ARTICLE INFO

Keywords:
3Rs
Biologics
Drug development
ICHS6(R1)
Nonclinical
Non-rodent
Rodent
Small molecules
Species selection
Toxicology

ABSTRACT

An international expert working group representing 37 organisations (pharmaceutical/biotechnology companies, contract research organisations, academic institutions and regulatory bodies) collaborated in a data sharing exercise to evaluate the utility of two species within regulatory general toxicology studies. Anonymised data on 172 drug candidates (92 small molecules, 46 monoclonal antibodies, 15 recombinant proteins, 13 synthetic peptides and 6 antibody-drug conjugates) were submitted by 18 organisations. The use of one or two species across molecule types, the frequency for reduction to a single species within the package of general toxicology studies, and a comparison of target organ toxicities identified in each species in both short and longer-term studies were determined. Reduction to a single species for longer-term toxicity studies, as used for the development of biologicals (ICHS6(R1) guideline) was only applied for 8/133 drug candidates, but might have been possible for more, regardless of drug modality, as similar target organ toxicity profiles were identified in the short-term studies. However, definition and harmonisation around the criteria for similarity of toxicity profiles is needed to enable wider consideration of these principles. Analysis of a more robust dataset would be required to provide clear, evidence-based recommendations for expansion of these principles to small molecules or other modalities where two species toxicity testing is currently recommended.

1. Introduction

The biopharmaceutical industry recognises the importance of an ongoing review of the regulatory recommendations for the nonclinical safety assessment of new investigative drugs. There is a responsibility to ensure human safety and that individual nonclinical studies and overall safety strategies reflect advances in the scientific field, such as employing new technologies (e.g., *in silico* and *in vitro* approaches), and the growing application of mechanistic toxicology to understand molecular pathways underlying toxic effects. This can also lead to opportunities to replace, refine and reduce (the 3Rs) the use of laboratory animals when appropriate. Review of best practices and identification of new and/or different approaches to safety assessment can enhance predictivity, as well as reduce animal use through efficiency improvements or streamlined processes. This may include a requirement for fewer studies, fewer animals per study or shortened timelines to reach decision-making milestones. Arguably, such reviews are conducted most effectively when pharmaceutical companies work together to share data and experience, in concert with regulatory authorities, for example as contributors to working parties and industry consortia. Recently, consortia have reviewed topics such as nonclinical to clinical data translation (Monticello et al., 2017), the effectiveness of nonclinical strategies in support of safe clinical trials (Butler et al., 2017), *in silico* prediction of toxicities (Briggs et al., 2015) and the most appropriate

Table 1
Relevant extracts from ICH general guidance regarding the species required for general toxicology studies.

ICHM3(R2): Nonclinical Safety Studies for the Conduct of Human Clinical Trials and Marketing Authorization for Pharmaceuticals

5. Repeated dose toxicity studies. In principle, the duration of the animal toxicity studies conducted in two mammalian species (one non-rodent) should be equal to or exceed the duration of the human clinical trials up to the maximum recommended duration of the repeated-dose toxicity studies.

Recommended Non-clinical Studies to Support Exploratory Clinical Trials. Approaches 1 and 2 describe requirements for a toxicity study in one species, usually rodent

ICHS6(R1): Pre-clinical Safety Evaluation of Biotechnology-Derived Pharmaceuticals

3.3 Animal Species/Model Selection. Safety evaluation programs should include the use of relevant species. A relevant species is one in which the test material is pharmacologically active due to the expression of the receptor or an epitope (in the case of monoclonal antibodies). Safety evaluation programs should normally include two relevant species. However, in certain justified cases one relevant species may suffice (e.g., when only one relevant species can be identified or where the biological activity of the biopharmaceutical is well understood). In addition, even where two species may be necessary to characterise toxicity in short-term studies, it may be possible to justify the use of only one species for subsequent long-term toxicity studies (e.g., if the toxicity profile in the two species is comparable in the short-term).

Addendum 2.1 General Principles. For monoclonal antibodies and other related antibody products directed at foreign targets (i.e., bacterial, viral targets etc.), a short-term safety study in one species (choice of species to be justified by the sponsor) can be considered.

Addendum 2.2 One or Two Species. If there are two pharmacologically relevant species for the clinical candidate (one rodent and one non-rodent), then both species should be used for short-term (up to 1 month duration) general toxicology studies. If the toxicological findings from these studies are similar or the findings are understood from the mechanism of action of the product, then longer-term general toxicity studies in one species are usually considered sufficient. The rodent species should be considered unless there is a scientific rationale for using non-rodents. Studies in two non-rodent species are not appropriate. The use of one species for all general toxicity studies is justified when the clinical candidate is pharmacologically active in only one species.

Addendum Note 2. If two species have been used to assess the safety of the ADC, an additional short-term study or arm in a short-term study should be conducted in at least one species with the unconjugated toxin. In these cases, a rodent is preferred unless the toxin is not active in the rodent. If only one pharmacologically relevant species is available, then the ADC should be tested in this species.

ICHS9: Nonclinical Evaluation for Anti-Cancer Pharmaceuticals

2.4 General Toxicology. For small molecules, the general toxicology testing usually includes rodents and non-rodents. In certain circumstances, determined case-by-case, alternative approaches can be appropriate (e.g., for genotoxic drugs targeting rapidly dividing cells, a repeat-dose toxicity study in one rodent species might be considered sufficient, provided the rodent is a relevant species).

Q&A Other considerations (applicable to ADCs).

4.3: *Are studies with the payload and/or linker only recommended?* If the toxicity of the payload or payload with linker has not been characterized, the payload or payload with linker could be evaluated in one species as a stand-alone study or could be added as an arm into toxicology studies of the ADC.

4.8: *If the ADC does not bind the target in the nonclinical species, what repeat dose in vivo toxicity study would be needed?* If the epitope is not present in nonclinical test species, a toxicology study in one species for the ADC should be sufficient.

4.10: *Generally, two species are used for toxicology testing. For an ADC, are there situations where one species may be acceptable?* When the antibody portion of an ADC binds only to human and NHP antigens, conducting a toxicity evaluation with the ADC in only the NHP (the only relevant species) would be appropriate, as discussed in ICH S6 (R1). For the payload, see the response to Question 4.3.

use of animals to assess toxicity of novel drugs and biopharmaceuticals (Brennan et al., 2018; Sewell et al., 2017).

In 2016, the UK National Centre for the Replacement, Refinement and Reduction of Animals in Research (NC3Rs) and member representatives of the Association of the British Pharmaceutical Industry (ABPI) formed an international working group to retrospectively review the utility of toxicology testing in two species for general toxicology studies, and specifically to investigate whether, in certain circumstances, data derived from a single species would be sufficient for safe progression of a drug candidate through clinical testing. The background to the project was described in detail previously (Prior et al., 2018), with the main focus being to explore whether both a rodent and a non-rodent species are necessary for general toxicology testing, as currently described by regulatory guidance. Initial questions were aligned around whether existing opportunities to use or reduce to a single species at different stages of drug development are being fully exploited by the pharmaceutical industry and supported by regulatory authorities and/or whether these opportunities could be expanded into areas where this is currently not accepted.

1.1. The use of one or two species for toxicity assessment

The number of species used for regulatory toxicity testing of new medicinal products is based on International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) guidance (see Table 1, ICHM3R2, 2009; ICHS6R1, 2011ICHS9, 2010), applied appropriately for specific drug classes and intended therapeutic applications. For small molecules and other drug modalities developed according to ICHM3(R2), and drug candidates for oncology indications for which ICHS9 guidelines apply, studies using two mammalian species (i.e., a rodent plus a non-rodent species) are usually recommended to support clinical development and licensing. Both guidelines also outline circumstances when a case-by-case approach might be warranted to expedite drug development; in particular a repeat-dose toxicity study in a single pharmacologically relevant (rodent) species for genotoxic drugs targeting rapidly dividing cells (ICHS9). Regulatory toxicology studies for biotechnology-derived pharmaceuticals (henceforth referred to as 'biologics') following the ICHS6(R1) guideline such as monoclonal antibodies (mAbs), should only be conducted in pharmacologically relevant species, since toxicity is generally driven by exaggerated pharmacology and/or immunogenicity (Baldrick, 2011; Brennan et al., 2018). Due to the high selectivity of biologics, often only one pharmacologically relevant species can be identified. This is frequently the non-human primate (NHP), due to the much higher sequence homology between the human and NHP proteome. Consequently, single species toxicology packages for biologics are relatively common and widely accepted. Nevertheless, if multiple pharmacologically-relevant species are identified for a biologic, testing in two species (rodent and non-rodent) is a regulatory recommendation. Should the toxicology profile in short-term studies (e.g., studies of up to one month dosing duration, to support phase I clinical trials) in both species be similar, then longer-term toxicity studies (e.g., studies of up to six months dosing duration) in a single species, preferably rodent, are usually considered sufficient to support phase II/III clinical trials (ICHS6(R1), 2011).

Over the recent decade(s), as technological advances have led to different approaches and therapies, the diversity of investigative drug modalities has increased dramatically beyond small molecules and first generation mAbs. For biologics such as antibody-drug conjugates (ADCs) (Beck et al., 2017), advanced therapy medicinal products (ATMPs) (Boran et al., 2017) and biosimilar mAbs (Stevenson et al., 2017), a case-by-case approach to nonclinical testing is appropriate, in line with ICHS6(R1) and other regulatory guidelines (Detela and Lodge, 2019; EMA, 2014; EMA, 2018; ICHS9, 2010). Interestingly, testing of ADCs may involve a two species approach for acute or short-term studies, regardless of the pharmacological relevance of the rodent species (usually rat) (Hinrichs and Dixit, 2015). The rat may be used to screen for the off-target tolerability of the small molecule component of the ADC, whilst a pharmacologically relevant non-rodent species (usually NHP) is used for assessment of on-target as well as off-target toxicities. If both rodent and non-rodent species are pharmacologically relevant, toxicity studies in two species are recommended in accordance with ICHS6(R1). Other drug types, such as oligonucleotides and peptides, are chemically synthesised but can be highly species selective, and therefore pharmacological relevance of the toxicology species selected is of critical importance. Whilst there is no separate guidance for these products, the nonclinical development path is largely consistent with that of small molecules (Mustonen et al., 2017). Case-by-case approaches outside of established guideline recommendations, or where guidelines are unavailable, are possible following prospective discussions with regulatory agencies on the nonclinical programme designs, dependent on the therapy area and scientific justification. There are a limited number of published examples of nonclinical programmes with no animal toxicology data, particularly for anti-cancer immunotherapies (Brizmohun, 2019; English, 2011), where no relevant toxicology species were identified.

Current drug development practices and regulations are intended to ensure the most appropriate nonclinical approaches are taken for a new investigative drug. The established regulatory framework aims to provide a high level of human (either healthy volunteer or patient) safety, based on preclinical toxicology data which may involve use of two, one or even no nonclinical species (Butler et al., 2017). Bearing in mind the large number of clinical trials performed each year in the US and Europe alone, serious adverse events (debilitating or requiring hospitalisation) in phase I trials are very low, in 0–0.31% of subjects treated with new investigational drugs (Emanuel et al., 2015). Nevertheless, the recent trial with BIAL 10–2474, a long acting fatty acid acyl hydrolase (FAAH) inhibitor (Chaikin, 2017) illustrates that there can still be a risk of severe adverse events occurring, despite extensive nonclinical studies. Although the reasons for this are not fully understood, the predictivity of nonclinical toxicology data for assessing risk to human safety is one factor of paramount importance, and the value of toxicity assessments using animal models is a matter of debate (Bailey et al., 2015; Mangipudy et al., 2014; van Meer et al., 2012). Other industry consortia have retrospectively investigated the concordance between animal toxicity data and human adverse effects. They concluded that when adverse effects were identified in humans, these were also identified in animals for between 48% (Tamaki et al., 2013) and 71% (Olson et al., 2000) of the dataset. More recently, a separate project has emphasised positive predictive values (PPV, the proportion of positive nonclinical findings that had positive clinical findings) as more relevant for prospective nonclinical to clinical translation (Monticello et al., 2017). For a dataset mainly consisting of small molecules (85%), mAbs (9%) and ADCs (3%), they found a PPV of 43%, which increased when the same target organ was identified in both the rodent and non-rodent toxicology species. Conversely, the absence of target organ toxicities in either test species strongly predicted a similar outcome in the clinic (a negative predictive value of 86%). It should be noted that there was no intention to assess predictivity of animal toxicity data within this project, since many of the member organisations of this working group were already members of the other consortia investigating this further (Monticello et al., 2017).

Historically, human risk assessment based on nonclinical toxicology data has mostly involved two species (rodent and non-rodent) toxicology packages, given that small molecules dominated the drug development landscape. With the advent of biologics, human risk assessment based on nonclinical data packages from only one (or occasionally no) animal species have become much more common. Whilst there are published case-studies of biologics that are cross-reactive in two species (Baldrick, 2017; Brennan et al., 2018; Sewell et al., 2017), there are no published data on the actual prevalence of two *versus* one species toxicology programmes, the potential benefit of two

Table 2a
Target organ categories, as defined within the questionnaire (drop-down menus).

Which target organ systems were affected?		
Adrenal glands	Bile ducts & liver	Clinical chemistry
Clinical signs (please describe)	Endocrine (thyroid/ pancreas)	Eyes and optic nerve
Female reproductive organs	GI tract/stomach/ oesophagus	Haematology
Heart or vascular tissue	Immune system#	Kidneys & ureters
Lungs & respiratory system	Male reproductive organs	Skeleton
Skin	Other (please describe)	

to include effects on the thymus, spleen, lymph nodes and bone marrow. Multiple option choices were allowed within these questions.

species toxicology packages or the predictivity of toxicology testing in two species for biologics across the pharmaceutical industry. Nor are there any published data on the incidence, potential impact and regulatory acceptance of reducing the toxicology programme from two to one species with a biologic when toxicity profiles in short-term studies are similar, or for genotoxic anti-cancer drugs conducting a repeat-dose toxicity study in a single rodent species. The definition of what a 'similar toxicity profile' in two species actually means remains vague (see Table 1) and there is a lack of clarity on how to apply this in practice. In particular, in ICHS6(R1) it is unclear whether or not the absence of toxicity in two species constitutes a similar toxicity profile. As a consequence, Sponsors may continue to use two species to avoid regulatory risk and potential delays in development timelines.

1.2. The NC3Rs/ABPI working group objectives

The NC3Rs/ABPI working group comprised experts in nonclinical safety assessment, representing 25 global pharmaceutical and biotechnology companies, six contract research organisations (CROs) or consultants, two academic institutions and four regulatory bodies and was ledby the NC3Rs. The group shared pre-registration, non-public domain data on rodent and non-rodent species used for toxicity assessments with any drug candidate in current or former portfolios. Information was also collected on the key toxicities identified in each species within studies at different phases of drug development, both in terms of the specific target organs affected and the impact of the data on progression of the drug candidate. These data were collected with the aim of investigating the incidence in similarity of toxicities between species, in order to perform retrospective assessments as to whether there had been potential opportunities to reduce from two to one species at later stages of development. An important objective was to assess whether existing opportunities to use a single specieswere being employed and also to assess if there were wider opportunities, regardless of the drug modality and independent of the current applicable guidelines.

2. Methods

2.1. Data collection

The working group designed a questionnaire to collect information on studies conducted to support safety assessment of any class of molecule within company portfolios, focussing on the impact/value of data for decision making from each species used. Following a pilot phase, the questionnaire was revised, mainly with clarifications to improve interpretation of questions and consistency of responses.

The questionnaire consisted of four main sections:

1) General information on the drug candidate, including modality (e.g. small molecule, mAb etc.), therapeutic area, current phase of development (if still in active development), or furthest nonclinical or clinical phase reached (if no longer in development), and regulatory guidelines followed. No data were collected that would allow identification of the drug candidate (such as chemical names or structures).
2) Information on the species used for general toxicity testing at any time during discovery and development of the drug candidate, including justification of species choice. The default NHP was the cynomolgus macaque, with other species noted where appropriate. Where general toxicity studies in rabbit were reported, this was in conjunction with NHP, for ocular/ophthalmology indications (and intravitreal dosing route). Those submitting this data confirmed the rabbit was used as a relevant 'small animal' species, in place of a rodent. As such, rabbit data and result interpretation have been included in association with rodent data (e.g., Fig. 2b).
3) Information on the toxicity studies in each species, including details of study design (e.g. study type, Good Laboratory Practice (GLP) compliance status, dosing duration, route of administration and the total number of animals used), and summarised (high level) study results. High level study results included the presence or absence of any target organ toxicities; the impact this had on progression of the drug candidate (through internal decision making) was also noted. The list of drop-down options for the target organ toxicities and the impact of data from the study are described in Tables 2a and 2b. Information on the nature, severity or incidence of toxicologically-relevant findings, achieved exposures or safety margins were not requested.
4) Several "retrospective", and potentially subjective, questions about the package of toxicity studies conducted were also included. These were a) 'If you used two species, with the benefit of hindsight, would you have been able to make the same decisions based on the data from one species only?' b) 'If you used two species, did you reduce to a single species at any point in the package?' and c) 'If you used two species, what additional value did data from the second species give you?'. The question regarding decisions made in hindsight was intended to be answered overlooking current regulatory recommendations i.e. the personal opinion of the individual/s completing the questionnaire. Survey respondents had the opportunity to provide additional information via free-text comment boxes.

2.2. Data inclusion criteria

To minimise the potential for selection bias, respondents were asked to start with their most recent molecule (regardless of stage of development) from mid-2017 (when the data were collected) working back to January 2012, to ensure the dataset reflected current guideline

Table 2b
Impact of data from studies, as defined within the questionnaire (drop-down menus).

What was the impact of data from this study?
Identified toxicities were the primary reason to stop progression
Data contributed to weight-of-evidence to stop progression
Lack of effects in this species gave confidence to progress
Identification of NOAEL in this species gave confidence to progress
Identified toxicities considered acceptable/monitorable for expected indication
Lack of translatability to human (species-specific effect)
Molecule progressed to next phase of animal studies
Contributed to the decision to progress in humans
FIH starting dose was based primarily on results from this study
FIH ceiling dose/exposure cap was based primarily on results from this study
FIH escalation frequency/magnitude was based primarily on results from this study
FIH additional clinical monitoring was included based on results from this study
Phase II doses were based primarily on results from this study
Phase II clinical monitoring was included based on results from this study
Other (please describe further):

Multiple option choices were allowed within these questions.

recommendations (e.g., to include ICHM3 and ICHS6 revisions). Companies were requested to provide information on at least five drug candidates (ideally), or more if available. Criteria for inclusion in the survey required a completed (reported) package of studies at one of these phases of development:

1) Pre-FIH: drug candidates having completed early toxicology studies which may be performed to assist molecule selection and provide pilot information for the pivotal GLP-compliant studies;
2) FIH: drug candidates having completed the pivotal (generally GLP-compliant) toxicology studies intended to support a phase I clinical trial;
3) Post-FIH: drug candidates having completed pivotal GLP-compliant longer-term dosing toxicology studies intended to support longer duration clinical trials. The study dosing duration within each of the phases above differed for individual drug candidates, as appropriate for each individual programme and drug modality. For example, a 4-week study may have been conducted to support a small molecule for the FIH phase and a 13 or 26-week study for post-FIH, whilst a 13-week study may have been conducted for a mAb for FIH phase and a 26-week study for post-FIH.

The studies suitable for inclusion in the questionnaire were restricted by the working group to general toxicology studies only (i.e. safety pharmacology, reproductive toxicity, genotoxicity or carcinogenicity studies were out of scope).

Completed questionnaires were submitted by individual companies and data were collated and anonymised by NC3Rs before being analysed by the working group. Any data provided by CROs/consultants were provided (with permission) from Sponsors not otherwise involved in the working group (to avoid duplication). Drug candidates were categorised by molecule type and only included in the dataset if at least three different companies submitted data on the same drug modality.

2.3. Target organ toxicities

For each molecule, the absence or presence of target organ toxicities in each study, species and at each phase, were noted. If multiple studies were performed for the same species within the same phase, and different organs were affected between studies, the total of all target organ toxicities were recorded for that phase.

For the drug candidates that were tested in two species (rodent and non-rodent) during the pre-FIH and FIH phase (115 of the 172 drug candidates), the target organ toxicities were compared between species. Only the presence or absence of target organ toxicities was collated, there was no information on nature, severity or incidence of toxicities, nor achieved exposures or safety margins. Likewise, there was no assessment of the relative importance of a toxicity in relation to other target organs for decision-making. Compared toxicities were classified into one of four categories:

1) None: no target organ toxicities identified in both species;
2) Same: toxicities identified in the same target organs in both species;
3) Similar: toxicities identified primarily in the same target organs (only one additional or different target organ toxicity was identified in one of the species);
4) Different: toxicities identified in both species; this could either be a) no target organ toxicities identified in one species and one or more target organs identified in the other, or b) multiple toxicities identified in each species with more than one target organ toxicity different between the two species.

For drug candidates with target organ toxicities categorised as similar or different, the species with the highest number of target organ toxicities was also noted (i.e., rodent, non-rodent or both species if the same number of toxicities were identified, with unique target organs in each species).

2.4. A blinded exercise to investigate the potential to conduct post-FIH studies in a single nonclinical species

An exercise was conducted to test the hypothesis that an assessment of the toxicities observed in two species during the FIH studies (data from the pre-FIH and FIH phases) could, in hindsight, inform on the requirement for two species toxicology testing for the longer-term post-FIH toxicity studies, or if reduction to a single species would have been acceptable without increasing risk of missing toxicity findings within the nonclinical safety decision-making. Panels of six to eight working group members reviewed results from the pre-FIH and FIH phase animal studies for drug candidates that had been tested in two species and had progressed to longer-term post-FIH studies (blinded to the latter data). This included the target organ toxicities identified in each of the

Table 3
Blinded exercise: criteria for the potential to reduce to one species or to retain two species for the longer-term post-FIH toxicity studies.

Working group decision	Comparisons with the actual outcomes of the post-FIH toxicity studies		
Decision from blinded exercise[#]	Concordant	Non-concordant	Missed opportunity
Drop to one species	No effects identified in the two species, or only in the species chosen to progress[a]	Different (new) effects in species chosen to drop (missed potential new toxicities)[b]	n/a
Retain two species	Different (new) effects in both species[c]	n/a	Effects identified only in one, or neither species[d]

[#] participants had access to pre-FIH and FIH toxicology data but were blind to the actual decision made and any post-FIH toxicology data. n/a: option not applicable for this category.

Concordant: the working group majority consensus decision, based on the available data (which was a supplied summary of the whole data package) were consistent with the toxicities identified in the longer-term post-FIH toxicity studies.

Non-concordant: the working group majority consensus decision, based on the available data (which was a supplied summary of the whole data package) were inconsistent with the toxicities identified in the longer-term post-FIH toxicity studies.

[a] Decision to drop to one species could be considered as low impact on nonclinical assessment of risk to humans, as no effects were identified in either species, or only in the species chosen to progress, in the longer-term post-FIH toxicity studies.

[b] Decision to drop to one species could be considered as potential risk to human safety, as effects were identified in the species chosen to drop (missed potential toxicities).

[c] Decision to retain two species could be considered as low impact on nonclinical assessment of risk to humans, as effects were identified in both species used in the longer-term post-FIH toxicity studies.

[d] Decision to retain two species could be considered as low impact on nonclinical assessment of risk to humans, as effects were identified in only one, or neither species used in the longer-term post-FIH toxicity studies, however, this is a missed opportunity to reduce to one species.

studies, along with the impact of data from each study on drug candidate progression (used for internal decision-making). The panels discussed the data to form a consensus opinion on whether the next (longer-term post-FIH) series of studies could have been performed in a single species, and to propose which species they would have chosen for these toxicity studies. The use of one species was considered acceptable if there were either no, the same, or similar toxicities from the totality of data available from pre-FIH and FIH studies. The rodent, rather than non-rodent, would be the species progressed if appropriate (as per the current ICHS6(R1) guideline, but expanded for this exercise to include any drug modality). The use of one species was also considered acceptable where the toxicities in two species at FIH were classified as different, when the difference was the absence of toxicities in one species but presence of toxicities in the other; the most-sensitive species was progressed in such cases. The actual results from the longer-term post-FIH studies were then revealed and the views reached by the panel were assessed for concordance relative to the toxicities observed, as outlined in Table 3.

3. Results

3.1. Overview of the dataset

Data were collected from 18 different pharmaceutical and biotechnology companies, including some submitted through CROs/consultants. Five or more drug candidates (with a mean of ten) were submitted to the dataset by 15 companies. The analysed dataset consisted of 172 drug candidates that met the inclusion criteria: 92 small molecules, 46 mAbs, 15 recombinant proteins, 13 synthetic peptides and 6 ADCs (Fig. 1a; comprising 53%, 27%, 9%, 8% and 3% of the total dataset, respectively).

The dataset included the following demographic information:

- The drug candidates were representative of a wide range of therapeutic areas, with oncology, central nervous system (CNS) and immunomodulatory indications being the most highly represented (Fig. 1b).
- Approximately one half (54%) of the drug candidates were defined as at the FIH stage (93/172 drug candidates), with 32 drug candidates at the pre-FIH stage (19%) and 47 drug candidates at the post-FIH stage (27%; Table 4).
- Approximately two-thirds (66%) of the drug candidates (114/172) were in active development while development of the remaining 58 drug candidates had been stopped (Table 4). The most common explanation for development having been stopped was nonclinical toxicology findings - the case for 13 pre-FIH, 16 FIH and 2 post-FIH drug candidates (Fig. 1c and Fig. 1d). This corresponds to 40%, 17% and 2% of all molecules at each of these stages respectively, or 87%, 43% and 33% of the stopped drug candidates at each of these stages respectively.

3.2. The number of nonclinical species used

For the majority of drug candidates, toxicology studies were conducted in either one or two species for the entire package, with drug candidates reaching varying stages of development (as described in section 3.1). For a small number of drug candidates (8/133), initial toxicology studies were conducted in two species, whilst studies supporting later stages of development were conducted in only one species. No drug candidates within the database used more than two species within the package.

3.2.1. Use of a single nonclinical species

Toxicology studies were conducted using only one species for 39 drug candidates (Fig. 2a): 3 small molecules, 32 mAbs, 3 recombinant proteins and 1 ADC (3%, 70%, 20% and 17% of each molecule-type, respectively). The single nonclinical toxicology species was a rodent for 4 of these drug candidates, whilst for 35 drug candidates it was a non-rodent:

- Small molecules (3 drug candidates): toxicology studies were conducted in rat only for two small molecules, with development terminated following pre-FIH studies. For the other small molecule, only the dog was used; however, development was terminated after conducting the FIH-enabling toxicology studies.
- mAbs (32 drug candidates): for the majority (30 mAbs) only the NHP was used for toxicology studies. Of the two exceptions, one drug candidate used a transgenic mouse model (in active development following post-FIH studies) and rat was used for the other (in active development following FIH-enabling toxicology studies).

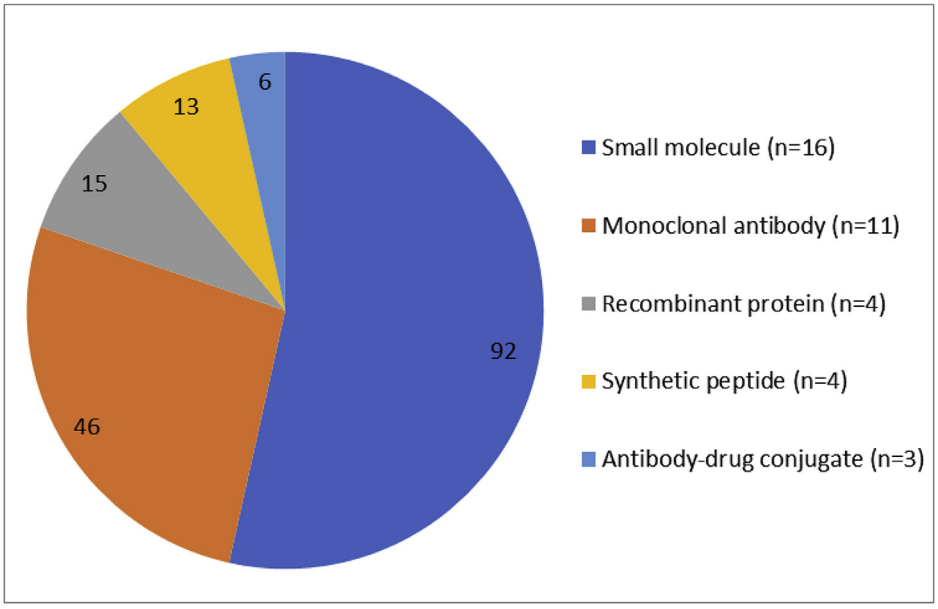

Fig. 1a. Total number of drug candidates per molecule type within the dataset. Number (n) of organisations submitting data for the specific molecule type is shown in parenthesis.

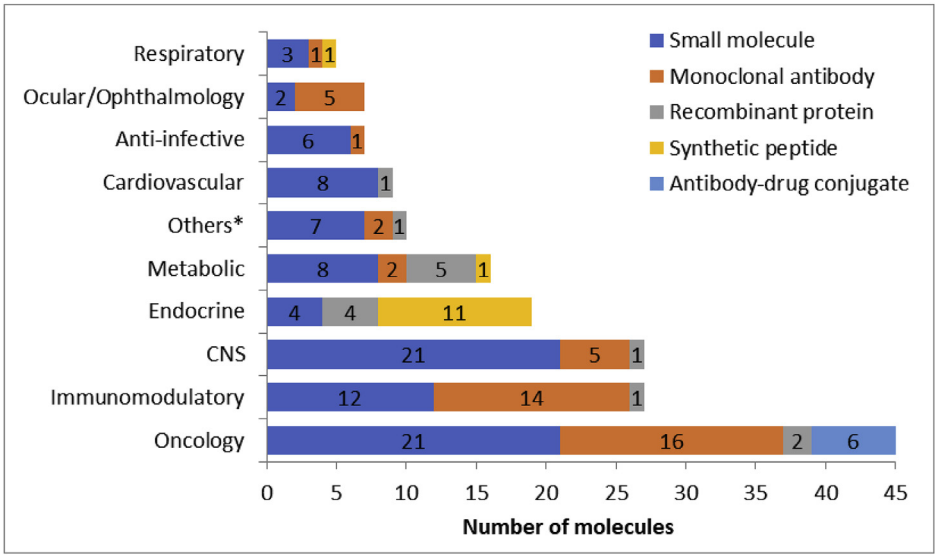

Fig. 1b. Therapeutic indications of the molecules within the dataset. Others*
includes "not disclosed" (3), "bleeding disorders/haematology" (2), "urology" (3), "gastrointestinal" (1) and "gynaecology" (1).

Table 4
Distribution of molecules (n = 172) across the three phases of development.

	Number of molecules in active development				Number of molecules stopped in development			
	Pre-FIH	FIH	Post-FIH	Total	Pre-FIH	FIH	Post-FIH	Total
Small molecule	12	26	25	63	11	18	0	29
Monoclonal antibody	0	19	10	29	1	11	5	17
Recombinant protein	3	4	1	8	1	5	1	7
Synthetic peptide	2	3	5	10	0	3	0	3
Antibody-drug conjugate	0	4	0	4	2	0	0	2
Total for phase	17	56	41	114	15	37	6	58

Pre-FIH: molecules that have completed early toxicology studies which may be performed to assist molecule selection and provide pilot information for the pivotal GLP-compliant studies.
FIH: molecules that have completed pivotal GLP-compliant toxicology studies intended to support a phase I clinical trial.
Post-FIH: molecules that have completed pivotal GLP-compliant long-term toxicology studies intended to support longer duration clinical trials.

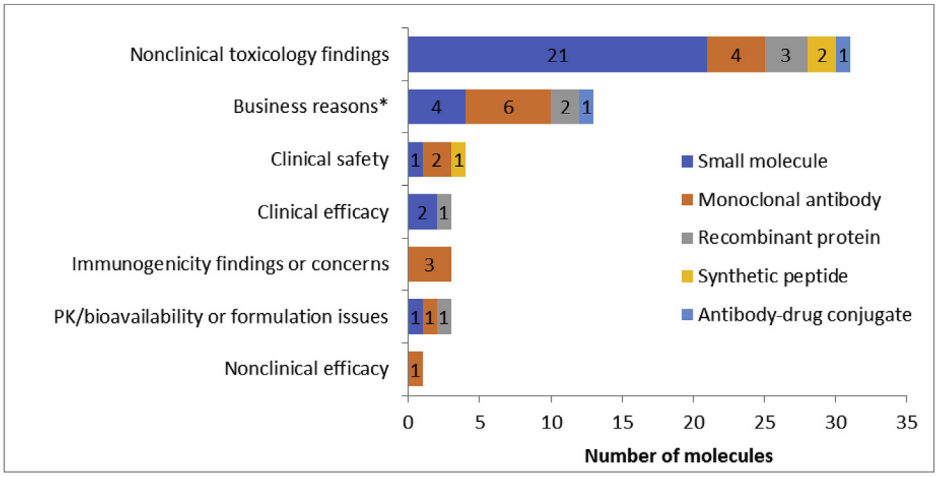

Fig. 1c. Primary reason for stopping molecules during development – by molecule type. Business
reasons*
includes three separate categories from the questionnaire: "not or no longer lead molecule", "competitive landscape" and "rationalisation of company portfolio".

- Recombinant proteins (3 drug candidates): all used NHP only. One in active development following pre-FIH studies, another in active development following FIH studies and the third (using rhesus macaque) was stopped following FIH studies.
- ADC (1 drug candidate): only NHP used and development was terminated after pre-FIH studies.

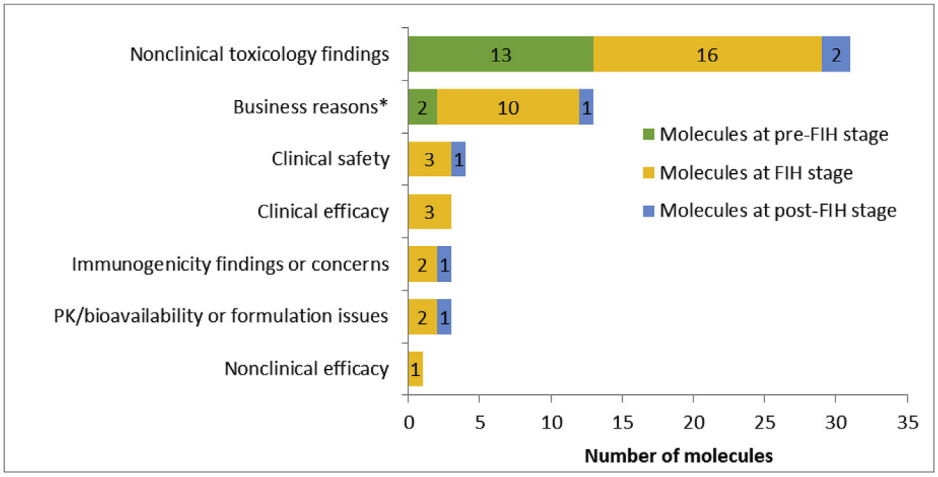

Fig. 1d. Primary reason for stopping molecules during development – by progression stage. Business reasons*
includes three separate categories from the questionnaire: "not or no longer lead molecule", "competitive landscape" and "rationalisation of company portfolio".

3.2.2. Use of two nonclinical species

Toxicology studies had been conducted in two species (a rodent/lagomorph and non-rodent) for the remaining 133 drug candidates (Fig. 2a): 89 small molecules, 14 mAbs, 12 recombinant proteins, 13 synthetic peptides and 5 ADCs (97%, 30%, 80%, 100% and 83% of each molecule-type, respectively). Overall the rat was the predominant rodent species (Fig. 2b), whilst dog and NHP were the main non-rodent species (Fig. 2c):

- Small molecules (89 drug candidates): rodent species was predominantly rat (83 drug candidates), with wild-type mouse being used for a further 5 and rabbit for another 1 small molecule (90%, 5% and 1% of small molecules, respectively). The non-rodent species was mainly dog (59 drug candidates) or NHP (29 drug candidates, including one rhesus macaque), whilst minipig was used for only 1 small molecule (64%, 32% and 1% of small molecules, respectively).
- mAbs (14 drug candidates): NHPs were always the non-rodent species. There was more variation in rodent species (Fig. 2b), including rat (7 mAbs), wild-type mouse (2 mAbs) and transgenic mouse (3 mAbs). A further two mAbs used rabbit as a second non-rodent species, as a relevant small animal model for the intravitreal dosing route.
- Recombinant proteins (12 drug candidates): rat (9 drug candidates) and NHP (10 drug candidates, including rhesus macaque for one) were used predominantly; wild-type mouse (3 drug candidates) and dog (2 drug candidates) were also used.
- Synthetic peptides (13 drug candidates): rat was the rodent for almost all (12 drug candidate) studies, with the wild-type mouse used for 1 drug candidate. The non-rodent species was almost equally split between dog (7 drug candidates) and NHP (6 drug candidates).
- ADCs (5 drug candidates): For all the ADCs, NHP was the non-rodent species, with the rat used as the rodent species in 4 cases and wild-type mouse in the other.

3.2.3. Reduction to a single species

For a small number of the drug candidates (8/133) for which two species were used during pre-FIH or FIH toxicology studies, reductions to one species were made later in development (Table 5):

- Small molecule (1 drug candidate): this was in active development; post-FIH studies were performed in dog only, whereas earlier studies were performed in rat in addition to dog.
- mAbs (5 drug candidates): for two mAbs, still active in development at post-FIH stage, post-FIH studies were performed in rat only, the NHP having been dropped after FIH studies. For three further mAbs, use of the rodent/lagomorph was dropped after initial pre-FIH or FIH studies, and further studies used NHP only; one of these was still

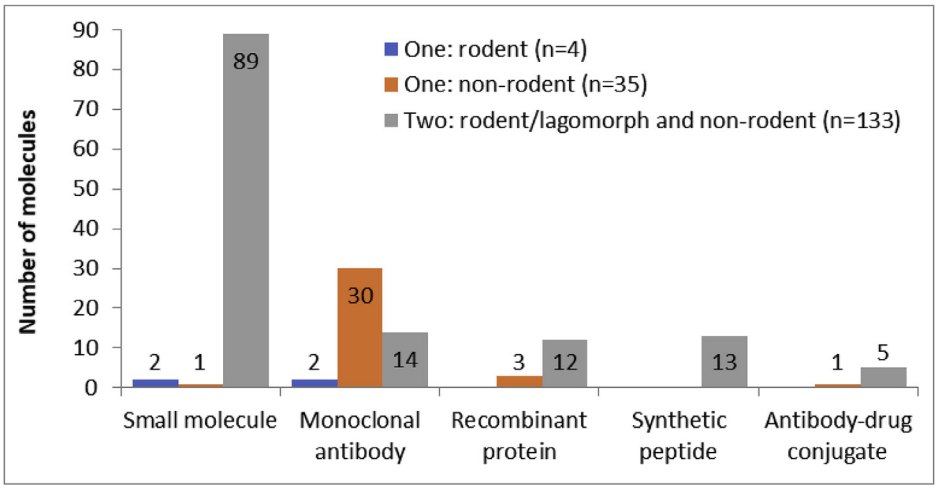

Fig. 2a. Number of species used for different molecule types in any phase.

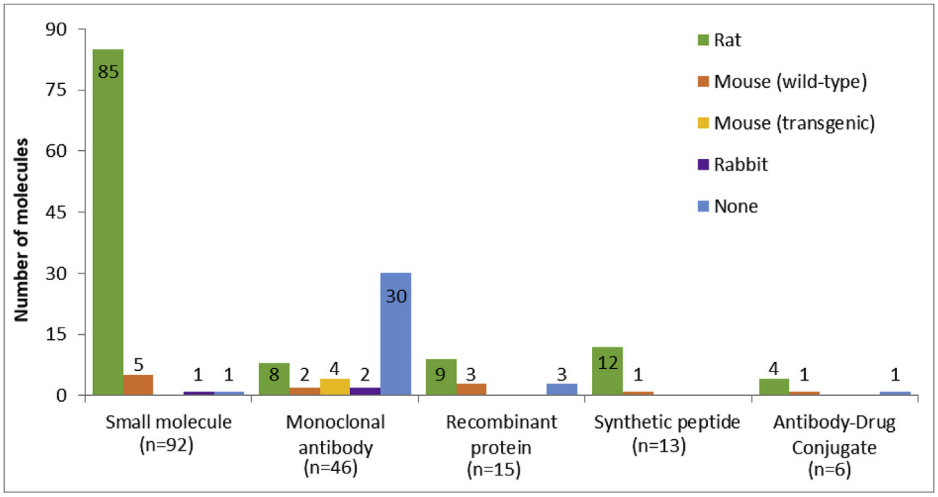

Fig. 2b. Rodent/lagomorph species used for different molecule types
Totals for each species are the sum of molecules using only one species (see section 3.2.1) and molecules using two species (see section 3.2.2). For example, 2 small molecules used only a rodent, as shown in Fig. 2a (the rat) whilst 83 used the rat in addition to a non-rodent, combining as 85 small molecules using rat displayed in the figure above. For each molecule type, the total number of molecules is identified in parenthesis; this number provides the denominator to calculate % values that are presented in the results text.

in active development at the FIH stage.

- ADCs (2 drug candidates): both were still in active development. FIH-enabling toxicity studies were conducted in NHP only, while earlier studies also included the rat for one molecule, and wild-type mouse for the other.

3.3. The value of data derived from the use of two nonclinical species

For 125/133 drug candidates, two species were retained throughout all stages of development. For these molecules, respondents were asked whether, with the benefit of hindsight, decisions could have been made using only one species (Fig. 2d). From 122 responses received (a response rate of 98%), for 81 drug candidates (66%) it was stated that in hindsight, decisions could have been made from a single species, with similar responses noted for the different drug modalities: 56 of the 85 small molecules (66%), 8 of the 9 mAbs (89%), 7 of the 12 recombinant proteins (58%), 8 of the 13 synthetic peptides (62%) and 2 of the 3 ADCs (67%). Respondents were asked which species they would have selected for progression. There was no clear trend for which species was felt to provide the most useful data, regardless of drug modality. Of the 56 small molecules, the rodent was chosen for 25 (45%) and the non-rodent for 31 (55%), while for biologicals, the number of molecules was insufficient to draw robust conclusions.

3.4. Target organ toxicities at the FIH stage

Two species were used in pre-FIH and FIH studies for 115 drug candidates: 75 small molecules, 13 mAbs, 11 recombinant proteins, 12 synthetic peptides and 4 ADCs. There were either no, the same, or similar target organ toxicities in both species for 45 of these drug candidates: 24 small molecules (32%), 11 mAbs (85%), 4 recombinant proteins (36%), 5 synthetic peptides (42%) and 1 ADC (25%) (Table 6a). For a subset of molecules being developed for the treatment of advanced cancer and following ICHS9, including 18 small molecules, 2 mAbs and 4 ADCs, the majority of the small molecule subset (15 or 83%) and ADCs (3 or 75%) displayed different toxicities between the species evaluated. The 2 mAbs displayed no toxicities in either species.

For the 11 drug candidates for which ICHS6(R1) was being followed and post-FIH studies had been conducted, 5 (2 mAbs, 1 recombinant protein and 2 synthetic peptides) showed different target organ toxicities across the two species used at the FIH stage. For the other 6 drug candidates (4 mAbs, 1 recombinant protein and 1 synthetic peptide), either the same, similar or no toxicities were observed across species in toxicology studies at the FIH phase. Despite this, only 2 mAbs were tested in a single species for the longer-term post-FIH studies.

There were 8 drug candidates (1 small molecule, 5 mAbs and 2 ADCs) that did reduce to one species for later (FIH or post-FIH) studies (Table 6b). For three mAbs, there were either no toxicities or the same target organtoxicities were recorded in both species at the preceding stage, and these drug candidates then progressed with toxicology studies in rodent only. For two other mAbs, toxicities were different in the two species, with more target organs identified in the non-rodent. As a result, the non-rodent was the species of choice for longer-term studies. For the ADCs, toxicities across the two species were the same for one

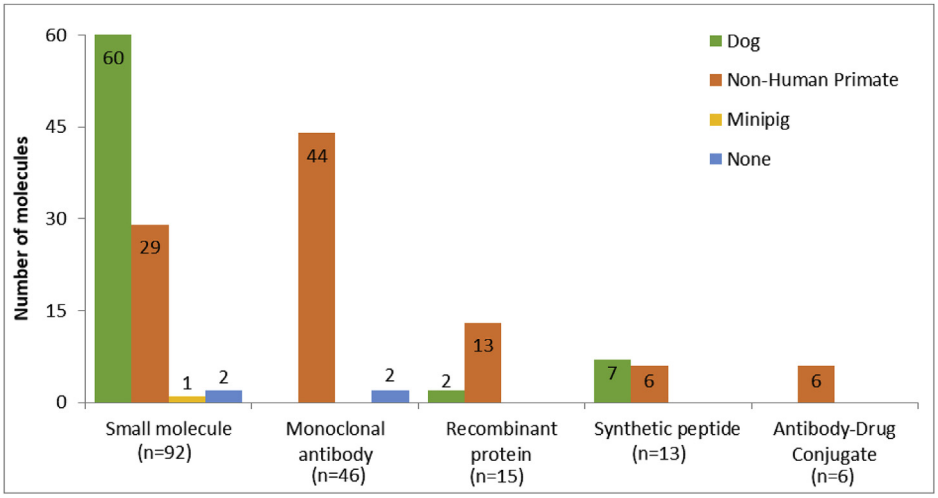

Fig. 2c. Non-rodent species used for different molecule types
Totals for each species are the sum of molecules using only one species (see section 3.2.1) and molecules using two species (see section 3.2.2). For example, 30 mAbs used only a non-rodent, as shown in Fig. 2a (the NHP) whilst 14 used the NHP in addition to a rodent/lagomorph, combining as 44 mAbs using NHP displayed in the figure above. For each molecule type, the total number of molecules is identified in parenthesis; this number provides the denominator to calculate % values that are presented in the results text.

Table 5
Numbers of molecules using two species (n = 133) that reduced to one species (n = 8) at one of the three phases of development.

	Guidelines followed[#]				Molecules at pre-FIH that reduced to one species	Molecules at FIH that reduced to one species	Molecules at post-FIH that reduced to one species
	Total	ICHM3 (R2)	ICHS6(R1)	ICHS9			
Small molecule	89	68	–	21	0/21	0/43	1[a]/25
Monoclonal antibody	14	–	14	–	1[b]/1	2[c]/7	2[d]/6
Recombinant protein	12	–	12	–	0/3	0/7	0/2
Synthetic peptide	13	3	10	–	0/2	0/6	0/5
Antibody-drug conjugate	5	–	2*	3	0/1	2[e]/4	–

Regulatory pathway followed, in addition to ICHM3(R2). Molecules within column 'ICHM3(R2) were only following those guidelines.
* these compounds also followed ICHS9 guidelines. -: no molecules in this category.
[a] molecule (for a bleeding disorder) still active in development; progressed in dog only (previous studies in rat also).
[b] molecule (ocular/ophthalmology indication) had stopped development; progressed in NHP only (previous studies in rabbit also).
[c] one mAb (for a CNS indication) was still active in development. This progressed in NHP only (previous studies in transgenic mouse with a surrogate molecule). The other mAb (for a respiratory indication) had stopped development. This had progressed in NHP (previous studies in wild-type mouse also).
[d] both mAbs (for a CNS indication and the other undisclosed indication) were still active in development. These progressed in rat only (previous studies in NHP also).
[e] both ADCs (for oncology indications) were still active in development. They progressed in NHP only (previous studies in rat for one ADC and wild-type mouse for other).

and different for the other, with more target organs identified in the non-rodent (NHP), which was the species that progressed in both cases. For the small molecule, toxicities were different at the preceding FIH stage, with more target organs identified in the rodent (rat). However, the non-rodent (dog) was progressed into longer-term studies as this was deemed more appropriate to assess a specific toxicity (cardiovascular-based), even though more toxicities had been observed in the rodent (this information was provided within the free text section of the questionnaire).

3.5. Blinded exercise investigating the potential for post-FIH studies to be conducted in a single nonclinical species

There were 38 drug candidates that had used two species at the FIH stage and progressed to longer-term post-FIH studies: 35 of these (23 small molecules, 6 mAbs, 2 recombinant proteins and 4 synthetic peptides) were assessed within an exercise to determine if, retrospectively (and with the benefit of hindsight), single species toxicology testing could have been considered for the longer-term (post-FIH) studies based on FIH data. There were insufficient data for the remaining 3 drug candidates, and therefore these were excluded from the exercise.

The panels of working group members decided it should have been possible to decrease to one species for the longer-term studies for 14 small molecules, 5 mAbs and 3 synthetic peptides (Table 7a). When the longer-term study data was unblinded, these decisions would have presented a low impact on human risk assessment for all the mAbs and synthetic peptides, but only for 9 of the small molecules (64%), as there wereno toxicities in either of the two species used in the longer-term studies, or toxicities were only identified in the species chosen to progress. For the remaining 5 small molecules, toxicities were identified in the 'dropped' species (detailed in Table 7b), and therefore this decision would have missed these potential new safety concerns.

The panels also made the hypothetical decision to retain two species for the longer-term studies for 9 small molecules, 1 mAb, 2 recombinant proteins and 1 synthetic peptide. For 7 (78%) of the small molecules and for 1 recombinant protein, this decision was concordant, as new toxicities were subsequently identified in both species in the longer-term studies. For the remaining 5 drug candidates (2 small molecules, 1 mAb, 1 recombinant protein and 1 synthetic peptide), there were no toxicities identified in either one or both species in the longer-term studies. For these molecules, the post-FIH toxicity evaluation would have been informed using one species, with a low impact on human risk assessment. As such, these cases were considered to be missed opportunities to reduce to one species.

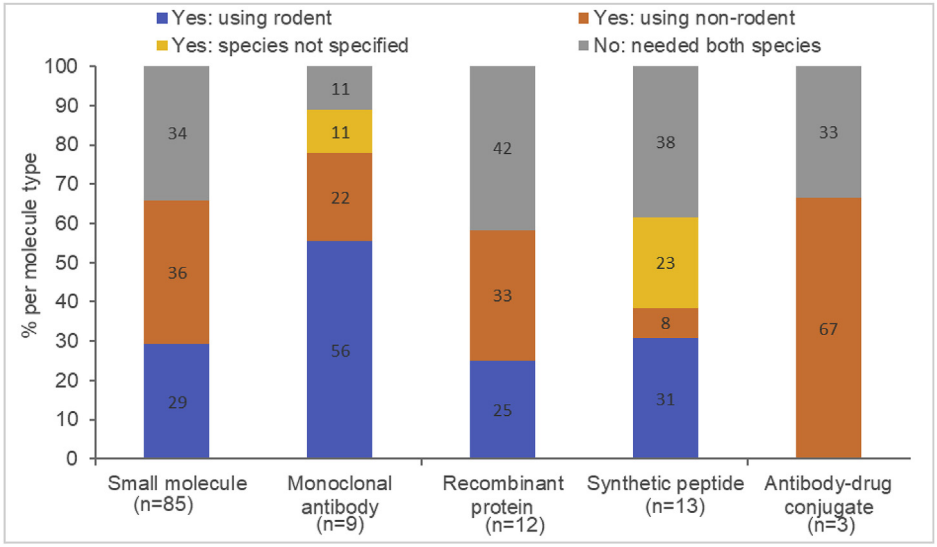

Fig. 2d. Survey responses to the question 'For the development of molecules in which two species were used, would the same decisions have been able to be made with data from one species only (in hindsight)?'
Yes or No answers to the survey question 'If you used two species, would you have been able to make the same decisions with data from one species only (in hindsight)? If Yes, which species was decision-making?'. Data are for the 125 compounds that used two species and had not reduced to one species (for three small molecules using two species, this question was not answered). For one mAb and three synthetic peptides the answer was Yes, but the decision-making species was not specified (defined as 'Yes, species not specified' in the figure above).

Table 6a
Target organ toxicities identified in the two species during pre-FIH and FIH toxicity studies.

	Target organ toxicities in the two species					
	None	Same	Similar	Different	If similar: species with the most toxicities	If different: species with the most toxicities
Small molecule (75)	3	11	10	51	Rodent (4); Non-rodent (4); both* (2)	Rodent (23); Non-rodent (15); both (13)
Oncology products (18)#	–	2	1	15	Rodent (1)	Rodent (7); Non-rodent (6); both (2)
Monoclonal antibody (13)	8	3	–	2	–	Non-rodent (1); both (1)
Oncology products (2)#	2	–	–	–	–	–
Recombinant protein (11)	1	1	2	7	Rodent (1); Non-rodent (1)	Rodent (4); both (3)
Synthetic peptide (12)	4	–	1	7	Non-rodent (1)	Rodent (4); Non-rodent (2); both (1)
Antibody-drug conjugate (4)#	–	1	–	3	–	Rodent (2); Non-rodent (1)

*both: little or no overlap in target organ toxicities in the two species.
-: no molecules in this category.
#These molecules intended for oncology indications were following ICHS9 guidelines. Subset of the total small molecules or monoclonal antibodies.
Definitions of the table categories.
None = no target organ toxicities identified in both species.
Same = toxicities identified in the same target organs in both species.
Similar = toxicities identified primarily in the same target organs (only one additional or different target organ toxicity was identified in one of the species).
Different = No target organ toxicities identified in one species and one or more in the other, or multiple toxicities identified in each species with more than one target organ toxicity different between the two species.

4. Discussion

The primary purpose of the expert working group was to identify if there are wider opportunities for use of a single nonclinical species in toxicology programmes, regardless of drug modality, relevant regulatory guidance and/or phase of drug development. To enable this, data were provided by eighteen different pharmaceutical and biotechnology companies, mainly based in the USA and Europe (53 and 44% of drug candidates in the dataset, respectively) and Japan, generally representing larger multi-national companies. In order to reduce potential for selection bias, contributors were asked to provide information for their most recent drug candidates to have completed at least one of three phases of development (pre-FIH, FIH or post-FIH). Information was subsequently supplied on molecules for a wide variety of therapeutic indications, spread across the three development phases. Whilst this could not entirely exclude the possibility of some selection bias regarding the most recent molecules, this is a common limitation of these types of data-sharing activities and as such should be viewed as providing a snap-shot of current industry practice. Additionally, the small sample sizes for some of the drug modalities and development phases has limited the conclusions that can be drawn. Nevertheless, the balance of drug modalities that data were provided for (53% small molecules and 47% biologics) and their associated range of therapeutic indications were considered by the working group as a reasonable representation of the industry, especially since two-thirds of the data were derived from drug candidates in active development at the time of the review. For drug candidates that were no longer in development, nonclinical toxicology findings were responsible for 87% of drug candidates terminated pre-FIH, and for 43% of drug candidates stopped at the FIH stage. Since a primary purpose of toxicity studies is identification of hazards to humans and risk assessment, this is not an unexpected outcome, and demonstrates the key role such studies play in successfully ensuring safety for phase I trials. These figures are consistent with those from other reports (Waring et al., 2015) and signify

Table 7a
Blinded exercise: outcomes for the potential to reduce to one species or to retain two species for the longer-term post-FIH toxicity studies.

Working Group Decision		Concordant	Non-concordant	Missed opportunity
Small molecules				
Drop to one species	14	9 (64%)	5 (36%)	n/a
Retain two species	9	7 (78%)	–	2 (22%)
Other molecules*				
Drop to one species	8	8 (100%)	–	n/a
Retain two species	4	1 (25%)	–	3 (75%)

See Table 3 for concordant and non-concordant definitions.
n/a: option not applicable for this category.
-: no molecules in this category.
* The mAbs, recombinant proteins and synthetic peptides have been collated into an 'other molecules' category. The eight molecules with decisions to 'drop to one species' were five mAbs and three synthetic peptides. The four molecules with decisions to 'retain two species' were one mAb, two recombinant proteins and one synthetic peptide. The concordant molecule was a recombinant protein.

Table 6b
Target organ toxicities identified in the two species for the eight molecules that reduced to one species at FIH or post-FIH phases#.

	Target organ toxicities in the two species (and species that progressed)				
	None	Same	Similar	Different	If different, species with the most toxicities
Small molecule	–	–	–	1 (non-rodent) [a]	rodent
Monoclonal antibody	2 (rodent)[b]	1 (rodent)[c]	–	2 (non-rodent)[d]	non-rodent
Antibody-drug conjugate	–	1 (non-rodent)[e]	–	1 (non-rodent)[e]	non-rodent

#as identified in Table 5.
-: no molecules in this category.
[a] molecule was still active in development (post-FIH); as the different effects were cardiovascular-based, the non-rodent (dog) was progressed as this was deemed more appropriate for cardiovascular assessments, even though more toxicities had been observed in the rodent.
[b] both mAbs were still active in development (one at FIH, the other at post-FIH).
[c] molecule was still active in development (post-FIH).
[d] both mAbs had stopped development (one at pre-FIH, the other at FIH).
[e] ADC was still active in development (FIH).

Table 7 b
Details of the toxicities identified during pre-FIH/FIH and post-FIH toxicity studies: for the five small molecules deemed 'non-concordant' in the blinded exercise.

Therapy area	Key findings (pre-FIH/FIH studies)[a]		Working group decision for species to progress	Key findings (post-FIH studies)	
	Rodent	Non-rodent	(Reason for choice)	Rodent	Non-rodent
Immunomodulatory	7 day dosing: clinical signs, clinical chemistry & haematology, immune system, endocrine, bile ducts/liver, kidneys/ureters. 4 week dosing: clinical chemistry & haematology	14 day dosing: clinical signs, clinical chemistry & haematology, skin, GI tract/stomach, immune system, bile ducts/liver. 4 week dosing: clinical chemistry	Non-rodent (clinical FIH starting and ceiling doses were based primarily on results from this species)	26 week dosing: mortality, clinical chemistry & haematology, immune system, **endocrine**	39 week dosing: clinical signs, clinical chemistry
Immunomodulatory	4 week dosing: clinical chemistry, GI tract/ stomach	4 week dosing: clinical chemistry, GI tract/ stomach	Non-rodent (clinical FIH ceiling dose was based primarily on results from this species)	13 week dosing: **skin**, GI tract/ stomach. 26 week dosing: **skin**, GI tract/ stomach	13 week dosing: no effects 39 week dosing: GI tract/stomach
CNS	6 week dosing: clinical chemistry, endocrine, bile ducts/liver	6 week dosing: clinical signs, clinical chemistry	Rodent (species with most toxicities)	13 week and 26 week dosing: clinical chemistry, endocrine, bile ducts/liver	13 week dosing: clinical signs, clinical chemistry. 39 week dosing: **idiopathic canine polyarteritis**
Cardiovascular	14 day dosing: bile ducts/liver. 4 week dosing: bile ducts/liver, kidneys	14 day dosing: clinical signs (vomiting). 4 week dosing: no effects	Rodent (species with most toxicities)	13 week and 26 week dosing: bile ducts/liver, kidneys	39 week dosing; bile ducts/liver, kidneys, **male reproductive organs**
Anti-infective	14 day dosing: GI tract/stomach	14 day dosing: no effects	Rodent (species with most toxicities)	13 week dosing: GI tract/stomach	13 week dosing: clinical chemistry, **liver**

[a] summarised findings from all pre-FIH and FIH studies conducted in each species. **Bold text** indicates the new finding (potential missed safety concern).
Rat was the rodent species used for all five small molecules. The non-rodent species used was NHP for the two immunomodulatory small molecules and dog for the remaining three. All five small molecules were still active in development.

that with the current nonclinical paradigm, a large proportion of drug candidates fail to reach the clinic due to toxicology findings in nonclinical animal studies, often at relatively late stages in drug discovery. Unfortunately, the relevance and translatability of nonclinical findings is not always clear and hence is why the current approach is conservative. However, the second most common explanation for stopping development (27% of stopped drug candidates at FIH) was for business/strategic reasons (no longer the lead molecule selected, competitive landscape or rationalisation of company portfolio). As previously highlighted (Morgan et al., 2018; Waring et al., 2015) this indicates that drug company portfolios are continually being re-evaluated strategically, such that factors other than adverse nonclinical findings often contribute to termination of drug candidates in development.

4.1. Nonclinical species used for development of small molecules

The small molecules were being developed following ICHM3(R2) in general, or in conjunction with ICHS9 for drug candidates for advanced cancer indications. These guidelines generally expect two species to be used for toxicity testing (rodents and non-rodents), with no explicit wording around opportunities for reducing to one species, other than case-by-case approaches within ICHS9. The species used for toxicity testing of small molecules in the dataset reflect those of other industry reviews (Baldrick, 2008; Butler et al., 2017), in that rat, dog and NHP are the predominant species used (92%, 65% or 32% of small molecules, respectively). Interestingly, minipig was used for only one small molecule, which would appear to be an under-representation of the use of this species for toxicology testing across the industry (Colleton et al., 2016; Heining and Ruysschaert, 2016; Monticello et al., 2017). Discussion within the working group indicated that some companies currently use, or are considering use of, minipig as the non-rodent species for toxicology testing, and therefore it is unclear why minipig data for only one drug candidate was submitted. It is possible that the retrospective nature of the dataset may not reflect a more recent transition to using minipig as a non-rodent species for toxicology testing and this may warrant more investigation.

There were three examples of a small molecule tested in a single species only. For two of these, only pre-FIH studies were conducted (in rodent) and development was stopped due to the toxicology findings, prior to the start of studies in a non-rodent, a common approach taken for early toxicity screening (Roberts et al., 2014). The other small molecule was tested in dog only, and development was stopped after FIH-enabling toxicology studies. No further information was provided to explain the use of this single species. With one further small molecule, FIH-enabling toxicology studies were conducted in two species (rat and dog), with longer-term post-FIH studies conducted in the dog only; this molecule was for a bleeding disorder and was still in active development. The exact reasons for reducing to a single species with this small molecule were not recorded within the survey, although since the primary toxicity of concern appeared to be cardiovascular-based, the dog was considered a more appropriate species for further investigation. Overall, these cases indicate that for a very small number of drug candidates, and for specific circumstances that are unclear from the data, companies have been able to scientifically justify novel package designs outside the established guideline recommendations, at least for internal decision-making.

4.2. Nonclinical species used for development of biologics

All 46 mAbs, 15 recombinant proteins and 6 ADCs, as well as most (10 out of 13) of the synthetic peptides, followed the principles outlined in ICHS6(R1)and ICHM3(R2). For the majority of the mAbs (30), 3 recombinant proteins and 1 ADC, toxicity studies were performed in NHP only. However, there were two examples of mAbs where only a rodent species was used. One mAb, in active development for an immunomodulatory indication, conducted the toxicology data package (up to 26-week studies) using a transgenic mouse model only. No further information was provided about the human-specificity of this mAb or why none of the conventional rodent and non-rodent species were pharmacologically relevant. The other mAb, in active development for an anti-infective indication, used rat only for the FIH-enabling study (a 14-day study only, as outlined in ICHS6(R1) as sufficient for foreign targets). Further information provided indicated that as no orthologous target existed in any nonclinical species, the rat was used to assess off-target toxicities at and above anticipated human exposures.

For the remaining 14 mAbs, 12 recombinant proteins, 10 synthetic peptides and 5 ADCs (53% of the drug candidates following ICHS6(R1) within the dataset) toxicity studies were performed in two species. For mAbs, this equates to 30% of the mAbs within the dataset with cross-reactivity to two species, consistent with working group member experience, but not previously published from a cross-industry perspective. Of the eleven different companies which submitted mAb data, seven provided data on drug candidates that were tested in two species. Multiple pharmacologically-relevant species for biologics raises a 3Rs challenge: when two species toxicology with biologics is required, this results in higher animal use overall. However, when only one pharmacologically relevant species is identified, this tends to be the NHP, (as confirmed in the current dataset). Whilst overall animal use would be lower, especially as smaller group sizes are generally employed (compared to rodent studies), selection of the NHP as the toxicology species should always be a balance between robust scientific justification (including demonstrable lack of suitability of any other species) and ethical considerations. Multi-species cross-reactive drug candidates offer other options to reduce NHP use during development, through early toxicity screening in rodents and potentially by reduction to a single rodent species for longer-term post-FIH toxicity studies. In addition, rodent crossreactivity would allow additional evaluations without NHP use, in particular reproductive, developmental or juvenile toxicity assessments, if warranted.

The use of two species in toxicity studies with recombinant proteins, synthetic peptides and ADCs reflects approaches which combine the principles outlined in ICHM3(R2) and ICHS6(R1). Although there is a lack of specific regulatory guidance for therapeutic peptides/proteins, the seven different companies which provided data for these drug modalities followed a similar approach and generally used two species within the toxicology package. For the three recombinant proteins that were assessed using only a single species (NHP), survey data indicated that this was the only pharmacologically-relevant species.

Five of the biologics (2 ADCs and 3 mAbs) tested in two species for pre-FIH or FIH studies, used NHP only for subsequent general toxicology studies. The reasons for these decisions were not stated, although for ADCs it is common for only short-term rodent studies to be conducted to explore off-target toxicities of the small molecule payload. Two further mAbs reduced to rodent only for longer-term post-FIH studies. For these, target organ toxicities identified at FIH in the rodent and NHP were absent or identical, thus fulfilling the criteria of similar as outlined within ICHS6(R1) and the principle that rodent is preferred to the non-rodent if there are no other reasons driving species selection. However, 9 other molecules following ICHS6(R1) retained both rodent and non-rodent species for longer-term post-FIH studies (4 mAbs, 2 recombinant proteins and 3 synthetic peptides). Target organ toxicities identified at FIH in two species were different for five of these molecules (2 mAbs, 1 recombinant protein and 2 synthetic peptides), which likely contributed to the decision to continue in two species for the post-FIH studies, although no further information was provided within the surveys to confirm this. However, the remaining 4 drug candidates (2 mAbs, a recombinant protein and a synthetic peptide), were associated with no or similar target organtoxicities in the two species at FIH and therefore might have presented the opportunity to reduce to one species. As previously discussed, the synthetic peptides tend to follow a small molecule-like package and there may be a perception that the longer-term toxicity studies should be performed in two species, as per

ICHM3(R2), rather than risk a reduction to one species, even when following ICHS6(R1). As no information was provided to explain the decisions for each drug candidate, it is also acknowledged that there may be sound scientific reasons why both species were progressed even when the target organ toxicities were the same or similar.

4.3. The value of data from two species for drug development

For the 125 drug candidates for which studies were conducted in two species throughout the package, a large proportion (66%) of the survey responders (for 56 small molecules and 25 biologics) stated that in hindsight, the same decisions relevant to human safety and clinical development could have been made based on data from one species (Fig. 2d). This perspective lends support to a notion that there were opportunities for a single species approach to be used more widely. However, the ability to predict which nonclinical species would provide the most useful/relevant data may be challenging in some cases, especially as there was no general preference for whether the rodent or non-rodent provided the key data to inform decision-making for a specific drug modality. Furthermore, the survey respondents did presumably have the benefit of wider knowledge about other nonclinical (and perhaps clinical) data within the package than that provided within the questionnaire. The prospective situation of potentially missing an important toxicity if a study in another species is not performed is difficult to argue against, when those data have historically been available to aid decision-making. However, the purpose of this project was to explore whether it could be appropriate to conduct toxicology studies in one nonclinical species, for any drug modality. The data collected indicate that there were numerous examples when the use of two species may not have yielded additional toxicology data that could not have been gained from using only one species. Although it is acknowledged that these inferences were made in hindsight and studies had been performed in two species to reach this conclusion, the data nevertheless raise a question as to whether there are further opportunities for changes from a two species approach towards a wider use of a single species, for small molecules as well as biologics, once the more sensitive species has been identified.

Although the retrospective responses indicate that for a large number of drug candidates data from a particular species may be more important for decision-making than another (perhaps due to severity of the finding(s), predicted exposure margins and human-relevance), it does not necessarily follow that the data from the other species were not also considered useful. In particular, nonclinical to clinical comparison exercises suggest that date from two species does increase the predictive value for either positive or negative effects in the clinic (Monticello et al., 2017). When responses to another question 'If you used two species, what additional value did data from the second species give you?' were reviewed, the answers often fell within an 'added confidence' category, the importance of which cannot be discounted for decision-making. Similar responses were also evident within the individual study data-impact questions (see Table 2b), which often reflected the different drug modalities. For mAbs, where toxicities are generally due to exaggerated pharmacology and/or immunogenicity, the absence or presence of an effect in a single pharmacologically relevant species is already considered sufficient to generate confidence for decision-making, and data from a second species was generally *not* felt to provide value (for 8/9 mAbs that used two species it was considered decisions could be made from one). Data from two species allows assessment of whether a toxicity occurs cross-species and can increase confidence that either a toxicity is more likely, or unlikely to be relevant to human. However, even for small molecules there were also survey responses indicating that the additional data from a second species were *not* felt to provide value (56/85 small molecules considered decisions could be made from one species). Scenarios for when a second species was not considered to provide value included when toxicities were similar between the species (for example, no new toxicities identified), or observed effects were known/expected from the mode-of-action or previous drug candidates in the series/literature. Whilst data from a second species may not always be needed to enable decisions to be made, two species toxicology is currently recommended for drug candidates covered by ICHM3(R2). Therefore, reducing to a single species presents a risk of not meeting regulatory expectations, especially if there are differences of opinion about validity of supporting scientific data for species choice.

4.4. The similarity of data from two species at FIH

A key factor for progression of drug candidates into longer-term post-FIH studies in one or two species revolves around the principles of comparable or similar toxicities being identified in the two species from short-term studies, as per ICHS6(R1) (see Table 1). However, there is no clear definition of 'similar' within regulatory guidance or the scientific literature, or how decisions on this can be reached in practice. There can also be significant differences in interpretation across different authorities, such that the decision to pursue a one-species post-FIH toxicity programme is at the discretion of the sponsor and/or reviewing regulatory agency. This is underscored by the scientific question on whether or not the absence of toxicity in two species constitutes a similar toxicity profile (per communications from company representatives within or outside the working group). Consequently, companies have taken different approaches to the choice of species for longer-term toxicity studies based on individual experience and willingness to accept regulatory risk, with some continuing in two species, whilst others may reduce to one.

A review of target organ toxicities identified in the FIH package of studies was performed, to determine how often toxicities were similar or different between the two species. The definitions of similarity agreed within the working group were either the absence of toxicities in both species, toxicities identified in the same target organs in both species or one different target organ toxicity identified between the species. A different toxicity profile was defined as no toxicity in one species and one or more recorded toxicities in the other, or multiple toxicities in each species with more than one target organ being different between the species. Survey information on toxicity profiles was limited to a high-level check list of target organ categories where toxicity was identified (see Table 2a). For the analysis, the number (and category) of target organs affected were simply counted and compared between species for each molecule. All target organ toxicities were assumed to be of equal importance, a conservative approach as it is recognised that in reality this would not be the case. Similarly, this exercise did not reflect any differences in exposure, results from other nonclinical data (e.g. safety pharmacology data) or translational value of the nonclinical toxicology findings. Nevertheless, using the data available, the 115 drug candidates with FIH package data (75 small molecules and 40 biologics) were classified as having either different target organ toxicities in two species (thus perhaps justifying the continuation of two species for longer-term studies) or similar toxicities in two species (providing evidence for the potential to reduce to one species for longer-term post-FIH studies). Around 40% had similar target organ toxicities in two species, although there was variability between the drug modalities. Of greatest contrast, 85% of mAbs (11/13) showed a high incidence of similar toxicities when two species were used, whereas the incidence of similar toxicities for the 75 small molecules that reached the FIH stage was considerably lower (32%). The intended therapeutic indication was also a factor. Notably, toxicities tended to be different between the species (75%) for drug candidates in development for oncology indications (18 small molecules, 2 mAbs and 4 ADCs with FIH data), to which the ICHS9 guideline applies. This may be due to the more severe toxicities observed or expected in toxicology studies with drug candidates targeting cancer indications, and differences in species susceptibility.

4.5. Could evaluation of toxicity in two species for the FIH-enabling toxicity studies be a decision point for selection of species for longer-term toxicity studies?

From the retrospective evaluations (sections 3.4 and 4.4), the working group was able to identify a sub-set of drug candidates (small molecules and biologics) that were associated with similar toxicities in the two species used for FIH-enabling toxicity studies. If ICHS6(R1) guideline principles had been applicable to all drug candidates (i.e., extended to small molecules) and the definition of similar toxicity used in the exercise was accepted, longer-term post-FIH studies could have been performed in a single species. To evaluate the feasibility of this approach, the survey data from 35 drug candidates that had completed both FIH and post-FIH studies were reviewed, with the participants blinded to the post-FIH phase toxicity data. Small panels of expert toxicologists reviewed the available data from the pre-FIH and FIH studies, to reach a consensus view on whether or not post-FIH studies in a single species would adequately characterise the toxicological hazard. The results from the longer-term post-FIH studies were then revealed and the consensus views were assessed relative to the potential risk to human safety (Table 3). Gaining consensus between the panels to drop a species if appropriate (thus reducing animal use) or to retain two species, entailed considerable discussion on each drug candidate. This reflected the diverse opinions and experience of individuals and their respective company policies. Panel participants concurred that a similar depth of discussion or consideration by a sponsor or individual reviewer would be expected in making their decision based on similar data. It was recognised that the same decision/conclusion may not necessarily be made where these decision makers have access to further pivotal data that was not provided within the surveys, such as *in vitro* data, mechanistic data, full study reports (including histopathology etc) for the studies that were considered, or data from other animal studies within the package, and that this exercise therefore reflected a conservative approach.

For 9 of the 14 small molecules (64%), and for all (8) other molecules, longer-term studies in a single species were considered appropriate. If this approach had been followed, it was considered as low risk to the nonclinical assessment of human safety, as generally no new and differential toxicities were identified in the species dropped from the longer-term post-FIH studies. However, there were five small molecules (36%) that were an exception to this, where new toxicities of concern were identified in species the expert panels had suggested to drop for longer-term studies. Although these five drug candidates were therefore classified as 'non-concordant decisions', this may reflect the conservative approach and may not be the situation when additional information is available. For example, the new toxicities may not translate to humans (such as the canine polyarteritis observed for drug candidate 3 in Table 7b) or may be expected exaggerated pharmacology and monitorable in the clinic (this additional information was provided for drug candidate 1 in Table 7b). Therefore, although with some small molecules new and differential toxicities were identified in the species the expert panel had suggested to drop for longer-term studies, this could be viewed as a worse-case scenario and decision making would likely be improved with the knowledge of the full breadth of data available for the respective drug candidates.

In order to progress this work, a larger dataset of drug candidates from individual drug modalities would be required, where FIH and post-FIH toxicology data were available. Other information that would also aid future evaluation could include broader details regarding the drug candidate and nonclinical information such as primary, secondary and safety pharmacology data, species relevance justification, toxicity findings including nature, severity and incidence, dose relationship, exposure data and safety margins, expected/exaggerated pharmacology, NOAELs in the different species and which species data drove the clinical safety margins. Such information from drug candidates that have conducted FIH and longer-term post-FIH toxicology packages in two species would allow investigation of the nature of any new toxicities in the longer-term studies and their relevance to human safety. Additionally, access to human safety data would allow for the most complete analysis of animal species decisions for post-FIH toxicology studies. Such an approach would enable expansion of both the similarity at FIH exercise and the blinded prospective exercise. A future prospective exercise could also be considered, similar to that currently being conducted to assess species usage for carcinogenicity studies (ICHS1, 2016). A hypothetical decision to use a single or two species is defined and justified ahead of testing and the data subsequently obtained from both species would be used to determine whether the correct decision would have been made from a reasonably sized database.

5. Concluding remarks and recommendations

The primary purpose of a nonclinical safety assessment during early drug development is to identify potential hazards for human risk assessment. While animal studies are currently embedded in the drug development process, the data presented herein provides evidence that under certain conditions single species toxicology programmes could be considered more widely, without being detrimental to human safety. However, this would require a major change in strategy, by some companies and/or regulatory agencies, especially for small molecule testing.

The goal of this industry survey was to gather data on the species used for repeat dose general toxicology studies conducted for human pharmaceuticals, using the current and relevant regulatory guidelines. The dataset covers small molecules and different classes of biologics across a wide range of therapeutic indications, many of which were still in active development. For the vast majority of small molecules, a two species (rodent and non-rodent) approach is typical, consistent with regulatory guidance (ICHM3 (R2) and ICHS9). The majority of single species approaches described were used to support testing of biologics (currently typically using the NHP), based on the recommendation that toxicology testing should only be conducted in a pharmacologically relevant species (ICHS6(R1)). The high specificity and species selectivity of biologics and associated lower likelihood of off-target toxicities, also contribute to acceptability of single species toxicology programmes. The dataset did contain mAbs for which two species approaches (rodent and non-rodent) were used, when cross-reactivity to multiple species was demonstrated. The data indicated that for many of these biologics, toxicities were similar between the species, indicating that a single species (preferably the rodent) may be sufficient to identify relevant toxicities in these cases. The dataset also contained drug candidates for other drug modalities for which a two species (rodent and non-rodent) approach is typical - therapeutic proteins/peptides and ADCs that have aspects of both small molecules and biologics and thus follow principles of both ICHM3(R2) and ICHS6(R1).

There were cases where, as supported by various guidelines, toxicology studies in two species were conducted for the FIH stage, with longer-term toxicity studies only being conducted in a single species. The data review by the working group showed that, with the benefit of hindsight and purposely working outside the relevant regulatory guidelines on species recommendations for individual drug modalities, it may have been possible for more of the drug candidates to also take this approach, without detrimental impact on human risk assessment. Generally, there was no trend towards one species being consistently more sensitive than another, which is in line with expectations given the wide range of targets, indications and pharmacologies likely represented in the dataset. However, the above conclusion was reached without access to the full datasets available for the drug candidates. Even within the working group, consensus was reached on the decision for each drug candidate only after much debate and differences of opinion. Ensuring that toxicities of potential human relevance have been adequately identified and the risks appropriately managed is

challenging and the level of discussion required in justification of potentially smaller packages of data may be a factor in the retention of two species approaches.

The challenge of designing the most appropriate toxicology testing programme for any novel human pharmaceutical is compounded by the lack of data-sharing between companies. In addition, toxicology study findings for terminated molecules rarely get published by the sponsoring company, an unfortunate situation, as such data (e.g., showing class effects) might influence the need for extensive nonclinical testing of subsequent similar drugs by other companies. This comprehensive evaluation of cross-industry and cross-drug modalities toxicology data enabled a retrospective evaluation of whether comparable decisions could have been reached on the toxicology profile of a new molecule using only one species (purposely working outside current regulatory guideline recommendations). From these data, the incidence of similar toxicities identified in both species at FIH stage indicated a wider number of drug modalities could potentially reduce to a single species for longer-term post-FIH studies. For those drug candidates where toxicities were different between species, or there was a lack of any effect in one of the species, there may still be opportunities for progression in a single species if additional information on the drug candidates (e.g., target and pharmacology) or significance of any differentially observed findings (i.e. target organ, severity) were available. In particular, where the difference was the absence of toxicities in one species but presence of toxicities in the other, the most-sensitive species could be the single species to progress. This does assume that aspects were equal (i.e. exposure levels etc) in the two species FIH studies, and that sensitivity in short term toxicity studies predicts the same differences in sensitivity for long term toxicity, which is not always the case, given that toxicities can resolve or new toxicities become apparent upon long-term dosing (Roberts et al., 2015).

Several additional issues need to be resolved before use of a single species can be routinely considered for all drug modalities after the FIH toxicology package is completed. It is acknowledged that for a few of the drug candidates in the survey (five small molecules), new toxicities were identified in the longer-term toxicity studies which would potentially have been missed had only a single species been tested. Other reviews of chronic (> 3-month dosing duration) toxicity studies from individual company portfolios indicate that additional target organ toxicities are occasionally revealed (Galijatovic-Idrizbegovic et al., 2016; Roberts et al., 2015) and that this information can lead to decisions to terminate late-stage drug candidates. Therefore, further work to establish a larger dataset would be necessary to provide scientific justification to adopt this flexibility for the use of a single species within longer-term post-FIH toxicology studies across a wider range of drug modalities, whilst still retaining the use of two species where both are necessary to fully identify the toxicity profile of an individual drug candidate. Additional guidance to define the criteria for similarity in toxicities between the species would be useful to facilitate and harmonise these decisions. In addition, criteria are needed on how to determine which species will be most appropriate to take into longer-term toxicity studies. Simple comparison of similarity in short-term studies does not always appear to be sufficient. High confidence that the species selected will detect all relevant long-term toxicity is needed to enable such single species programs.

Proposals to decrease the number of species used at the later-stages of the toxicology programme, or at any point within the drug development package, may be more acceptable/achievable in the future if more human-predictive data were available to support or even replace some of the animal data. The use of target safety evaluations early in drug discovery is an effective way to increase knowledge of safety risks from literature and databases, to reduce the need for *in vivo* studies and prioritise resources (Hornberg et al., 2014). It is hoped that the emergence of non-animal technologies (such as *in silico* modelling and *in vitro* human 3D-tissue models) could provide more relevant and predictive data and thus replace the need for data in a second species in the future (Ewart et al., 2018). In the meantime, continued sharing of experiences of different approaches to enable and support clinical drug development, and further international data-sharing and analysis activities (to also include other regions within ICH) are required to identify the criteria for justifying the use of a single species, or conversely, justifying the use of two species. Such further evidence may provide confidence that, when applied correctly, reducing animal toxicology packages would not be detrimental to human safety and in turn, continue to drive the most appropriate use of animals for regulatory general toxicology studies.

Funding body information

The role of the lead author (Helen Prior) is funded by the NC3Rs and the ABPI and the work described within the submitted manuscript was conducted in the course of the lead authors' employment.

The funding sources have had no role in the data collection, analysis or interpretation, or in the writing of the manuscript. The funding sources have approved the decision to submit this article for publication.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Acknowledgements

This article reflects the views of the authors and in particular should not be construed to represent views or policies of the FDA. At the time of initiation and conduct of the research/data gathering/analysis, K Jones was employed by Chugai Pharma Europe Ltd., and subsequently in the latter stages of the manuscript preparation is a consultant to Chugai Pharmaceutical., Japan.

The project was co-funded by member companies of the ABPI, working in partnership with the NC3Rs. ABPI represents the research-based pharmaceutical industry in the UK, developing strategy, monitoring policy and responding to consultations. Expert groups cover nonclinical issues in pharmaceutical research and drug development, in particular, focussing on discovery, preclinical safety, animal research and welfare.

The NC3Rs is an independent scientific organisation. It supports the UK science base by driving and funding innovations and technological developments that replace or reduce the need for animals in research and testing, and lead to improvements in welfare where animals continue to be used. The Centre promotes robust and ethical scientific practice through collaborating with research funders, academia, industry, regulators and animal welfare organisations, both in the UK and internationally.

References

Bailey, J., Thew, M., Balls, M., 2015. Predicting human drug toxicity and safety via animal tests: can any one species predict drug toxicity in any other, and do monkeys help? *Alternatives to laboratory animals*. ATLA 43, 393–403.

Baldrick, P., 2008. Safety evaluation to support First-In-Man investigations II: toxicology studies. Regul. Toxicol. Pharmacol. : RTP (Regul. Toxicol. Pharmacol.) 51, 237–243.

Baldrick, P., 2011. Safety evaluation of biological drugs: what are toxicology studies in primates telling us? Regul. Toxicol. Pharmacol. : RTP (Regul. Toxicol. Pharmacol.) 59, 227–236.

Baldrick, P., 2017. Getting a molecule into the clinic: nonclinical testing and starting dose considerations. Regul. Toxicol. Pharmacol. : RTP (Regul. Toxicol. Pharmacol.) 89, 95–100.

Beck, A., Goetsch, L., Dumontet, C., Corvaia, N., 2017. Strategies and challenges for the next generation of antibody-drug conjugates. Nat. Rev. Drug Discov. 16, 315–337.

Boran, T., Menezes-Ferreira, M., Reischl, I., Celis, P., Ferry, N., Gansbacher, B., Krafft, H.,

Lipucci Di Paola, M., Sladowski, D., Salmikangas, P., 2017. Clinical Development and Commercialization of Advanced Therapy Medicinal Products (ATMPs) in the EU: how are the product pipeline and regulatory framework evolving? Human gene therapy. Clinical development 28 (3), 126–135.

Brennan, F.R., Cavagnaro, J., McKeever, K., Schutten, M., Vahle, J., Weinbauer, G., Black, L.E., 2018. Safety testing of monoclonal antibodies in non-human primates: case studies highlighting their impact on risk assessment for humans. mAbs 10, 1–17.

Briggs, K., Barber, C., Cases, M., Marc, P., Steger-Hartmann, T., 2015. Value of shared preclinical safety studies - the eTOX database. Toxicol Rep 2, 210–221.

Brizmohun, N., 2019. Achilles on getting 'new wave' immunotherapy from concept to clinic in three years. Pink Sheet 15 March 2019.

Butler, L.D., Guzzie-Peck, P., Hartke, J., Bogdanffy, M.S., Will, Y., Diaz, D., Mortimer-Cassen, E., Derzi, M., Greene, N., DeGeorge, J.J., 2017. Current nonclinical testing paradigms in support of safe clinical trials: an IQ Consortium DruSafe perspective. Regul. Toxicol. Pharmacol. : RTP (Regul. Toxicol. Pharmacol.) 87 (Suppl. 3), S1–S15.

Chaikin, P., 2017. The bial 10-2474 phase 1 study-A drug development perspective and recommendations for future first-in-human trials. J. Clin. Pharmacol. 57, 690–703.

Colleton, C., Brewster, D., Chester, A., Clarke, D.O., Heining, P., Olaharski, A., Graziano, M., 2016. The use of minipigs for preclinical safety assessment by the pharmaceutical industry: results of an IQ DruSafe minipig survey. Toxicol. Pathol. 44, 458–466.

Detela, G., Lodge, A., 2019. EU regulatory pathways for ATMPs: standard, accelerated and adaptive pathways to marketing authorisation. *Molecular therapy*. Methods & clinical development 13, 205–232.

EMA, 2014. EU Guideline on similar biological medicinal products containing monoclonal antibodies – nonclinical and clinical issues (EMA/CHMP/BMWP/42832/2005 Rev1). https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-similar-biological-medicinal-products-containing-monoclonal-antibodies-non-clinical_en.pdf.

EMA, 2018. Guideline on quality, non-clinical and clinical requirements for investigational advanced therapy medicinal products in clinical trials (EMA/CAT/852602/2018). https://www.ema.europa.eu/en/documents/scientific-guideline/draft-guideline-quality-non-clinical-clinical-requirements-investigational-advanced-therapy_en.pdf.

Emanuel, E.J., Bedarida, G., Macci, K., Gabler, N.B., Rid, A., Wendler, D., 2015. Quantifying the risks of non-oncology phase I research in healthy volunteers: meta-analysis of phase I studies. BMJ (Clinical research ed.) 350 h3271.

English, V., 2011. Immunocore: A Regulatory First? *MedNous*. pp. 3–15 April 2011.

Ewart, L., Dehne, E.M., Fabre, K., Gibbs, S., Hickman, J., Hornberg, E., Ingelman-Sundberg, M., Jang, K.J., Jones, D.R., Lauschke, V.M., Marx, U., Mettetal, J.T., Pointon, A., Williams, D., Zimmermann, W.H., Newham, P., 2018. Application of microphysiological systems to enhance safety assessment in drug discovery. Annu. Rev. Pharmacol. Toxicol. 58, 65–82.

Galijatovic-Idrizbegovic, A., Miller, J.E., Cornell, W.D., Butler, J.A., Wollenberg, G.K., Sistare, F.D., DeGeorge, J.J., 2016. Role of chronic toxicology studies in revealing new toxicities. Regul. Toxicol. Pharmacol. : RTP (Regul. Toxicol. Pharmacol.) 82, 94–98.

Heining, P., Ruysschaert, T., 2016. The use of minipig in drug discovery and development: pros and cons of minipig selection and strategies to use as a preferred non-rodent species. Toxicol. Pathol. 44, 467–473.

Hinrichs, M.J., Dixit, R., 2015. Antibody drug conjugates: nonclinical safety considerations. AAPS J. 17, 1055–1064.

Hornberg, J.J., Laursen, M., Brenden, N., Persson, M., Thougaard, A.V., Toft, D.B., Mow, T., 2014. Exploratory toxicology as an integrated part of drug discovery. Part I: why and how. Drug Discov. Today 19, 1131–1136.

ICHM3(R2), June 2009). Nonclinical safety studies for the conduct of human clinical trials and marketing authorization for pharmaceuticals. In: International Conference on Harmonisation (ICH). Topic M3(R2).

ICHS1, 2016. Regulatory Notice on Changes to Core Guideline on Rodent Carcinogenicity Testing of Pharmaceuticals.

ICHS6(R1), June 2011. Preclinical safety evaluation of biotechnology-derived pharmaceuticals. In: International Conference on Harmonisation (ICH). Topic S6(R1).

ICHS9, March 2010. Nonclinical evaluation for anticancer pharmaceuticals. In: International Conference on Harmonisation (ICH). Topic S9.

Mangipudy, R., Burkhardt, J., Kadambi, V.J., 2014. Use of animals for toxicology testing is necessary to ensure patient safety in pharmaceutical development. Regul. Toxicol. Pharmacol. : RTP (Regul. Toxicol. Pharmacol.) 70, 439–441.

Monticello, T.M., Jones, T.W., Dambach, D.M., Potter, D.M., Bolt, M.W., Liu, M., Keller, D.A., Hart, T.K., Kadambi, V.J., 2017. Current nonclinical testing paradigm enables safe entry to First-In-Human clinical trials: the IQ consortium nonclinical to clinical translational database. Toxicol. Appl. Pharmacol. 334, 100–109.

Morgan, P., Brown, D.G., Lennard, S., Anderton, M.J., Barrett, J.C., Eriksson, U., Fidock, M., Hamren, B., Johnson, A., March, R.E., Matcham, J., Mettetal, J., Nicholls, D.J., Platz, S., Rees, S., Snowden, M.A., Pangalos, M.N., 2018. Impact of a five-dimensional framework on R&D productivity at AstraZeneca. Nat. Rev. Drug Discov. 17, 167–181.

Mustonen, E.K., Palomaki, T., Pasanen, M., 2017. Oligonucleotide-based pharmaceuticals: non-clinical and clinical safety signals and non-clinical testing strategies. Regul. Toxicol. Pharmacol. : RTP (Regul. Toxicol. Pharmacol.) 90, 328–341.

Olson, H., Betton, G., Robinson, D., Thomas, K., Monro, A., Kolaja, G., Lilly, P., Sanders, J., Sipes, G., Bracken, W., Dorato, M., Van Deun, K., Smith, P., Berger, B., Heller, A., 2000. Concordance of the toxicity of pharmaceuticals in humans and in animals. Regul. Toxicol. Pharmacol. : RTP (Regul. Toxicol. Pharmacol.) 32, 56–67.

Prior, H., Baldrick, P., deHaan, L., Downes, N., Jones, K., Mortimer-Cassen, E., Kimber, I., 2018. Reviewing the utility of two species in general toxicology related to drug development. Int. J. Toxicol. 37, 121–124.

Roberts, R., Callander, R., Duffy, P., Jacobsen, M., Knight, R., Boobis, A., 2015. Target organ profiles in toxicity studies supporting human dosing: does severity progress with longer duration of exposure? Regul. Toxicol. Pharmacol. : RTP (Regul. Toxicol. Pharmacol.) 73, 737–746.

Roberts, R.A., Kavanagh, S.L., Mellor, H.R., Pollard, C.E., Robinson, S., Platz, S.J., 2014. Reducing attrition in drug development: smart loading preclinical safety assessment. Drug Discov. Today 19, 341–347.

Sewell, F., Chapman, K., Couch, J., Dempster, M., Heidel, S., Loberg, L., Maier, C., Maclachlan, T.K., Todd, M., van der Laan, J.W., 2017. Challenges and opportunities for the future of monoclonal antibody development: improving safety assessment and reducing animal use. mAbs 9, 742–755.

Stevenson, J.G., Popovian, R., Jacobs, I., Hurst, S., Shane, L.G., 2017. Biosimilars: practical considerations for pharmacists. Ann. Pharmacother. 51, 590–602.

Tamaki, C., Nagayama, T., Hashiba, M., Fujiyoshi, M., Hizue, M., Kodaira, H., Nishida, M., Suzuki, K., Takashima, Y., Ogino, Y., Yasugi, D., Yoneta, Y., Hisada, S., Ohkura, T., Nakamura, K., 2013. Potentials and limitations of nonclinical safety assessment for predicting clinical adverse drug reactions: correlation analysis of 142 approved drugs in Japan. J. Toxicol. Sci. 38, 581–598.

van Meer, P.J., Kooijman, M., Gispen-de Wied, C.C., Moors, E.H., Schellekens, H., 2012. The ability of animal studies to detect serious post marketing adverse events is limited. Regul. Toxicol. Pharmacol. : RTP (Regul. Toxicol. Pharmacol.) 64, 345–349.

Waring, M.J., Arrowsmith, J., Leach, A.R., Leeson, P.D., Mandrell, S., Owen, R.M., Pairaudeau, G., Pennie, W.D., Pickett, S.D., Wang, J., Wallace, O., Weir, A., 2015. An analysis of the attrition of drug candidates from four major pharmaceutical companies. Nat. Rev. Drug Discov. 14, 475–486.