

Opportunities to Apply the 3Rs in Safety Assessment Programs

Fiona Sewell, Joanna Edwards, Helen Prior, and Sally Robinson

Fiona Sewell, PhD, is a Programme Manager in Toxicology and Regulatory Sciences at the UK National Centre for the Replacement, Refinement and Reduction of Animals in Research (NC3Rs), London, UK. Joanna Edwards, PhD, is a Programme Manager in Technology Development at the NC3Rs, London, UK. Helen Prior, PhD, is a Programme Manager in Drug Development at the NC3Rs, London, UK. Sally Robinson, PhD, is Head of Laboratory Animal Sciences at AstraZeneca, Alderley Park, UK.

Address correspondence and reprint requests to Dr. Fiona Sewell, NC3Rs, Gibbs Building, 215 Euston Rd, London, NW1 2BE, UK or email fiona.sewell@nc3rs.org.uk.

Abstract

Before a potential new medicine can be administered to humans it is essential that its safety is adequately assessed. Safety assessment in animals forms an integral part of this process, from early drug discovery and initial candidate selection to the program of recommended regulatory tests in animals. The 3Rs (replacement, reduction, and refinement of animals in research) are integrated in the current regulatory requirements and expectations and, in the EU, provide a legal and ethical framework for in vivo research to ensure the scientific objectives are met whilst minimizing animal use and maintaining high animal welfare standards. Though the regulations are designed to uncover potential risks, they are intended to be flexible, so that the most appropriate approach can be taken for an individual product. This article outlines current and future opportunities to apply the 3Rs in safety assessment programs for pharmaceuticals, and the potential (scientific, financial, and ethical) benefits to the industry, across the drug discovery and development process. For example, improvements to, or the development of, novel, early screens (e.g., in vitro, in silico, or nonmammalian screens) designed to identify compounds with undesirable characteristics earlier in development have the potential to reduce late-stage attrition by improving the selection of compounds that require regulatory testing in animals. Opportunities also exist within the current regulatory framework to simultaneously reduce and/or refine animal use and improve scientific outcomes through improvements to technical procedures and/or adjustments to study designs. It is important that approaches to safety assessment are continuously reviewed and challenged to ensure they are science-driven and predictive of relevant effects in humans.

Key words: 3Rs; drug development; pharmaceuticals, safety assessment; reduction; refinement; replacement

Introduction

It is a scientific, ethical, and regulatory requirement that before any potential new medicine can be administered to humans its safety must be adequately assessed. The assessment may

include a variety of in vitro, ex vivo, and in silico approaches and almost always includes in vivo tests in both rodent and/or nonrodent species. The conduct of safety studies in animals is highly regulated, and the pharmaceutical industry recognizes the need

to continuously reassess and challenge the design and implementation of such studies so they are performed to the most up-to-date scientific knowledge and procedures. The 3Rs (replacement, reduction, and refinement), first described by Russell and Burch in 1959 (Russell and Burch 1959), currently provide a legal and ethical framework for *in vivo* research in the EU. Their consideration and implementation can not only improve animal welfare but also offer scientific and business benefits through reduced costs and improved efficiency (Table 1).

Drug development is a long and costly process, and it is therefore desirable to avoid wasting money, resources, and animals on drug candidates that are not suitable for development as potential medicines. This can be achieved by ensuring early attrition of unsuitable compounds. By early involvement of safety assessment in drug discovery, studies can be performed to identify overt toxicities and thus eliminate unsuitable candidates before significant investment is made in the compound. A survey from 2004 (Kola and Landis 2004) indicated success rates from first-in-human (FIH) studies to registration (over the period 1991–2000) was only one in nine compounds (11%), with the greatest reasons for failure being efficacy (~30%) and safety (toxicology and clinical safety, ~30%). A decade later (data from 2003–2011), there was no change, with approximately 1 in 10 compounds (10%) that entered FIH studies achieving approval by the US Food and Drug Administration (Hay et al. 2014), with the same reasons for program suspensions. From a safety perspective, the major causes of attrition throughout the development pipeline are cardiovascular and liver toxicities leading to the need for the development of screens (animal and nonanimal) that are more predictive of these liabilities (Cook et al. 2014; Hornberg et al. 2014; Redfern et al. 2010; Waring et al. 2015).

Before regulatory authorities approve administration of potential new pharmaceuticals (conventional pharmaceuticals or biopharmaceuticals) or diagnostic agents to humans, or permit marketing authorizations, they generally require that the safety of the drug candidate has been assessed in animals. Animal use within the drug discovery and development process occurs in a logical and step-wise manner. Pharmacokinetic and pharmacodynamic properties are first defined, followed by efficacy studies before moving into candidate selection and regulatory safety testing in animals. This article concentrates on the use of animals for safety assessment within the drug discovery and development process.

Early Screens and Candidate Selection

Early *ex vivo*, *in vitro*, or *in silico* screens that focus on better candidate selection can reduce the number of compounds that enter the cascade of regulatory tests in animals. Even small improvements against the current >90% failure rate (Hay et al. 2014; Kola and Landis 2004) would cut down on the number of animals used to develop compounds that fail in the clinic and

could have a substantial impact on overall use of animals in pharmaceutical development. Stopping compounds earlier and/or selecting the compounds most likely to be successful in the drug discovery/development process reduces late-stage attrition but also contributes to an overall reduction in animal use (by stopping compound progression before the *in vivo* testing stage) and/or refinement (as the compounds that do progress may have fewer adverse effects than compounds with potential problems that may be stopped in earlier screens). There is a rapidly developing field of *in vitro* models and *in silico* tools for modelling that can be utilized during drug discovery to replace the need for *in vivo* testing. Some examples with 3Rs potential are described below.

Use of In Vitro Screens and/or In Silico Tools

A recent survey (Goh et al. 2015) reported a steady increase in the use of *in vitro* tests by the pharmaceutical industry between 1980 and 2013, with 99% directed towards absorption, distribution, metabolism, excretion (ADME), safety pharmacology, and genotoxicity endpoints, presumably in response to scientific and technological advances in these areas over this period. As these tests, such as Ames (for DNA mutational risk assessment) and hERG (cell lines transfected with *human ether-à-go-go related gene* for cardiovascular arrhythmia risk assessment), are commonly used, they are not addressed further within this article (but the reader is directed towards Roth and Singer 2014 for a review of these and potential new technologies). The use of pharmacological profiling, the screening of compounds against a broad range of targets (e.g., receptors, ion channels, enzymes, and transporters) that are distinct from the intended therapeutic target(s), is also widely used and can identify specific molecular interactions that could cause adverse reactions in humans (Bowes et al. 2012).

A more mechanistic approach, employing the Adverse Outcome Pathway (AOP) concept, has the potential to improve safety assessment and reduce reliance on animal methods through the development of new and more predictive safety assessment processes that could help support go/no-go decisions (Burden et al. 2015). AOPs link a molecular initiating event (this could be the intended drug target, or an unintended off-target event) to an apical endpoint (which could be the anticipated therapeutic effect or an unexpected side-effect) through a series of scientifically proven, causally linked events. Knowledge contained within AOPs could therefore be used to explain an adverse effect observed in *in vivo* safety studies, but ultimately this could be used to develop nonanimal methods (*in vitro* or *in silico*) that could be used in place of safety assessment in animals to predict the likely adverse outcome based on the known primary site of action/pharmacology as well as off-target effects.

Table 1. The definition of the 3Rs

	Standard definition ^a	Contemporary definition
Replacement	Nonanimal methods	Accelerating the development and use of human-relevant tools based on latest technologies
Reduction	Minimum number of animals consistent with scientific aims	Appropriately designed and considered animal experiments that are robust and reproducible
Refinement	Minimum pain, suffering, distress or lasting harm	New <i>in vivo</i> technologies that can benefit animal welfare and science

^aRussell and Burch 1959.

Predictive Mathematical/Computational Modelling

A number of reviews have outlined the potential uses of in silico technologies in toxicology testing (Hartung and Hoffmann 2009; Pelkonen et al. 2011; Raunio 2011; Valerio 2009, 2011). Many of these methods (e.g., data/power analyses) may be used as part of an integrative approach in combination with in vivo or in vitro tests (e.g., analysis of data from the results of in vivo experiments). For example, physiologically-based pharmacokinetic (PBPK) modelling can be used to predict human internal exposure and thus be used to set relevant doses for animal toxicity studies or to inform concentration setting for in vitro assays to enable better in vitro to in vivo extrapolation (Jones et al. 2015). However, there are so-called nontesting in silico methods that do not involve any physical experiments and do not require use of the drug candidate itself, such as computer modelling and structure activity relationships (SAR) or quantitative (Q)SAR, which could be used to help develop and build new more predictive in silico/computational models in the future. The development of such technologies requires input from large, good-quality databases and could make use of the wealth of data held within pharmaceutical companies and in public databases. For example, there have been multidisciplinary efforts to better utilize existing data from in vitro, in vivo, and ex vivo experiments to develop more human-relevant computational approaches for cardiovascular disease, which could presumably also be used to predict cardiovascular liabilities earlier in drug development (Rodriguez et al. 2016). Furthermore, mechanistic information held within AOPs could be used in conjunction with structural alerts from (Q)SAR models, for example, to identify and prevent drug candidates with unacceptable safety profiles from progressing through the pharmaceutical development pipeline.

Human Tissue Models

The use of human tissue and/or 3D models such as organ-on-chip, spheroid models, and human induced pluripotent stem cells (iPSC) are becoming more widely used for safety assessment purposes in the fields of safety pharmacology and toxicology (Holmes et al. 2015; Roth and Singer 2014).

Human iPSCs have the capability to differentiate into a large range of specific tissues and could be used as relevant and predictive early screens. For example, assays based on iPSC-derived cardiomyocytes may complement or replace currently used assays based either on primary cardiomyocytes from animals or cell-lines overexpressing ion channels (Sinnecker et al. 2014). The measurement of the field potential duration from iPSC-derived cardiomyocytes by microelectrode arrays is being evaluated as part of the Comprehensive in vitro Proarrhythmia Assay (CiPA) initiative (Cavero and Holzgrefe 2014). This international consortium of industry and regulators has the objective to engineer (early in the drug discovery and development process) assays allowing the evaluation of the proarrhythmic risk of compounds by studying drug effects on multiple ion channels (not limited to hERG) and incorporating these effects into an in silico model of a human ventricular action potential (Fermini et al. 2016). If successful, the evaluation of proarrhythmic risks will be moved to earlier in the development process, allowing for removal of compounds with undesirable effects on cardiac repolarisation and alleviating the risk and cost of subsequent human trials. These techniques also provide opportunities for other

areas of toxicity such as hepatotoxicity and neurotoxicity (McGivern and Ebert 2014).

Bespoke Early Models

Studies In Alternative (Invertebrate) Species

Whilst in vitro screens using cells and tissues are useful, and organ-on-a-chip models are incorporating more sophisticated interplays between systems, they may not fully represent the in vivo situation due to the more complex interplay between systems and processes at the whole animal level. Use of invertebrate species such as social amoeba (*Dictyostelium*), fruit flies (*Drosophila*), and nematodes (*C. elegans*) are increasingly used as alternative early options prior to studies using vertebrate species and thus offer a refinement (Cocorocchio et al. 2016; Strange 2016; Kwok et al. 2006; Willoughby et al. 2013).

Studies In Nonmammalian (Vertebrate) Species

The use of lower sentient vertebrates is considered a refinement over higher sentient vertebrates, since the ability to suffer would be less. The zebrafish (*Danio rerio*) in particular has gained in popularity for screening early in the testing cascade, as its small size and transparent body allows testing within a 96-well format, combining the scale and throughput of in vitro systems with the physiological complexity of vertebrate whole animal research (Garcia et al. 2016). A wide range of such tests has been developed in the area of safety pharmacology (Redfern et al. 2008) and reproductive toxicity (He et al. 2014), amongst others.

Early Studies in Mammalian (Vertebrate) Species

Models in alternative (invertebrate and nonmammalian vertebrate) species can be useful as early screens; however, strategies for conducting hypothesis-driven in vivo studies in the classical toxicology rodent species earlier in discovery have also been published (Bass et al. 2009; Hornberg et al. 2014; Roberts et al. 2014). These studies may be designed to address potential on-target or off-target issues with the candidate drug. Data generated from studies performed in support of discovery can help inform the decision on whether to develop a candidate drug. For compounds that progress, information from these studies may be included in regulatory submissions as supporting data. Early indications of toxicity liabilities will allow early detection of unsuitable candidates, so that only the candidates that are more likely to progress are taken forward, thus reducing late stage attrition and associated development cost and animal use. Studies performed on “tool” compounds can help to improve candidate selection and reduce the resources and animals used (e.g., by early identification of candidates from a pharmacological class or chemical series that has inherent toxicology liabilities associated with it). Data generated from such studies will feed back into programmes for future compounds to refine the discovery and development strategy and thus avoid unnecessary animal use and expenditure.

Regulatory (GLP Standard) Studies

Data from animal studies are used to characterize potential safety risks to humans and to help determine a safe starting dose for FIH clinical trials. The general requirements for the conduct of toxicology, safety pharmacology, and other associated studies are outlined within regulatory guidances

Table 2. International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) Guidelines

Subject area	Short name	Reference
Nonclinical safety studies	ICH M3 (R2)	Nonclinical Safety Studies for the Conduct of Human Clinical Trials and Marketing Authorization for Pharmaceuticals. International Conference on Harmonisation (ICH). Topic M3(R2): June 2009
	ICH M3 (R2) Q&A	Nonclinical Safety Studies for the Conduct of Human Clinical Trials and Marketing Authorization for Pharmaceuticals. European Medicines Agency. Committee for medicinal products for human use (CHMP). ICH guidelines M3(R2) Q&A. July 2011
Toxicokinetics and Pharmacokinetics	ICH S3A	Note for Guidance on Toxicokinetics: The Assessment of Systemic Exposure in Toxicity Studies. International Conference on Harmonisation (ICH). Topic S3: October 1994
	ICH S3A Q&A	Note for Guidance on Toxicokinetics: The Assessment of Systemic Exposure in Toxicity Studies. Questions and Answers. S3A Implementation Working Group. Draft ICH Consensus Guideline. Step 2. January 2016.
Reproductive toxicity	ICH S5(R2)	Direction of toxicity to reproduction for medicinal products and toxicity to male fertility. International Conference on Harmonisation (ICH). Topic S5(R2): November 2005
Biotechnology-derived pharmaceuticals	ICH S6 (R1)	Preclinical Safety Evaluation of Biotechnology-Derived Pharmaceuticals. International Conference on Harmonisation (ICH). Topic S6(R1). June 2011
Safety pharmacological studies	ICH S7A	Safety pharmacological studies for human pharmaceuticals. International Conference on Harmonisation (ICH). Topic S7A. November 2000
Delayed ventricular repolarization (QT interval prolongation)	ICH S7B	The Nonclinical Evaluation of the Potential for Delayed Ventricular Repolarisation (QT Interval Prolongation) by Human Pharmaceuticals. International Conference on Harmonisation (ICH). Topic S7A. May 2005
Anticancer pharmaceuticals	ICH S9	Nonclinical Evaluation for Anticancer Pharmaceuticals. International Conference on Harmonisation (ICH). Topic S9. March 2010

coordinated under the auspices of the International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) (Table 2) and must be carried out according to Good Laboratory Practice (GLP) standards. Although this has harmonized study requirements and designs within the included regions (currently EU, Japan, US, Canada, and Switzerland), flexibility exists within the guidelines such that the most appropriate approach can be taken for an individual drug candidate. Consequently, different perceptions and interpretation of requirements can lead to variations in the number of animals used for similar studies, which highlights the value of cross-company sharing of study designs so that good practice can be learned and shared (Chapman et al. 2009, 2016; Sewell et al. 2014; Sparrow et al. 2011). Additionally, as companies are aiming for global registration, data packages often reflect the needs (or perceived expectation) of the regulatory authority that requires the most data.

Challenging the Regulatory Requirements

The ICH guidelines are continually reviewed and revised to incorporate recent scientific developments, reflect change in practice, and add new topics, as well as to clarify and provide additional guidance or support. Therefore, a forum exists to question regulatory requirements that may have become redundant and no longer add value. For example, the latest ICH M3(R2), which primarily refers to small molecules, includes removal of the requirement for stand-alone acute toxicity testing, a reduced maximum duration for chronic studies in nonrodents (9 months rather than 12 months), harmonization of the criteria for selecting the high dose in toxicity studies to allow a reduced limit dose (2 g/kg to 1 g/kg), and integration of bone marrow micronucleus endpoints in rodent general toxicity studies (ICH, M3(R2); Ledwith and DeGeorge 2011). For biotechnology-derived products, the ICH S6(R1) addendum clarifies that 6-month toxicity testing, usually in a single

species, is the maximum needed for biologics. Additionally, global cross-company collaborations under the auspices of the International Life Sciences Institute-Health and Environmental Sciences Institute (ILSI-HESI), the NC3Rs (UK National Centre for the Replacement, Refinement and Reduction of Animals), industry consortia (such as the European Federation of Pharmaceutical Industries and Associations (EFPIA) and the International Consortium for Innovation and Quality in Pharmaceutical Development), and international scientific organizations such as the Society of Toxicology and Safety Pharmacology Society actively discuss the need for new scientific approaches and work together to influence regulatory change and implement 3Rs improvements, as demonstrated in examples below.

Species Relevance

Regulatory authorities usually require safety and tolerability data from both a rodent and a nonrodent species before progressing to FIH clinical trials (ICH, M3(R2)). It is expected that pharmacological or adverse effects observed in a test species may also occur in humans, and it has been previously demonstrated that the use of two phylogenetically separate animal species will increase the likelihood of detection of adverse effects (Olson et al. 2000). The species selected should be based on a pharmacokinetic and metabolism profile similar to humans and relevant pharmacology (i.e., the target has a similar role to that in humans), though for biologics, such as monoclonal antibodies (mAbs), species selection is usually based on the results of in vitro biological assays. The choice of species is generally the rat and dog (Baldrick 2008; Horner et al. 2013), although the minipig (Colleton et al. 2016) and nonhuman primate (NHP) are also used when relevant; the latter usually for biologics testing due to the highly specific human target (Chapman et al. 2009). Reviews of nonclinical data have shown that the nonrodent data identify additional toxicities to that

detected in rodents, thus providing additional concordance and predictivity for adverse events identified in humans (Horner et al. 2013; Olson et al. 2000; Tamaki et al. 2013). However, with appropriate justification, it can be relevant for some packages (particularly for small molecules in areas with unmet medical need such as some cancers) to submit rodent-only data (Newell et al. 1999) (ICH, S9) or chronic-dosing studies in a single rodent species for large molecules (Chapman et al. 2013) (ICH, S6(R1)). This not only reduces the number of animals used but provides efficiencies in costs and resource, and potentially accelerates the availability of new medicines to humans. Across the industry, a number of international cross-company data-sharing initiatives are ongoing (Monticello 2015) or in discussion (EFPIA, MHRA, NC3Rs) to provide large databases of information on species relevance within toxicology studies and their predictivity to clinical effects. Generation of evidence bases may provide opportunities to enhance toxicology programs and reduce the use of animals in the future. For mAb biosimilar products, the need for in vivo studies has been questioned altogether (Chapman et al. 2016; van Aerts et al. 2014; van Meer et al. 2015), and the EU regulatory guidelines provide a path that allows the development and authorization of biosimilars with the submission of nonclinical in vitro data alone in certain circumstances (EMA 2014a, 2014b).

Acute Toxicity Tests

Stand-alone acute toxicity studies in animals are no longer required in pharmaceutical development for countries following ICH guidelines, due to its removal from ICH M3(R2) in 2010. They were historically required to support the registration of any new medicine and to identify doses that caused major adverse effects and the minimum dose causing lethality. The stated reasons for these studies were to set doses for further nonclinical studies and to support FIH clinical trials and evaluation of the effects of overdose in humans. However, in 2003 a cross-company initiative led by the NC3Rs and AstraZeneca questioned the value of these studies, since it was thought that more useful information for dose setting was already gained from other nonclinical studies (e.g., from short-term maximum tolerated dose “MTD” studies). The working group shared information on study objectives, design, timing, and data on outcome and recommended that acute toxicity studies should not be required prior to FIH, and that any short-term or dose-escalation data that is currently used for dose setting in other animal studies should be acceptable to allow assessment of acute toxicity and data should be provided by the clinical route only (original requirement was for two routes of administration). These were presented to regulators from the EU, US, and Japan and led to the revision of ICH M3 (ICH, M3(R2)) and the removal of the Committee for Medicinal Products for Human Use (CHMP) guideline on single dose/acute toxicity (EMA 2010b; EMEA 2008; EudralexVol3B), where specific reference was made to the working group’s evidence base and publication (Robinson et al. 2008). Further work concluded that acute toxicity studies do not provide useful information to support the clinical management of overdose in humans once the medicine is marketed (Chapman et al. 2010; Robinson and Chapman 2009). The impact of this project is evidenced through the proportion of clinical trial applications for drugs going into humans for the first time in the UK, which contain the results from acute toxicity tests reducing from 86% in 2007 to 8% in 2014 (unpublished data supplied by MHRA).

Opportunities within the Current Regulatory Framework

Minimized Study Designs

There are opportunities to implement the 3Rs within the current regulatory framework. For example, ICH M3(R2) supports a number of approaches for exploratory clinical trials (e.g., microdosing) that require a reduced number of safety studies in animals (EMA 2003; Garner 2010; ICH M3(R2)). Approaches such as micro-sampling allow very specific questions to be assessed in humans (e.g., confirmation of PK profile) and can be useful in making early go/no-go decisions (see later section on microsampling). An early ‘stop’ decision at this stage would have required fewer animals than progressing to standard FIH studies. However, compounds progressing further would then require the normal safety assessment package to be conducted. Currently there is little guidance on study design for parameters such as appropriate group size or the number of dose groups required. Along with differences in approaches to the use of additional animals, such as the inclusion of toxicokinetic (TK) satellites or off-treatment recovery animals, there can be variability in the total number of animals used to meet the same regulatory requirement. Practices have been reviewed in the past, and recommendations for minimal study designs have been published (Baldrick 2008, 2011; Chapman et al. 2009; Sparrow et al. 2011). For example, for biologics it may be possible to move away from the small molecule approach of control plus three dose groups for main study animals to instead use one or two relevant dose levels plus control (Baldrick 2011). However, the biggest opportunity to reduce the number of rodents used has been highlighted as reducing the use of satellite animals for TK sampling (Sparrow et al. 2011). The introduction of microsampling has also allowed TK samples to be obtained from the main study animals directly, reducing the need for separate satellite animal groups (Chapman et al. 2014a, 2014b; ICH S3A Q&A). However, there are still opportunities to promote this approach and reduce animal use further through more subtle changes to the study design.

Group Size

Though the published recommended study designs based on cross-company data sharing include suggestions for group sizes (Chapman et al. 2009; Prior et al. 2016; Sparrow et al. 2011), the general regulations relating to pharmaceutical development themselves do not specify preferred group size. The only ICH guideline to suggest a specific dosing group size is ICH S9 for oncology indications, which suggests group sizes of “at least 3 animals/sex/group, with an additional 2/sex/group for recovery” for nonrodents. Though ICH S5 discusses the number of litters that should be evaluated for reproductive studies, stating that “the number of animals per sex per group should be sufficient to allow meaningful interpretation of the data,” it recognizes that there is “very little scientific basis underlying specified group sizes in past and existing guidelines nor in this one” (ICH, S5 (R3)). Typically, the published literature recommends groups sizes for general toxicology studies up to 3 months in duration of 3M+3F for nonrodents and 10M+10F for rodents. Though these minimal study designs appear to be acceptable for the shorter-term studies to support FIH clinical trials, larger group sizes (e.g., 4M+4F for nonrodents) are included in longer-term animal studies (6 months) that are designed to support human clinical trials later in development to alleviate the possibility of an animal having to be removed from the study during the longer treatment period. These studies are also likely to detect

more subtle chronic effects that may require more animals to assess. However, there are no published statistical analyses to support these small changes in group size.

Recovery Animals

It is a regulatory expectation that recovery from adverse effects is assessed at some point during the drug development process, but this is not necessarily required prior to FIH and may not require the use of dedicated recovery phase animals. Scientific assessment through the use of literature; previous knowledge or experience may suffice (ICH M3(R2) Q&A). Regulatory guidelines provide flexibility for adapting the off-dose recovery phase to the needs of the program and offer limited direction as to how, where, and when recovery should be included (ICH M3(R2), M3(R2) Q&A, S6(R1), S9). The ICH M3(R2) Q&A document and other publications (Horner et al. 2014; Sewell et al. 2014) provide additional guidance on when inclusion of recovery phase animals may (or may not) be warranted.

A recent global cross-company initiative led by the NC3Rs in collaboration with the MHRA examined the value of recovery phase animals in studies to support FIH for both small molecules and biologics (Sewell et al. 2014). The decision to include recovery animals was primarily driven by a regulatory expectation or standard company approach (which may be based on individual company rationale), yet the data obtained rarely had an impact on internal or regulatory decision making. Absence of recovery data had little or no impact on regulatory submission; compounds that did not include any recovery animals in any study to support FIH were still able to enter clinical trials. The experts concluded that recovery should not be included by default and should only be included for scientific reasons. Inclusion should be considered across the whole development package, and whilst it could be argued that animal numbers could be reduced through early assessment of reversibility, it may be more appropriate to include the recovery assessment later in development once more information on the toxicological profile is known. This recommendation is supported by an analysis of 77 candidate drugs from AstraZeneca that showed that the majority (>86%) of lesions fully or partially resolved by the end of the recovery period in studies to support FIH, but that additional toxicities were identified in 39% of the longer-term chronic studies (Horner et al. 2014).

Where an assessment requiring the use of recovery animals is warranted, typically suggested study designs include recovery animals in the control and high dose only (Baldrick 2011; Chapman et al. 2012; Sparrow et al. 2011). Though one dose group should suffice, it is important that a *relevant* dose group is used, and this may not always be the highest dose group (e.g., low dose may be more appropriate for biologics to avoid oversaturation of the target). However, it can also be questioned whether a control group is always necessary, since the purpose of the recovery groups is to assess recovery from *treatment-related* effects that requires comparison between main study animals with recovery animals of the same dose (Konigsson 2010; Sewell et al. 2014; Tomlinson et al. 2016). This may be more applicable to studies with short recovery periods and with sexually mature nonrodents, where there is minimal risk of age-related phenotypic drifts.

Incorporation of Multiple Endpoints

The regulatory guidelines for both biopharmaceuticals and anticancer pharmaceuticals (ICH S6(R1), S9) encourage the

incorporation of safety pharmacology measurements into toxicology studies. This can reduce the overall number of animals used compared to stand-alone studies but can also provide more data from the same animals to allow direct interpretation of results (e.g., pharmacokinetic/pharmacodynamic profiles). This approach is also increasingly being adopted for small molecule pharmaceuticals (Authier et al. 2013), where the ability to investigate effects after repeat-dosing is considered an additional advantage. The methods suitable for inclusion within toxicology studies are comprehensively reviewed by Redfern et al. (2013), with the most widely implemented technologies/methods including jacketed telemetry for cardiovascular assessment (Kaiser et al. 2015; Prior et al. 2009) and Functional Observational Battery for neurobehavioural assessments (Moscardo et al. 2009, 2010; Moser et al. 1997).

The timing and design of developmental and reproductive toxicity studies are conducive to combination with, or incorporation into, other studies, and a number of suggestions have been proposed for reducing the number of animals used (Chapman et al. 2013b; ICH S5(R3)). For example, for small molecules, combining the male and female fertility study and the embryo fetal development (EFD) study can reduce rodent usage by 20% per compound. For biologics, the most recent ICH S6(R1) advocates the use of an enhanced pre/postnatal development study (ePPND) for NHPs, which includes dosing from day 20 of gestation to birth to combine EFD and pre/postnatal development (PPND) endpoints in a single study. Inclusion of male and female fertility endpoints into the standard 28-day or longer general toxicity studies can even eliminate the need for a separate fertility study altogether. For compounds with low toxicity or low systemic exposures, male and female fertility, EFD, and PPND studies may be incorporated into a single study (ICH S5(R3)), reducing animal use by 50%. The ICH S5 update concept paper (ICH S5(R3)) also suggests a number of ways to enhance human risk assessment, particularly for compounds with long half-lives (i.e., biologics) whilst contributing substantially to reduction in animal usage.

It has been suggested that PPND endpoints are incorporated with juvenile animal studies so that they are no longer carried out as stand-alone studies. The paper also considers whether EFD testing is needed in two mammalian species and whether in some cases it may be sufficient to conduct EFD studies in a single species, if supported by data from other test systems.

Severity

Dose Selection

Since the objective of toxicology studies in animals is to identify potential toxic effects in humans, some of the animals used may suffer adverse effects. There are five general criteria for defining the high dose in a toxicology study. These are (1) maximum tolerated dose (MTD), (2) limit dose, (3) top dose based on saturation of exposure, (4) maximum feasible/practical dose, or (5) dose providing a 50-fold margin of exposure. For a full description of the options for selecting the high dose in general toxicity studies, see ICH guidance M3 (R2). Careful consideration should be given to dose selection so that the impact on the animal can be minimized while still achieving the scientific objective of the study. The current CHMP guidance on repeated dose toxicity studies indicates that doses should be selected to establish a dose- or exposure-response to treatment (EMA 2010a). This can generally be achieved by the use of three groups of animals receiving the test item, at low, intermediate,

and high doses, plus a control group that receives vehicle alone. Experience has shown that three appropriately chosen doses will usually cover the span between no effect and adverse effects, although there are exceptions, and sometimes more dose levels will be required or in very specific cases (e.g., for some large molecules), fewer dose levels may suffice. The CHMP guideline also indicates that the high dose should be selected to enable identification of target organ toxicity, or other nonspecific toxicity, or until limited by volume or limit dose (EMA 2010a). In addition to establishing toxicity, it is necessary from a scientific perspective to establish the no observed effect level (NOEL) and/or the no observed adverse effect level (NOAEL) that may be used along with other information, such as the pharmacologically active dose, to determine the first dose in human studies. Determining an appropriate dose therefore requires relevant experience and judgement, as it is often influenced by the nature of the test item, its target pharmacology, and its intended therapeutic use in humans. Selecting a dose that is too high or selecting a dose that does not produce toxicity may risk repetition of the study, thus requiring the use of additional animals. It may also prevent identification of target organs and early indicators that can be used to monitor potential effects in human studies. There is an inherent risk that selecting doses from initial studies using small numbers of animals or short dosing duration may not predict what happens when larger numbers of animals or longer dosing durations are used in the subsequent regulatory studies. Dose level selection and staggered dosing approaches in safety studies offers a significant opportunity for refinement.

A group of toxicologists in collaboration with NC3Rs and the Laboratory Animal Science Association (LASA) have shared practical advice for study directors and other toxicologists working in the field of regulatory toxicology to maximize the implementation of refinement in dose level selection for regulatory toxicology studies (NC3Rs/LASA 2009). The guidance document is available on the NC3Rs website (www.nc3rs.org.uk) and is aimed specifically at scientists that are new to, or training in, the role of study director. Drawing on the knowledge and experience of the working group members, the guidance document is intended to supplement the process of training and mentoring of study directors to improve the scientific outcome of general regulatory toxicology studies and to promote the application of the 3Rs. The guidance provided has the potential to make substantial progress in reducing and refining animal use in this area, through avoiding unnecessary exposure of animals to marked adverse effects, thereby reducing inadvertent morbidity and mortality and avoiding potential repetition of toxicology studies.

Maximum Tolerated Dose (MTD) Studies

Short-term toxicity studies are used in pharmaceutical development to set doses for longer term toxicology studies in animals and to determine safe starting doses for FIH. Data from these studies are lower severity alternatives to conventional single-dose toxicology studies, because they avoid death as an endpoint (Robinson et al. 2008). Such studies provide information on adverse effects that are observed at the MTD, the highest dose that will be tolerated within a given study for the study duration.

Defining the MTD in the studies of shortest duration informs dose setting in subsequent toxicity studies and is crucially important in the application of the 3Rs, since it reduces the chances of the larger numbers of animals that are used in regulatory studies being exposed to unanticipated pain and

distress (NC3Rs/LASA 2009). The MTD is usually determined by parameters such as clinical signs and reductions in body weight and food consumption. However, there are limited published criteria or guidance on the intensity and duration of clinical signs that would optimise the selection of an MTD, especially in studies of short duration (Chapman et al. 2013a; FELASA 1994; Morton 2000; NC3Rs/LASA 2009). Though body weight loss (BWL) is an objective measurement and is often used as a primary endpoint in these studies, there is no industry agreement on what level of BWL constitutes an MTD, although cross-company experience indicates typical upper limits of between 15% and 25% loss. Despite the crucial importance of defining a short-term MTD from a scientific and ethical perspective, variation exists across the industry and regulators on the interpretation of clinical signs and BWL indicative of the MTD. In 2013, a cross-pharmaceutical company working group led by the NC3Rs shared data on BWL in toxicity studies to assess the impact on the animal and the study outcome. Information on 151 studies was used to develop an alert/warning system for BWL in short-term toxicity studies. The data analysis supports BWL limits for short-term dosing (up to 7 days) of 10% for rat and dog and 6% for NHPs (Chapman et al. 2013a). BWL loss above these limits did not add scientific value and was almost always associated with additional clinical signs requiring the animal to be killed, indicating that MTD had already been exceeded. Implementation of these criteria offers the opportunity to reduce the severity of these studies from potentially severe to moderate. However, currently it is not clear how widely these criteria are being applied in practice, and further dissemination may be required to encourage uptake. BWL as an objective indicator of MTD is supported by a similar cross-company initiative within the chemicals industry (mainly agrochemicals), led by the NC3Rs, where data on clinical signs observed during acute inhalation toxicity studies in rats was shared. Statistical analyses showed that BWL >10% is highly predictive (positive predictive value of 94%) of death or severe toxicity at higher doses, showing that the MTD had already been reached or exceeded (Sewell et al. 2015). More data sharing is required to establish criteria for longer study durations. However, it may be that the regulatory guidelines need to be updated to incorporate clear criteria of what constitutes an MTD in order to see a change in practice.

Improvement of Procedures

Microsampling

The biggest opportunity to reduce animal numbers, particularly rodents, in regulatory toxicology studies is in the collection of blood samples for TK evaluation (Harstad et al. 2016; Sparrow et al. 2011), as separate groups of animals, termed satellite animals, are often included for this purpose. TK analysis is carried out to assess systemic exposure and is a required component of repeat-dose toxicity studies (ICH S3A). Conventional sampling approaches require between 200 and 300 μ L/sample and for a complete TK profile the volume required can exceed the maximum blood volume allowed to be taken in rodents. Advances in bioanalytical techniques and new sampling methods mean that small molecules and biologics can be detected in blood samples of <50 μ L, termed microsamples. This enables TK sampling to be carried out in main study animals, reducing the number of satellite animals required, and in some instances removing the need for satellite animals entirely. Furthermore, drug exposures can be directly correlated with toxic effects within the same

Table 3. Example study designs for a 13-week study in rats with assessment of recovery: (a) conventional vs. (b) microsampling approaches

(a) Conventional sampling				
Dose group	Low	Medium	High	Control
No. animals	10 M + 10 F	10 M + 10 F	10 M + 10 F	10 M + 10 F
No. TK satellites	9 M + 9 F	9 M + 9 F	9 M + 9 F	9 M + 9 F
No. recovery animals			5 M + 5 F	5 M + 5 F
Total				172
(b) Microsampling				
Dose group	Low	Medium	High	Control
No. animals	10 M + 10 F	10 M + 10 F	10 M + 10 F	10 M + 10 F
No. recovery animals			5 M + 5 F	5 M + 5 F
Total				100

animals, as is currently the case for nonrodents. Example study designs show how the use of microsampling can remove the need for satellite animals in a 13-week regulatory general toxicology rodent study, potentially reducing animal use by 42% for this type of study (Table 3).

Significant refinements can be realized through implementation of microsampling approaches, as the procedure can be completed in less time and with only minimal restraint, reducing stress for the animal. For rodents, the duration animals have to be placed into a warming chamber prior to blood collection can be reduced (Powles-Glover et al. 2014a, 2014b). Additionally, the use of closed blood collection systems removes the potential for overfilling of tubes and can minimize the total blood loss as well as enabling researchers to more accurately assess the total volume of blood loss. Microsampling can also allow sampling from less invasive sites that otherwise would not be suitable (e.g., ear vein sampling in dogs, rabbits, or minipigs) (Smith et al. 2011).

Surveys of the scientific community carried out by the NC3Rs in 2013 and 2015 have highlighted that microsampling is increasingly being implemented in a wider range of studies, with the majority of survey respondents in 2015 (17/25 compared with 6/22 in 2013) using it for regulatory studies and a 4-fold increase in respondents using it in safety pharmacology studies (8/25 in 2015 compared with 2/22 in 2013) (J Edwards, unpublished data). There has also been an increase in microsampling from nonrodent species, from 14% of respondents in 2013 (3/22) to 44% in 2015 (11/25). Recent publications have highlighted the applicability of microsampling to both adult and juvenile studies in rat (Powles-Glover et al. 2014b) and have confirmed that there is no significant impact of microsampling on a variety of toxicology endpoints, including haematology, plasma biochemistry, and pathology. A Question and Answer document to accompany the ICH S3A guideline on TK is currently in development to consolidate the regulatory perspective on microsampling and support its use in practice (ICH S3A Q&A).

Social Housing

Social housing is common practice for both nonrodents and rodent species in order to meet social behavioral needs and ensure physiological well-being. The most current updates to the EU legislation on the use of animals in scientific research specifically states “except those which are naturally solitary, shall be socially housed in stable groups of compatible individuals” (2010/63/EU 2010) and “single housing of social species should be the exception” (ILAR 2011). However, there may be some occasions where single housing may be considered. For

example, this may be due to practicalities during telemetry recording studies (described in more detail below) or where dietary administration of a drug candidate is required to allow more accurate estimates of food (and therefore drug) consumption. However, in rodents, there appears to be little inter-individual variation in food intake between group-housed animals, and no difference in intake is seen in singly versus group-housed animals (Klir et al. 1984). In fact, there is evidence from the literature that variation in food intakes is greater between sexes than the variation observed between individually and group-housed animals (Boggiano et al. 2008; Krohn et al. 2011; Perez et al. 1997), and that individual housing causes pathophysiological changes to organs that could interfere with interpretation of data with respect to potentially toxic effects (Perez et al. 1997; Wyndham et al. 1983).

Social Housing During Telemetry Recordings

The assessment of cardiovascular function (electrocardiogram, heart rate, and blood pressure) within a nonrodent species is a regulatory requirement for most new chemical entities prior to first administration in humans (ICH S7A, S7B). This is usually performed as a safety pharmacology telemetry study (Leishman et al. 2012) and/or integrated into toxicology studies (Guth et al. 2009), particularly for biopharmaceuticals or anticancer agents when a stand-alone safety pharmacology study is not always required (ICH S6(R1), S9; Vargas et al. 2008). Although it is general practice to socially house nonrodents on nontelemetry recording days, the majority of the industry will separate animals for data collection on the specific telemetry recording days within a study, partly due to limitations in the equipment used and perceptions on animal activity/data variability (Prior et al. 2016). However, these barriers can be overcome as many companies upgrade to new equipment (that allows for social housing) in the future and data from companies successfully socially housing are shared (Kaiser et al. 2015; Klumpp et al. 2006; Xing et al. 2015). Social housing could also be considered for studies using telemetered rodents, and automated cage systems for noninvasive behavioural assessments have recently been introduced (Tse et al. 2016) that are suitable for data collection in a group-housed environment.

Conclusions

Implementation and consideration of the 3Rs have the potential to provide clear benefits to the pharmaceutical industry by improving the efficiency of drug development and registration processes. This can occur by providing new, more human-

relevant predictive early screens that can reduce late-stage attrition but also through refinements within the existing regulatory framework to make improvements to study designs and technical procedures.

There is a need to regularly challenge and review practices and regulatory requirements so that they incorporate the most up-to-date scientific advances and techniques, with better focus on the scientific question. Cross-company and cross-sector approaches that allow data sharing can improve practices and may provide the evidence and confidence to stimulate change and move towards more science-driven processes. Global harmonization of regulations (particularly non-ICH members) and improved communication between regulators and researchers are also required to ensure new opportunities are identified and used in practice across the whole industry.

References

- 2010/63/EU. 2010. Directive 2010/63/EU of the European Parliament and of the Council of 22 September 2010 on the protection of animals used for scientific purposes.
- Authier S, Vargas HM, Curtis MJ, Holbrook M, Pugsley MK. 2013. Safety pharmacology investigations in toxicology studies: an industry survey. *J Pharmacol Toxicol Methods* 68:44–51.
- Baldrick P. 2008. Safety evaluation to support First-In-Man investigations II: toxicology studies. *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* 59:227–236.
- Bass AS, Cartwright ME, Mahon C, Morrison R, Snyder R, McNamara P, Bradley P, Zhou YY, Hunter J. 2009. Exploratory drug safety: a discovery strategy to reduce attrition in development. *J Pharmacol Toxicol Methods* 60:69–78.
- Boggiano MM, Cavigelli SA, Dorsey JR, Kelley CE, Ragan CM, Chandler-Laney PC. 2008. Effect of a cage divider permitting social stimuli on stress and food intake in rats. *Physiol Behav* 95:222–228.
- Bowes J, Brown AJ, Hamon J, Jarolimek W, Sridhar A, Waldron G, Whitebread S. 2012. Reducing safety-related drug attrition: the use of in vitro pharmacological profiling. *Nat Rev Drug Discov* 11:909–922.
- Burden N, Sewell F, Andersen ME, Boobis A, Chipman JK, Cronin MT, Hutchinson TH, Kimber I, Whelan M. 2015. Adverse outcome pathways can drive non-animal approaches for safety assessment. *J Appl Toxicol* 35:971–975.
- Cavero I, Holzgrefe H. 2014. Comprehensive in vitro proarrhythmia assay, a novel in vitro/in silico paradigm to detect ventricular proarrhythmic liability: a visionary 21st century initiative. *Expert Opin Drug Saf* 13:745–758.
- Chapman K, Adjei A, Baldrick P, da Silva A, De Smet K, DiCicco R, Hong SS, Jones D, Leach MW, McBlane J, Ragan I, Reddy P, Stewart DI, Suitters A, Sims J. 2016. Waiving in vivo studies for monoclonal antibody biosimilar development: National and global challenges. *MAbs* 8:427–435.
- Chapman K, Burnett J, Corvaro M, Mitchell D, Robinson S, Sangster T, Sparrow S, Spooner N, Wilson A. 2014a. Reducing pre-clinical blood volumes for toxicokinetics: Toxicologists, pathologists and bioanalysts unite. *Bioanalysis* 6:2965–2968.
- Chapman K, Chivers S, Gliddon D, Mitchell D, Robinson S, Sangster T, Sparrow S, Spooner N, Wilson A. 2014b. Overcoming the barriers to the uptake of nonclinical microsampling in regulatory safety studies. *Drug Discov Today* 19:528–532.
- Chapman K, Creton S, Kupferschmidt H, Bond GR, Wilks MF, Robinson S. 2010. The value of acute toxicity studies to support the clinical management of overdose and poisoning: A cross-discipline consensus. *Regul Toxicol Pharmacol* 58: 354–359.
- Chapman K, Pullen N, Coney L, Dempster M, Andrews L, Bajramovic J, Baldrick P, Buckley L, Jacobs A, Hale G, Green C, Ragan I, Robinson V. 2009. Preclinical development of monoclonal antibodies: Considerations for the use of non-human primates. *MAbs* 1:505–516.
- Chapman K, Sewell F, Allais L, Delongas JL, Donald E, Festag M, Kervyn S, Ockert D, Noguez V, Palmer H, Popovic M, Roosen W, Schoenmakers A, Somers K, Stark C, Stei P, Robinson S. 2013a. A global pharmaceutical company initiative: An evidence-based approach to define the upper limit of body weight loss in short term toxicity studies. *Regul Toxicol Pharmacol* 67:27–38.
- Chapman KL, Andrews L, Bajramovic JJ, Baldrick P, Black LE, Bowman CJ, Buckley LA, Coney LA, Couch J, Maggie Dempster A, de Haan L, Jones K, Pullen N, de Boer AS, Sims J, Ian Ragan. C 2012. The design of chronic toxicology studies of monoclonal antibodies: Implications for the reduction in use of non-human primates. *Regul Toxicol Pharmacol* 62: 347–354.
- Chapman KL, Holzgrefe H, Black LE, Brown M, Chellman G, Copeman C, Couch J, Creton S, Gehen S, Hoberman A, Kinter LB, Madden S, Mattis C, Stemple HA, Wilson S. 2013b. Pharmaceutical toxicology: Designing studies to reduce animal use, while maximizing human translation. *Regul Toxicol Pharmacol* 66:88–103.
- Cocorocchio M, Ives R, Clapham D, Andrews PL, Williams RS. 2016. Bitter tastant responses in the amoeba *Dictyostelium* correlate with rat and human taste assays. *Altex*. 33: 225–236.
- Colleton C, Brewster D, Chester A, Clarke DO, Heining P, Olaharski A, Graziano M. 2016. The use of minipigs for pre-clinical safety assessment by the pharmaceutical industry: Results of an IQ DruSafe minipig survey. *Toxicol Pathol* 44: 458–466.
- Cook D, Brown D, Alexander R, March R, Morgan P, Satterthwaite G, Pangalos MN. 2014. Lessons learned from the fate of AstraZeneca's drug pipeline: A five-dimensional framework. *Nat Rev Drug Discov* 13:419–431.
- EMA. 2010a. Committee for Medicinal Products for Human Use (CHMP). Guideline on repeated dose toxicity CPMP/SWP/1042/99 Rev 1 Corr* (2010).
- EMA. 2010b. Committee for Medicinal Products for Human Use (CHMP). Questions and answers on the withdrawal of the 'Note for guidance on single dose toxicity.' 24 June 2010. EMA/CHMP/SWP/81714/2010.
- EMA. 2014a. Committee for Medicinal Products for Human Use (CHMP). Guideline on similar biological medicinal products containing biotechnology-derived proteins as active substance: non-clinical and clinical issues. 18 December 2014. EMA/CHMP/BMWP/42832/2005 Rev1.
- EMA. 2014b. Committee for Medicinal Products for Human Use (CHMP). Guideline on similar biological medicinal products. 23 October 2014. CHMP/437/04 Rev 1.
- EMA. 2003. Committee for proprietary medicinal products (CPMP). Position paper on non-clinical safety studies to support clinical trials with single microdose. CPMP/SWP/2599/02. January 2003.

- EMEA. 2008. Concept Paper. Need for revision of the guideline single dose toxicity 3BS1a.(EMEA/CHMP/SWP/302413/08). Eudralex.Vol3B. Eudralex Vol3B; 3BS1a single dose toxicity.
- FELASA. 1994. Pain and distress in laboratory rodents and lagomorphs. Report of the Federation of European Laboratory Animal Science Associations (FELASA) Working Group on Pain and Distress accepted by the FELASA Board of Management November 1992. *Lab Anim* 28:97–112.
- Fermi B, Hancox JC, Abi-Gerges N, Bridgland-Taylor M, Chaudhary KW, Colatsky T, Correll K, Crumb W, Damiano B, Erdemli G, Gintant G, Imredy J, Koerner J, Kramer J, Levesque P, Li Z, Lindqvist A, Obejero-Paz CA, Rampe D, Sawada K, Strauss DG, Vandenberg JI. 2016. A new perspective in the field of cardiac safety testing through the comprehensive in vitro proarrhythmia assay paradigm. *J Biomol Screen* 21:1–11.
- Garcia GR, Noyes PD, Tanguay RL. 2016. Advancements in zebrafish applications for 21st century toxicology. *Pharmacol Ther* 161:11–21.
- Garner RC 2010. Practical experience of using human microdosing with AMS analysis to obtain early human drug metabolism and PK data. *Bioanalysis* 2:429–440.
- Goh J-Y, Weaver R, Dixon L, Platt N, Roberts R. 2015. Development and use of in vitro alternatives to animal testing by the pharmaceutical industry 1980-2013. *Toxicol Res* 4:1297–1307.
- Guth BD, Bass AS, Briscoe R, Chivers S, Markert M, Siegl PK, Valentin JP. 2009. Comparison of electrocardiographic analysis for risk of QT interval prolongation using safety pharmacology and toxicological studies. *J Pharmacol Toxicol Methods* 60:107–116.
- Harstad E, Andaya R, Couch J, Ding X, Liang X, Liederer BM, Messick K, Nguyen T, Schweiger M, Tarrant J, Zhong S, Dean B. 2016. Balancing blood sample volume with 3Rs: implementation and best practices for small molecule toxicokinetic assessments in rats. *ILAR J* 57:XXX–XXX.
- Hartung T, Hoffmann S. 2009. Food for thought ... on in silico methods in toxicology. *Altox* 26:155–166.
- Hay M, Thomas DW, Craighead JL, Economides C, Rosenthal J. 2014. Clinical development success rates for investigational drugs. *Nat Biotechnol* 32:40–51.
- He JH, Gao JM, Huang CJ, Li CQ. 2014. Zebrafish models for assessing developmental and reproductive toxicity. *Neurotoxicol Teratol* 42:35–42.
- Holmes A, Bonner F, Jones D. 2015. Assessing drug safety in human tissues - What are the barriers? *Nat Rev Drug Discov* 14:585–587.
- Hornberg JJ, Laursen M, Brenden N, Persson M, Thougard AV, Toft DB, Mow T. 2014. Exploratory toxicology as an integrated part of drug discovery. Part II: Screening strategies. *Drug Discov Today* 19:1137–1144.
- Horner S, Robinson S, Lees D, Callander R, Roberts R. 2014. Target organ profiles in toxicity studies supporting human dosing: An assessment of recovery and chronic dosing. *Regul Toxicol Pharmacol* 70:270–285.
- Horner S, Ryan D, Robinson S, Callander R, Stamp K, Roberts RA. 2013. Target organ toxicities in studies conducted to support first time in man dosing: An analysis across species and therapy areas. *Regul Toxicol Pharmacol* 65:334–343.
- ICH. M3(R2). Nonclinical Safety Studies for the Conduct of Human Clinical Trials and Marketing Authorization for Pharmaceuticals. International Conference on Harmonisation (ICH). Topic M3(R2): June 2009.
- ICH. M3(R2) Q&A. Nonclinical Safety Studies for the Conduct of Human Clinical Trials and Marketing Authorization for Pharmaceuticals. European Medicines Agency. Committee for medicinal products for human use (CHMP). ICH guidelines M3(R2) Q&A. July 2011.
- ICH. S3A. Note for Guidance on Toxicokinetics: The Assessment of Systemic Exposure in Toxicity Studies. International Conference on Harmonisation (ICH). Topic S3: October 1994.
- ICH. S3A Q&A. Note for Guidance on Toxicokinetics: The Assessment of Systemic Exposure in Toxicity Studies. Questions and Answers. S3A Implementation Working Group. Draft ICH Consensus Guideline. Step 2. January 2016.
- ICH. S5(R3). Final Concept Paper. Detection of toxicity to reproduction for medicinal products and toxicity to male fertility. International Conference on Harmonisation (ICH). Topic S5 (R3): February 2015.
- ICH. S6(R1). Preclinical Safety Evaluation of Biotechnology-Derived Pharmaceuticals. International Conference on Harmonisation (ICH). Topic S6(R1). June 2011.
- ICH. S7A. Safety pharmacological studies for human pharmaceuticals. International Conference on Harmonisation (ICH). Topic S7A. November 2000.
- ICH. S7B. The Non-clinical Evaluation of the Potential for Delayed Ventricular Repolarisation (QT Interval Prolongation) by Human Pharmaceuticals. International Conference on Harmonisation (ICH). Topic S7A. May 2005.
- ICH. S9. Nonclinical Evaluation for Anticancer Pharmaceuticals. International Conference on Harmonisation (ICH). Topic S9. March 2010.
- ILAR. 2011. Institute for Laboratory Animal Research, Guide for the care and use of laboratory animals. 8th ed. Washington, DC: The National Academies Press.
- Jones HM, Chen Y, Gibson C, Heimbach T, Parrott N, Peters SA, Snoeys J, Upreti VV, Zheng M, Hall SD. 2015. Physiologically based pharmacokinetic modeling in drug discovery and development: A pharmaceutical industry perspective. *Clin Pharmacol Ther* 97:247–262.
- Kaiser RA, Tichenor SD, Regalia DE, York K, Holzgreffe HH. 2015. Telemetric assessment of social and single housing: Evaluation of electrocardiographic intervals in jacketed cynomolgus monkeys. *J Pharmacol Toxicol Methods* 75:38–43.
- Klir P, Bondy R, Lachout J, Hanis T. 1984. Physiological changes in laboratory rats caused by different housing. *Physiol Bohemoslov* 33:111–121.
- Klumpp A, Trautmann T, Markert M, Guth B. 2006. Optimizing the experimental environment for dog telemetry studies. *J Pharmacol Toxicol Methods* 54:141–149.
- Kola I, Landis J. 2004. Can the pharmaceutical industry reduce attrition rates? *Nat Rev Drug Discov* 3:711–715.
- Konigsson K. 2010. Reduction in dog numbers in regulatory one-month toxicology studies without compromising scientific quality: Challenging the status quo. *Toxicology* 278:341.
- Krohn TC, Salling B, Hansen AK. 2011. How do rats respond to playing radio in the animal facility? *Lab Anim* 45:141–144.
- Kwok TC, Ricker N, Fraser R, Chan AW, Burns A, Stanley EF, McCourt P, Cutler SR, Roy PJ. 2006. A small-molecule screen in *C. elegans* yields a new calcium channel antagonist. *Nature* 441:91–95.
- Ledwith BJ, DeGeorge JJ. 2011. Changes to ICH guideline M3: new and revised guidance on nonclinical safety studies to support human clinical trials and marketing authorization. *Clin Pharmacol Ther* 89:295–299.
- Leishman DJ, Beck TW, Dybdal N, Gallacher DJ, Guth BD, Holbrook M, Roche B, Wallis RM. 2012. Best practice in the conduct of key nonclinical cardiovascular assessments in drug development: Current recommendations from the

- Safety Pharmacology Society. *J Pharmacol Toxicol Methods* 65:93–101.
- McGivern JV, Ebert AD. 2014. Exploiting pluripotent stem cell technology for drug discovery, screening, safety, and toxicology assessments. *Adv Drug Deliv Rev* 69–70:170–178.
- Monticello TM. 2015. Drug development and nonclinical to clinical translational databases: Past and current efforts. *Toxicol Pathol* 43:57–61.
- Morton DB. 2000. A systematic approach for establishing humane endpoints. *ILAR J* 41:80–86.
- Moscardo E, Fasdelli N, Giarola A, Tontodonati M, Dorigatti R. 2009. An optimised neurobehavioural observation battery integrated with the assessment of cardiovascular function in the beagle dog. *J Pharmacol Toxicol Methods* 60:198–209.
- Moscardo E, McPhie G, Fasdelli N, Giarola A, Tontodonati M, Dorigatti R, Meecham K. 2010. An integrated cardiovascular and neurobehavioural functional assessment in the conscious telemetered cynomolgus monkey. *J Pharmacol Toxicol Methods* 62:95–106.
- Moser VC, Becking GC, MacPhail RC, Kulig BM. 1997. The IPCS collaborative study on neurobehavioral screening methods. *Fundam Appl Toxicol* 35:143–151.
- NC3Rs/LASA. 2009. Guidance on dose level selection for regulatory general toxicology studies for pharmaceuticals. NC3Rs. London.
- Newell DR, Burtles SS, Fox BW, Jodrell DI, Connors TA. 1999. Evaluation of rodent-only toxicology for early clinical trials with novel cancer therapeutics. *Br J Cancer* 81:760–768.
- 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* 32:56–67.
- Pelkonen O, Turpeinen M, Raunio H. 2011. In vivo-in vitro-in silico pharmacokinetic modelling in drug development: current status and future directions. *Clin Pharmacokinet* 50:483–491.
- Perez C, Canal JR, Dominguez E, Campillo JE, Guillen M, Torres MD. 1997. Individual housing influences certain biochemical parameters in the rat. *Lab Anim* 31:357–361.
- Powles-Glover N, Kirk S, Jardine L, Clubb S, Stewart J. 2014a. Assessment of haematological and clinical pathology effects of blood microsampling in suckling and weaned juvenile rats. *Regul Toxicol Pharmacol* 69:425–433.
- Powles-Glover N, Kirk S, Wilkinson C, Robinson S, Stewart J. 2014b. Assessment of toxicological effects of blood microsampling in the vehicle dosed adult rat. *Regul Toxicol Pharmacol* 68:325–331.
- Prior H, Bottomley A, Champeroux P, Cordes J, Delpy E, Dybdal N, Edmunds N, Engwall M, Foley M, Hoffmann M, Kaiser R, Meecham K, Milano S, Milne A, Nelson R, Roche B, Valentin JP, Ward G, Chapman K. 2016. Social housing of non-rodents during cardiovascular recordings in safety pharmacology and toxicology studies. *J Pharmacol Toxicol Methods*. 81:75–87.
- Prior H, McMahon N, Schofield J, Valentin JP. 2009. Non-invasive telemetric electrocardiogram assessment in conscious beagle dogs. *J Pharmacol Toxicol Methods* 60:167–173.
- Raunio H. 2011. In silico toxicology: Non-testing methods. *Front Pharmacol* 2:33.
- Redfern W, Bialecki R, Ewart L, Hammond T, Kinter L, Lindgren S, Pollard C, Rolf M, Valentin J-P. 2010. Impact and prevalence of safety pharmacology-related toxicities throughout the pharmaceutical life cycle. *Toxicologist* 114:1080.
- Redfern WS, Ewart L, Lainée P, Pinches M, Robinson S, Valentin JP. 2013. Functional assessments in repeat-dose toxicity studies: The art of the possible. *Toxicol Res* 2:209–234.
- Redfern WS, Waldron G, Winter MJ, Butler P, Holbrook M, Wallis R, Valentin JP. 2008. Zebrafish assays as early safety pharmacology screens: Paradigm shift or red herring? *J Pharmacol Toxicol Methods* 58:110–117.
- Roberts RA, Kavanagh SL, Mellor HR, Pollard CE, Robinson S, Platz SJ. 2014. Reducing attrition in drug development: smart loading preclinical safety assessment. *Drug Discov Today* 19:341–347.
- Robinson S, Chapman K. 2009. Are acute toxicity studies required to support overdose for new medicines? *Regul Toxicol Pharmacol* 55:110.
- Robinson S, Delongea JL, Donald E, Dreher D, Festag M, Kervyn S, Lampo A, Nahas K, Noguev V, Ockert D, Quinn K, Old S, Pickersgill N, Somers K, Stark C, Stei P, Waterson L, Chapman K. 2008. A European pharmaceutical company initiative challenging the regulatory requirement for acute toxicity studies in pharmaceutical drug development. *Regul Toxicol Pharmacol* 50:345–352.
- Rodríguez B, Carusi A, Abi-Gerges N, Ariga R, Britton O, Bub G, Bueno-Orovio A, Burton RA, Carapella V, Cardone-Noott L, Daniels MJ, Davies MR, Dutta S, Ghetti A, Grau V, Harmer S, Kopjar I, Lambiase P, Lu HR, Lyon A, Minchole A, Muszkiewicz A, Oster J, Paci M, Passini E, Severi S, Taggart P, Tinker A, Valentin JP, Varro A, Wallman M, Zhou X. 2016. Human-based approaches to pharmacology and cardiology: an interdisciplinary and intersectorial workshop. *Europace*. 18:1287–1298.
- Roth A, Singer T. 2014. The application of 3D cell models to support drug safety assessment: Opportunities and challenges. *Adv Drug Deliv Rev* 69–70:179–189.
- Russell WMS, Burch RL. 1959. *The Principles of Humane Experimental Technique*. Wheathamsted, UK: Universities Federation for Animal Welfare, (reprinted 1992).
- Sewell F, Chapman K, Baldrick P, Brewster D, Broadmeadow A, Brown P, Burns-Naas LA, Clarke J, Constan A, Couch J, Czupalla O, Danks A, DeGeorge J, de Haan L, Hettlinger K, Hill M, Festag M, Jacobs A, Jacobson-Kram D, Kopytek S, Lorenz H, Moesgaard SG, Moore E, Pasanen M, Perry R, Ragan I, Robinson S, Schmitt PM, Short B, Lima BS, Smith D, Sparrow S, van Bekkum Y, Jones D. 2014. Recommendations from a global cross-company data sharing initiative on the incorporation of recovery phase animals in safety assessment studies to support first-in-human clinical trials. *Regul Toxicol Pharmacol* 70:413–429.
- Sewell F, Ragan I, Marczylo T, Anderson B, Braun A, Casey W, Dennison N, Griffiths D, Guest R, Holmes T, van Huygevoort T, Indans I, Kenny T, Kojima H, Lee K, Prieto P, Smith P, Smedley J, Stokes WS, Wnorowski G, Horgan G. 2015. A global initiative to refine acute inhalation studies through the use of 'evident toxicity' as an endpoint: Towards adoption of the fixed concentration procedure. *Regul Toxicol Pharmacol* 73:770–779.
- Sinneker D, Laugwitz KL, Moretti A. 2014. Induced pluripotent stem cell-derived cardiomyocytes for drug development and toxicity testing. *Pharmacol Ther* 143:246–252.
- Smith C, Sykes A, Robinson S, Thomas E. 2011. Evaluation of blood microsampling techniques and sampling sites for the analysis of drugs by HPLC-MS. *Bioanalysis* 3:145–156.
- Sparrow SS, Robinson S, Bolam S, Bruce C, Danks A, Everett D, Fulcher S, Hill RE, Palmer H, Scott EW, Chapman KL. 2011. Opportunities to minimise animal use in pharmaceutical regulatory general toxicology: A cross-company review. *Regul Toxicol Pharmacol* 61:222–229.

- Strange K. 2016. Drug discovery in fish, flies, and worms. *ILAR J* 57:XXX-XXX.
- 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.
- Tomlinson L, Ramaiah L, Tripathi NK, Barlow VG, Vitsky A, Poitout-Belissent FM, Bounous DI, Ennulat D. 2016. STP best practices for evaluating clinical pathology in pharmaceutical recovery studies. *Toxicol Pathol* 44:163-172.
- Tse K, Rimmer V, Grant C, Simpson D, Klein S, Armstrong D, Lukins T, Chartsias A, O'Callaghan E, Sillito R, Craig D, Vickers C, Chapman K, Redfern W. (In press) Rodent big brother: a home cage automated behavioural monitoring system for safety pharmacology and toxicology studies. Abstract at Safety Pharmacology Society Annual Meeting, Prague 2015.
- Valerio LG Jr. 2009. In silico toxicology for the pharmaceutical sciences. *Toxicol Appl Pharmacol* 241:356-370.
- Valerio LG Jr. 2011. In silico toxicology models and databases as FDA Critical Path Initiative toolkits. *Hum Genomics* 5:200-207.
- van Aerts LA, De Smet K, Reichmann G, van der Laan JW, Schneider CK. 2014. Biosimilars entering the clinic without animal studies. A paradigm shift in the European Union. *MAbs* 6:1155-1162.
- van Meer PJ, Ebbers HC, Kooijman M, Gispens-de Wied CC, Silva-Lima B, Moors EH, Schellekens H. 2015. Contribution of animal studies to evaluate the similarity of biosimilars to reference products. *Drug Discov Today* 20:483-490.
- Vargas HM, Bass AS, Breidenbach A, Feldman HS, Gintant GA, Harmer AR, Heath B, Hoffmann P, Lagrutta A, Leishman D, McMahon N, Mittelstadt S, Polonchuk L, Pugsley MK, Salata JJ, Valentin JP. 2008. Scientific review and recommendations on preclinical cardiovascular safety evaluation of biologics. *J Pharmacol Toxicol Methods* 58:72-76.
- Waring MJ, Arrowsmith J, Leach AR, Leeson PD, Mandrell S, Owen RM, Pairedeau G, Pennie WD, Pickett SD, 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.
- Willoughby LF, Schlosser T, Manning SA, Parisot JP, Street IP, Richardson HE, Humbert PO, Brumby AM. 2013. An in vivo large-scale chemical screening platform using *Drosophila* for anti-cancer drug discovery. *Dis Models Mech* 6: 521-529.
- Wyndham JR, Everitt AV, Everitt SF. 1983. Effects of isolation and food restriction begun at 50 days on the development of age-associated renal disease in the male Wistar rat. *Arch Gerontol Geriatr* 2:317-332.
- Xing G, Lu J, Hu M, Wang S, Zhao L, Zheng W, Schofield J, Oldman K, Adkins D, Yu H, Platz S, Ren J, Skinner M. 2015. Effects of group housing on ECG assessment in conscious cynomolgus monkeys. *J Pharmacol Toxicol Methods* 73:72-79.