



Minimum Information in In Vivo Research

Patrizia Voehringer and Janet R. Nicholson

Contents

1	Introduction	198
2	General Aspects	201
3	Behavioural Experiments	207
4	Anaesthesia and Analgesia	210
5	Ex Vivo Biochemical and Histological Analysis	212
6	Histology	212
7	Ex Vivo Biochemical Analysis	214
8	Perspective	215
	References	216

Abstract

Data quality, reproducibility and reliability are a matter of concern in many scientific fields including biomedical research. Robust, reproducible data and scientific rigour form the foundation on which future studies are built and determine the pace of knowledge gain and the time needed to develop new and innovative drugs that provide benefit to patients. Critical to the attainment of this is the precise and transparent reporting of data. In the current chapter, we will describe literature highlighting factors that constitute the minimum information that is needed to be included in the reporting of in vivo research. The main part of the chapter will focus on the minimum information that is essential for reporting in a scientific publication. In addition, we will present a table distinguishing information necessary to be recorded in a laboratory notebook or another form of internal protocols versus information that should be reported in a paper. We will use examples from the behavioural literature, in vivo studies where the use of

P. Voehringer · J. R. Nicholson (✉)
Boehringer Ingelheim Pharma GmbH & Co. KG, Biberach an der Riss, Germany
e-mail: patrizia.voehringer@boehringer-ingelheim.com;
janet.nicholson@boehringer-ingelheim.com

© The Author(s) 2019

A. Bernaldo de Siqueira et al. (eds.), *Good Research Practice in Non-Clinical Pharmacology and Biomedicine*, Handbook of Experimental Pharmacology 257,
https://doi.org/10.1007/164_2019_285

197

anaesthetics and analgesics are used and finally *ex vivo* studies including histological evaluations and biochemical assays.

Keywords

Behavior · Data quality · In vivo · Publication · Reporting · Reproducibility · Standards

1 Introduction

Data quality, reproducibility and reliability are a matter of concern in many scientific fields including biomedical research. Robust, reproducible data and scientific rigour form the foundation on which future studies are built and determine the pace of knowledge gain and the time needed to develop new and innovative drugs that provide benefit to patients (Freedman and Gibson 2015). In particular, research involving animals is essential for the progression of biomedical science, assuming that experiments are well designed, performed, analysed, interpreted as well as reported.

However, it has been described many times over the last few years that in preclinical research – particularly preclinical animal research – many findings presented in high-profile journals are not reliable and cannot be replicated (Begley and Ellis 2012; Peers et al. 2012; Prinz et al. 2011). This has led to the so-called reproducibility crisis which, according to some, may largely be due to the failure to adhere to good scientific and research practices and the neglect of rigorous and careful application of scientific methods (Begley and Ioannidis 2015; Collins and Tabak 2014). In this context, various reasons have been suggested to contribute to and perhaps explain the lack of reliability and reproducibility in preclinical research including inadequacies in the design, execution and statistical analysis of experiments as well as deficiencies in their reporting (Glasziou et al. 2014; Ioannidis et al. 2014; Jarvis and Williams 2016).

It has been reported that only a minority of animal studies described in the scientific literature use critical experimental design features such as randomisation and blinding despite these components being essential to the production of robust results with minimal risk of experimental bias (Hirst et al. 2014; Macleod et al. 2015). Furthermore, in a study by Bebarta et al., it was described that studies, which did not utilise randomisation and blinding, were more likely to display differences between control and treatment groups, leading to an overestimation of the magnitude of the treatment effects (Bebarta et al. 2003). Another kind of bias that may compromise the validity of preclinical research is reporting bias, consisting of publication bias as well as selective analysis and outcome reporting bias. In many cases, animal studies with negative, neutral or inconclusive results are not reported at all (publication bias), or only the analysis yielding the best statistically significant effect is selectively presented from a host of outcomes that were measured (selective analysis and outcome reporting bias) (Tsilidis et al. 2013). This under-representation of negative research findings can be misleading concerning the interpretation of presented data, often associated with falsely inflated efficacy estimates of an

intervention (Korevaar et al. 2011). Furthermore, unnecessary repetitions of similar studies by investigators unaware of earlier efforts may result.

In 2005, Ioannidis stated that it can be proven that most published research findings are irreproducible or even false due to the incorrect and inadequate use of statistics for their quantification. Specifically, underlying factors such as flexible study designs, flexible statistical analyses and the conductance of small studies with low statistical power were described (Button et al. 2013; Ioannidis 2005). Along these lines, Marino expressed the view that poor understanding of statistical concepts is a main contributory factor to why so few research findings can be reproduced (Marino 2014). Thus, it is urgently required that best practices in statistical design and analysis are incorporated into the framework of the scientific purpose, thereby increasing confidence in research findings.

Additionally, transparent, clear and consistent reporting of research involving animals has become a further substantial issue. Systematic analysis has revealed that a significant proportion of publications reporting in vivo research lack information on study planning, study execution and/or statistical analysis (Avey et al. 2016; Kilkenny et al. 2009; Landis et al. 2012). This failure in reporting makes it difficult to identify potential drawbacks in the experimental design and/or data analysis of the underlying experiment, limiting the benefit and impact of the findings. Moreover, when many of these factors are intertwined, this can lead to negative consequences such as higher failure rates and poor translation between preclinical and clinical phases (Hooijmans and Ritskes-Hoitinga 2013).

Importantly, from an ethical perspective, laboratory animals should be used responsibly. In this context, it is of utmost importance to implement Russell and Burch's 3Rs (reduction, refinement, replacement) principle in the planning and execution of animal studies (Carlsson et al. 2004; Tannenbaum and Bennett 2015; Wuerbel 2017) as well as more efficient study designs, improved research methods including experimental practice, animal husbandry and care. Also the availability of sufficient information and detailed descriptions of animal studies may help to improve animal welfare and to avoid unnecessary animal experiments and wasting animals on inconclusive research.

In the past decade, several guidelines and frameworks have been released in order to improve the scientific quality, transparency and reproducibility of animal experiments (Hooijmans et al. 2010; Kilkenny et al. 2010a; Nature 2013; NIH, Principles and Guidelines for Reporting Preclinical Research). The ARRIVE (Animal Research: Reporting In Vivo Experiments) guidelines focus on the clear and transparent reporting of the minimum information that all scientific publications reporting preclinical animal research should include such as study design, experimental procedures and specific characteristics of the animals used (Kilkenny et al. 2010b). Similarly, the Gold Standard Publication Checklist (GSPC) also aims at improving the planning, design and execution of animal experiments (Hooijmans et al. 2011). The ARRIVE guidelines were launched in 2010 by a team led by the UK National Centre for the Replacement, Refinement and Reduction of Animals in Research (NC3Rs) and have steadily gained credence over the past years. Endorsed by more than 1,000 biomedical journals, the ARRIVE guidelines are now the most

widely accepted key reporting recommendations for animal research (NC3Rs, ARRIVE: Animal Research: Reporting In Vivo Experiments). In addition, various leading scientific journals have begun to change their review practices and place greater emphasis on experimental details prompting authors to report all relevant information on how the study was designed, conducted and analysed (Curtis and Abernethy 2015; Curtis et al. 2015; McGrath and Lilley 2015; McNutt 2014a, b; Nature 2013). Such initiatives may help to ensure transparency and reproducibility of preclinical animal research, thereby improving its reliability and predictive value as well as maximising a successful translation into clinically-relevant applications. However, the compliance with these guidelines remains low several years later. An evaluation of papers published in *Nature* and *PLOS* journals in the 2 years before and after the ARRIVE guidelines were communicated suggests that there has been only little improvement in reporting standards and that authors, referees and editors generally are ignoring the guidelines (Baker et al. 2014). Quite recently, a similar analysis by Leung et al. has shown that the reporting quality in animal research continues to be low and that supporting the ARRIVE guidelines by several journals has not resulted in a considerable improvement of reporting standards (Leung et al. 2018). Obviously, despite the widespread endorsement of the guiding principles by multiple journals in various research areas, the impact of this endorsement on the quality of reporting standards of animal studies is only modest (Avey et al. 2016; Delgado-Ruiz et al. 2015; Liu et al. 2016; Schwarz et al. 2012; Ting et al. 2015). In part, this may be caused by the fact that the recommendations have limitations regarding feasibility and applicability across the diversity of scientific fields that comprise biomedical research making them impractical for some kind of studies. Moreover, researchers may not be convinced that it is necessary to apply effort in order to achieve maximum transparency and reproducibility of animal-based research. It is crucial to increase the awareness of the existence of animal research reporting guidelines as well as the importance of their implementation. A serious problem of guiding principles in general and the ARRIVE guidelines in particular is that most biomedical research journals endorse them but do not rigorously enforce them by urgently requiring comprehensive and detailed reporting of the performed research. A direct consequence of enforced compliance may be increased time and financial burdens making an balanced weighting between what is ideal and what is feasible and practical absolutely essential (Leung et al. 2018).

Nevertheless, the scientific community needs effective, practical and simple tools, maybe in the form of guidelines or checklists, to promote the quality of reporting preclinical animal research. Ideally, such guiding principles should be used as references earlier in the research process before performing the study, helping scientists to focus on key methodological and analytical principles and to avoid errors in the design, execution and analysis of the experiments.

A recent study by Han et al. showed that the mandatory application of a checklist improved the reporting of crucial methodological details, such as randomisation, blinding and sample size estimation, in preclinical in vivo animal studies (Han et al. 2017). Such positive examples support optimism that when reporting is distinctly required, important improvements will be achieved (Macleod 2017). Accordingly,

the strict adherence to reporting guidelines will become useful to address the concerns about data reproducibility and reliability that are widely recognised in the scientific community.

In the present chapter, we discuss the minimum information that should be provided for an adequate description of *in vivo* experiments, in order to allow others to interpret, evaluate and eventually reproduce the study. The main part of the chapter will focus on the minimum information that is essential for the reporting in a scientific publication. In addition, a table will be presented distinguishing information necessary to be recorded in a laboratory notebook or another form of internal record versus information that should be reported in a paper. Examples of specific research areas such as behavioural experiments, anaesthesia and analgesia and their possible interference with experimental outcomes as well as *ex vivo* biochemical and histological analysis will be described.

2 General Aspects

Over the last decade, several guiding principles, such as the GSPC and the ARRIVE guidelines, have been developed in order to improve the quality of designing, conducting, analysing and particularly reporting preclinical animal research. These recommendations have in common that all major components of animal studies that can affect experimental outcomes, including conditions of animal housing, husbandry and care, have to be efficiently reported. In the following section, the most important aspects mentioned in these guidelines are summarised (Hooijmans et al. 2010; Kilkenny et al. 2010a). Finally, a table will be presented comparing information that is necessary to be recorded in a laboratory notebook or another form of internal protocols versus information that should be reported in a scientific publication (Table 1).

At the beginning of a preclinical animal study report, readers should be introduced to the research topic within the context of the scientific area as well as the motivation for performing the current study and the focus of the research question, specific aims and objectives. Primarily, it should be explained why the specific animal species and strain have been chosen and how this animal model can address the scientific hypotheses, particularly with regard to the clinical relevance of the project.

Any studies involving the use of laboratory animals must be formally approved by national regulatory authorities. Therefore, it is necessary to provide information indicating that the protocol used in the study has been ethically reviewed and approved. Additionally, any compliance to national or institutional guidelines and recommendations for the care and use of animals that cover the research should be stated (Jones-Bolin 2012).

In order to allow the replication of a reported study, a detailed description of the experimental animals has to be provided, including species, strain (exact genetic code/nomenclature), gender, age (at the beginning and the end of the experiment), weight (at the start of the experiment) and the origin and source of the animals.

Table 1 Necessary information for including in a publication and recording in a laboratory notebook

	Publication	Laboratory notebook
<i>(A) Experimental model</i>		
Model name	✓	✓
Background and purpose of test	✓	✓
Species	✓	✓
Sex of animals used	✓	✓
Genetic background		
• Standard name	✓	✓
• Original and current parental strain		✓
• Number of backcrosses from original to current		✓
Genetic manipulation		
• Knockout/transgenic	✓	✓
• Constitutive/inducible	✓	✓
• Cre line	✓	✓
• Doxycycline/tamoxifen	✓	✓
Previous use in an experiment		
• Drug and test naïve Y/N	✓	✓
• Description of previous procedures including details of drug washout period/return to baseline values	✓	✓
Source of animals		
• Name of commercial vendor/collaborator/in-house breeding	✓	✓
• Age and weight when received	✓	✓
• Health/immune status	✓	✓
In-house colony		
• Breeding/husbandry	✓	✓
• Breeding scheme (state female genotype first)		✓
• Duos/trios		✓
• Are all animals littermates Y/N (if Y then how)		✓
• How many cohorts are planned for each study		✓
• How far apart are the cohorts		✓
• Are all experimental groups equally represented in all cohorts (Y/N)		✓
• How often are breeders changed		✓
• Birth check frequency		✓
• Sexing age		✓
• Age at weaning		✓
• Are litters mixed at weaning		✓
Age		✓
• At start/end of experiment	✓	✓
<i>(B) Experimental details</i>		
Habituation to vivarium period (days) (if sourced from external) prior to experimental procedure	✓	✓
Assignment to experimental groups		

(continued)

Table 1 (continued)

	Publication	Laboratory notebook
• Randomisation method	✓	✓
• Matching for group assignment (name of variable matched)		✓
• Procedures to minimise bias (e.g. litter, cohort, cage, treatment order)		✓
• SOPs available (Y/N)		✓
Experimenter blinding procedures		
• Procedures to keep treatments blind	✓	✓
• Procedures to keep experimenter blind	✓	✓
• Blinding code and decoding timeline		✓
• SOPs available		✓
Training of experimenters		
• Are experimenters trained and certified in each procedure?	✓	✓
• Method of training and certification		✓
• How often is certification renewed?		✓
Sample		
• Sample size	✓	✓
• Power analysis conducted for each measure for each test	✓	✓
Experimental protocols for each test	✓	✓
• Description	✓	✓
• Tests order and rationale	✓	✓
• Duration of habituation to testing room	✓	✓
• SOPs available (Y/N)		✓
Food/water access during experiment (description)	✓	✓
• Ad libitum or restricted access to food and water during experiment	✓	✓
Adverse/noteworthy events during test	✓	✓
Exclusion criteria	✓	✓
Data processing and analysis	✓	✓
• QC methods		✓
• Primary and secondary measures for each test	✓	✓
• Analysis for each measure for each test	✓	✓
• Check to see if data meets statistical test assumptions	✓	✓
• Treatment of outliers	✓	✓
• Experimental units of analysis (animal/cage/litter)	✓	✓
• Notebooks and data storage		✓
Drug	✓	✓
Name of drug used	✓	✓
Source of drug	✓	✓
Drug batch/sample number	✓	✓
Storage prior to preparation	✓	✓
Drug preparation		✓
• Vehicle name and details of preparation	✓	✓

(continued)

Table 1 (continued)

	Publication	Laboratory notebook
• Doses and rationale	✓	✓
• Dose volume	✓	✓
• Route of administration	✓	✓
• Time of administration and pretreatment time	✓	✓
• Drug storage (cold/dark/duration)	✓	✓
Blood sampling time and method (for bioanalysis)		
• Blood sampling method	✓	✓
• Blood sample volume	✓	✓
• Type of collection tube	✓	✓
• Plasma/serum preparation method	✓	✓
• Plasma/serum freezing and storage	✓	✓
Anaesthesia method and monitoring	✓	✓
Euthanasia method and monitoring	✓	✓
Genotyping tissue collection		
• Age at genotyping	✓	✓
• Method of genotyping	✓	✓
• Is genotyping repeated at end of study? (Y/N)	✓	✓
Tail samples kept (Y/N)		✓
Animal ID		
• Method used to ID animals, frequency of checking		✓
<i>(C) Animal facility</i>		
Microbial/pathogen status (if specific pathogen-free (SPF) specify pathogens)	✓	✓
Housing		
• Housing room used		✓
• Experimental rooms used		✓
• Species/sex of animals housed in same room		✓
• Caging type	✓	✓
• Controls in place for position of cages? (e.g. light differences, proximity to door)		✓
• Use of ventilated racks	✓	✓
• Number of animals per cage	✓	✓
• Are cages homogeneous for genotype	✓	✓
• Are animals regrouped at any time? If so, at what age?	✓	✓
Enrichment	✓	
• Type of bedding	✓	✓
• Toys in cage? Running wheel?	✓	✓
• Shredded paper?	✓	✓
• Igloos? Other?	✓	✓
Light/dark cycle	✓	
• Time of lights on/off	✓	✓
• Light/dark change with dawn and dusk light gradient? If Y, over what time frame?	✓	✓

(continued)

Table 1 (continued)

	Publication	Laboratory notebook
Music/sound used. If so, specify details	✓	✓
Humidity	✓	✓
Type of chow	✓	✓
Water (acidified/tap/distilled/autoclaved/filtered/other?)	✓	✓
Air exchange frequency		✓
Handling		
Frequency and duration of handling	✓	✓
Husbandry		
• No. cage changes/week		✓
• No. health checks/week		✓
Health reports from facility		✓
Personal protective equipment, description		✓
<i>(D) Approvals and authorisation</i>		
For example, IACUC or AAALAC approval number and date	✓	✓
Ethical approval statement/animal license application	✓	✓
<i>(E) Equipment</i>		
Description of equipment used	✓	✓
• Model number	✓	✓
• Vendor	✓	✓
Calibration		
• Method	✓	✓
• Frequency	✓	✓

Adapted from Brunner et al. (2016)

These biological variables are scientifically important since they often represent critical factors affecting health or disease of the animals and therefore may influence research outcomes (GV-SOLAS 1985; Oebrink and Reh binder 2000). For the same reason, it is also essential to comment on the animals' experience and to state if they are drug naïve or if they have received any previous procedures or treatments. Additionally, information about the health, microbiological and immune status of the animals can be of high relevance for study outcomes and the ability to replicate findings and therefore should be given (GV-SOLAS 1999). This means, e.g. to depict that the animals are kept under specific pathogen-free (SPF) conditions (accompanied by a list of pathogens excluded) and that their health and microbiological status is checked and monitored according to the FELASA recommendations (Nicklas et al. 2002). When using genetically modified animals, it is important to describe their genetic background, how these animals were generated and which control animals were selected.

There is increasing evidence that elements of the laboratory environment as well as housing and husbandry practices can significantly affect the animals' biology and

ultimately research outcomes (Hogan et al. 2018; Reardon 2016). This implicates an exact specification of the environmental conditions in which the animals were housed and where the experiments were conducted. The animal facility should be described concerning temperature, relative humidity, ventilation, lighting (light/dark cycle, light intensity) and noise (Baldwin et al. 2007; Speakman and Keijer 2012; Swoap et al. 2004; Van der Meer et al. 2004). In more detail, the specific housing conditions of the animals should be represented including type and size of the cages, bedding material, availability and type of environmental enrichment, number of animals per cage (and reasons for individual housing when applicable) as well as frequency of cage changes and handling procedures (Balcombe et al. 2004; Nicholson et al. 2009; Perez et al. 1997; Rock et al. 1997; van Praag et al. 2000). In addition, the reporting of nutrition and water regimes needs to be specified regarding the type (composition, special diets, purification) as well as access to food and water (ad libitum, restricted, amount of food/water and frequency and time of feeding or water supply).

When describing the procedures carried out in animal studies, several aspects require thorough consideration and need to be presented for each experiment and each experimental group, including controls. When has the experiment been performed (day and time of intervention and time interval between intervention and data sampling or processing)? Where has the experiment been performed (home cage, laboratory, special device/equipment for investigation)? What kind of intervention has been carried out? Here, details about the methodological techniques such as surgical procedures or sampling methods (including specialist equipment and suppliers) should be provided. Importantly, drugs and compounds used in the experiments need to be specified concerning name, manufacturer and concentration as well as the formulation protocol, dosage, application volume, frequency and route of administration. Additionally, when anaesthetics and analgesics are required for animal welfare reasons, it is crucial to include information about the name of these agents, administered doses, route and frequency of application as well as monitoring procedures of the animals' physiological signs that are used to guarantee a sufficient level of anaesthesia and analgesia (Flecknell 2018; Gaertner et al. 2008). Similarly, the method of euthanasia applied at the end of the study should be described (Sivula and Suckow 2018).

To ensure the quality and validity of preclinical animal research, it is crucial to indicate if the performed study is a confirmatory or hypothesis-testing one and to implement appropriate experimental study designs (Johnson and Besselsen 2002). This comprises a clear definition of the experimental unit (individual animal or group of animals in one cage) as well as the number of treatment and control (positive, negative, vehicle) groups. In this context, the reporting of animal numbers (total number per experiment as well as per experimental group) is essential to assess biological and statistical significance of the results and to re-analyse the data. Additionally, any power and sample size calculations used for the determination of adequate animal numbers that allow the generation of statistically meaningful results should be reported (Button et al. 2013). Moreover, any actions undertaken to minimise the effects of subjective bias when allocating animals to experimental

groups (e.g. randomisation) and when assessing results (e.g. blinding) should be stated (Bello et al. 2014; Hirst et al. 2014; Moser 2019). Randomisation is the best method to achieve balance between treatment and control groups, whereas blinded assessment of outcomes (assessing, measuring or quantifying) improves qualitative scoring of subjective experimental observations and promotes comparable handling of data. Both strategies enhance the rigour of the experimental procedure and the scientific robustness of the results.

When reporting the results of the experiments, statistics needs to be fully described including the statistical method/test used to analyse the primary and secondary outcomes of the study (Marino 2014). The exact number of analysed animals and a measure of precision (mean, median, standard deviation, standard error of the mean, confidence interval) should be presented. This is of high relevance for interpreting the results and for evaluating the reliability of the findings. Importantly, the number of excluded animals as well as reasons and criteria to exclude them from the experiment, and hence analysis, should be well documented. Furthermore, the description of outcomes should comprise the full spectrum of positive and negative results as well as whether there were attempts to repeat or confirm the data. Equally, all relevant adverse events and any modifications that were made to the experimental protocol in order to reduce these unwanted effects should be reported.

Finally, when discussing and interpreting the findings, it is important to take into account the objectives and hypotheses of the study as predetermined in the experimental study design. Additionally, a comment on the overall scientific relevance of the outcomes as well as their potential to translate into clinical significance should be included. In order to demonstrate how animal welfare issues have been addressed in the current study, any implications of the experimental methods or results for the replacement, refinement or reduction of the use of laboratory animals in research need to be described (Taylor 2010).

In conclusion, the meaningful and accurate reporting of preclinical animal studies encompasses a plethora of aspects, ranging from a detailed description of the experimental animal to a complete documentation of the statistical analysis. Creating transparency in this way can help to evaluate studies in terms of their planning, methodology, statistical verification and reproducibility. It is highly recommended to make all raw data, analyses and protocols available to the whole research community in order to provide insight into the full workflow of the scientific project.

3 Behavioural Experiments

Behavioural animal studies are of great importance to increase the scientific knowledge about the complex processes underlying animal behaviour in general as well as to investigate potential drug effects on behavioural outcomes. Furthermore, translational research aims to identify disease-relevant endpoints in behavioural animal studies that are robust, reliable and reproducible and ultimately can be used to assess

the potential of novel therapeutic agents to treat human diseases (Sukoff Rizzo and Silverman 2016).

However, performing behavioural experiments in animals is largely challenging for scientists since studies of this nature are extremely sensitive to external and environmental factors (Crabbe et al. 1999). Specific housing conditions, e.g. the lack of environmental stimulation, can interfere with brain development and selectively alter brain functions, thereby affecting the expression of certain behaviour (Wuerbel 2001). Resulting stereotypies and other abnormal repetitive behaviours can be severely confounding in behavioural experiments and have an impact on the validity, reliability and reproducibility of scientific outcomes (Garner 2005).

Additionally, when measuring behaviour in animals, there are multiple other factors that may influence the generation of a behavioural response which can be classified as ‘trait’, ‘state’ and ‘technical’ factors (Sousa et al. 2006). ‘Trait’ factors include genetic (e.g. genetic background, gender) as well as developmental characteristics (e.g. stress experience, handling, housing conditions, social hierarchy) of the animals. ‘State’ factors comprise the time of the experiment, the experience and training status of the investigator, characteristics of the animal (e.g. age, health status, pharmacological treatment) as well as features of the experimental setup (e.g. construction, illumination, test environment, cleansing). ‘Technical’ factors encompass data acquisition (e.g. automated vs. manual observation, calibration, choice of behavioural parameters) as well as data analysis (e.g. distribution, normalisation of data).

In preclinical research settings, it is difficult to standardise all such factors, which may contribute to the poor reproducibility of behavioural observations in animals across different laboratories (Wahlsten 2001). Standardisation is assumed to minimise the variability of results and to increase sensitivity and precision of the experimental procedure. However, contrary to the assumption that rigorous standardisation of animal experiments may help to ensure their reproducibility, it has been proposed that rather, systematic variation of experimental conditions (heterogenisation) can lead to the generation of robust and generalisable results across behavioural animal studies since the external validity is enhanced, thereby improving reproducibility (Richter et al. 2010; Voelkl et al. 2018; Wuerbel 2000). Nevertheless, considering that a strict and universal standardisation of laboratory environmental and experimental conditions is exceptionally unlikely, it is of major importance to take into account any possible determinants that might exert an effect on animals’ performance when designing, conducting and analysing behavioural experiments and to report these factors accurately and transparently.

As mentioned above, there is increasing evidence that the laboratory environment and distinct husbandry and housing conditions may influence animal welfare and hence behaviour. Moreover, test outcomes of behavioural animal studies are highly dependent on small but important details regarding these conditions that are usually poorly reported. One such example is light conditions: light is a fundamental environmental factor regulating animal activity and physiology, and it has been found in rats that intense light conditions can lead to retinal damage, suppression of social play behaviour and locomotion as well as dissociation of circadian rhythms

(Castelhano-Carlos and Baumans 2009). Similarly, environmental sounds that are inevitably present in animal research facilities also exert considerable effects on animals' physiology and behaviour influencing sleeping patterns, locomotor activity, learning and anxiety reactions. Provision of a stable and controlled light and noise environment for the animals will contribute to their wellbeing and to the reproducibility of experimental outcomes, making a clear reporting of light and noise conditions obligatory.

Standard husbandry practices such as regularly performed cage-changing as well as commonly-used experimental procedures such as injections can significantly affect behavioural parameters in rodents, as measured by increased arousal behaviour and locomotor activity (Duke et al. 2001; Gerdin et al. 2012). These stress-related responses may have a considerable influence on the validity and quality of experimental outcomes and should be considered by researchers when designing study protocols and comparing data. Similarly, it has been shown that a change in housing conditions, including a combination of standard vs. individually ventilated cages and single vs. social housing, has a major impact on several physiological parameters and behavioural features of mice such as body weight, locomotor activity and anxiety-related behaviour (Pasquarelli et al. 2017). Thus, it is mandatory to clearly state as well as maintain a well-defined housing protocol during the experiment in order to ensure better comparison, reliability and reproducibility of experimental results across research facilities.

Environmental cage enrichment, which should be transparently reported when describing animals' housing conditions, is strongly recommended by various guidelines regulating laboratory animal care and accommodation, as it is reported to enhance animal welfare, to protect against the development of stereotypies, to reduce anxiety and to positively influence brain development as well as learning and memory behaviour (Simpson and Kelly 2011). And indeed, it has been shown in rats and mice that environmental enrichment does not result in enhanced individual data variability nor generate inconsistent data in replicate studies between multiple laboratories, indicating that housing conditions can be improved without impacting the quality or reproducibility of behavioural results (Baumans et al. 2010; Wolfer et al. 2004).

Much evidence concerning the reproducibility of behavioural animal studies comes from the area of rodent phenotyping (Kafkafi et al. 2018). Some behavioural phenotypes, such as locomotor activity, can be highly reproducible across several laboratories, suggesting high stability and therefore better reproducibility (Wahlsten et al. 2006). In contrast, other behavioural phenotypes, such as anxiety-like behaviour, are more problematic to measure since they show increased susceptibility to a multitude of environmental factors that can affect the animals' performance. Indeed, it has been reported that animal handling procedures, particularly the specific handling method itself, can elicit profound effects on animals' anxiety levels and stress responses, indicating that the use of handling methods that will not induce strong anxiety responses will minimise confounding effects during experiments (Hurst and West 2010).

One of the most commonly used methods to investigate anxiety behaviour in rodents is the elevated plus maze (EPM) test (Lister 1987; Pellow et al. 1985). Besides strain, gender and age differences, it has been shown that the manipulation of the animals prior to the experiment (e.g. exposure to stressors, housing, handling procedures) and the averseness of the test conditions themselves (e.g. increased light levels) as well as repeated testing in the EPM can strongly influence the manifestation of anxiety behaviour (Bessa et al. 2005; File 2001; Hogg 1996). These crucial factors should not be excluded from experimental descriptions when reporting. Additionally, illumination of the EPM is a critical aspect that needs to be clearly specified. In fact, Pereira et al. concluded that it is not the absolute level of luminosity upon the arms, but the relative luminosity between the open and closed arms that predicts the behavioural performance of rats in the maze (Pereira et al. 2005).

Overall, it has been suggested that animal behaviour that is more closely linked to sensory input and motor output will probably be less affected by minimal modifications within the laboratory environment, whereas behaviour that is associated with emotional and social processes will be more sensitive (Wahlsten et al. 2006).

4 Anaesthesia and Analgesia

For numerous animal experiments such as surgeries or imaging studies, the use of anaesthetics and analgesics in order to reduce animal suffering from pain and distress is an ethical obligation and crucial to the 3Rs concept (Carbone 2011). However, it is known that these drugs (as well as untreated pain itself) can severely affect the animals' biology and physiology, thereby influencing experimental data and introducing variability into research outcomes. Focusing on animal pain management means both an issue of generating high-quality, reproducible data and a substantial animal welfare concern. Dealing with this ethical and methodological conflict can pose a challenging task for scientists.

The ARRIVE guidelines recommend the reporting of anaesthesia and analgesia in order to achieve a full and detailed description of the experimental procedures performed in preclinical animal studies and to allow the critical evaluation and reproduction of published data. However, there is evidence that the current scientific literature lacks important details concerning the use of animal anaesthetics and analgesics, underestimating their potential interference with experimental results (Carbone and Austin 2016; Uhlig et al. 2015). In many cases, it is not clear whether scientists actively withhold treatment of animals with anaesthetic or analgesic drugs or just fail to include this information in the reporting, perhaps due to assumed insignificance to the experimental outcome. This creates the false impression that the selection of appropriate anaesthetic and analgesic regimens is not considered as a crucial methodological concern for generating high-quality research data. Furthermore, under-reporting of anaesthesia and pain management may also shape ongoing

practice among researchers and encourage under-treatment of animals, which represents a serious problem concerning animal welfare.

Surgical pain and insufficient analgesia act as stressors and can elicit various effects on the animals' immune system, food and water consumption, social behaviour, locomotor activity as well as metabolic and hormone state, among others, which may all influence the experimental outcomes of animal studies (Leach et al. 2012; Liles and Flecknell 1993). The use of anaesthetics and analgesics relieves surgical pain, thus contributing to the refinement of the experimental methods. Additionally, following the surgical procedure, an appropriate long-term pain management, which could last for several days, is required to ensure animal wellbeing. However, anaesthetic and analgesic drugs themselves may also confound experimental results, e.g. by regulating inflammatory pathways or exerting immunomodulatory effects (Al-Hashimi et al. 2013; Fuentes et al. 2006; Galley et al. 2000; Martucci et al. 2004). In cancer studies on tumour metastasis in rats, it has been shown that analgesic drugs such as tramadol are able to prevent the effect of experimental surgery on natural killer cell activity and on the enhancement of metastatic diffusion, which needs to be taken into account when using this kind of animal model (Gaspani et al. 2002). Furthermore, as demonstrated for inhalation anaesthesia using sevoflurane in rats, the expression of circadian genes may be severely influenced, which needs to be borne in mind in the design of animal studies analysing gene expression (Kobayashi et al. 2007).

As indicated in these few examples, the selection of appropriate anaesthetic and analgesic procedures is a key factor in preclinical animal studies and has to be carefully considered in the context of the specific research question and study protocol (Gargiulo et al. 2012). Scientists need to know which particular anaesthetic and analgesic drugs were used, including name, dose, application frequency and route of administration. Importantly, concerning long-term pain management after surgery, it is recommended to specify the duration of the analgesic treatment. Moreover, when it is decided to withhold analgesics because of interference with the research project, it is essential to include the reasons for this decision when reporting the study so that this information is available to those who may subsequently wish to replicate and extend such studies (Stokes et al. 2009).

Hypothermia, hypotension, hypoxemia and respiratory depression are frequently observed side effects during animal anaesthesia that can develop to serious health problems culminating in unexpected death (Davis 2008). These risks need to be incorporated when planning and performing experiments and highlight the importance of adequate animal monitoring procedures to eliminate the incidence of complications during anaesthesia. Additionally, the reporting of such events and their practical management (e.g. the use of warming pads) is crucial for scientists trying to reproduce and evaluate research data.

Animal imaging studies have specific requirements concerning anaesthesia that are related to the use of particular methodological techniques and the duration of the experiments. The primary reason for general anaesthesia in imaging studies is the need for the restraint and immobility of the animals in order to avoid movement artefacts and to obtain signals with maximal reproducibility (Gargiulo et al. 2012).

However, anaesthetic agents can unintentionally affect physiological parameters of animals and confound the outcomes of different imaging modalities (Hildebrandt et al. 2008). As shown for positron emission tomography (PET) neuroimaging studies, the use of anaesthetics such as ketamine or isoflurane may alter neuromolecular mechanisms in animal brains, thereby leading to an incorrect representation of normal properties of the awake brain (Alstrup and Smith 2013). Moreover, repeated anaesthesia procedures and the preparation of the animals for the study may influence the processes under investigation. Physical restraint stress before the experiment can increase the anaesthetic induction doses and negatively influence the quality of some molecular imaging procedures such as PET due to altered kinetics and biodistribution of radiotracers (Hildebrandt et al. 2008). The latter effect has also been observed to be dependent on the choice of anaesthetics, the duration of fasting periods as well as to result from hypothermia observed as an adverse event from anaesthesia (Fueger et al. 2006).

As for surgical procedures, the careful selection of the most appropriate anaesthesia method addressing all the needs and goals of the specific research project and imaging modality is important (Vesce et al. 2017). Since anaesthetics can influence various physiological and pharmacological functions of the animals, monitoring of anaesthetic levels and of vital functions during imaging studies has proven useful. In order to achieve reproducible experimental conditions in imaging studies, a clear and consistent reporting of methodological details concerning the animals, fasting conditions, anaesthesia regimens and monitoring is absolutely essential.

5 Ex Vivo Biochemical and Histological Analysis

Numerous *ex vivo* methods, including biochemical and histological analyses, are used routinely to complement *in vivo* studies to add additional information or to address scientific questions which are difficult to address in an *in vivo* setting. The starting point for such studies is a living organism, and as such, many of the previously described considerations in, e.g. the ARRIVE guidelines are entirely applicable and should be included when reporting data from such studies. In the following section, we will highlight examples of studies where specific methodological details have been evinced to be important for outcome and as such should be included in any reporting of data from studies where similar *ex vivo* analyses have been carried out.

6 Histology

Histology is the microscopic study of animal and plant cells and tissues. It comprises a multistage process of cell or tissue collection and processing, sectioning, staining and examining under a microscope to finally quantification. Various methods are routinely applied in numerous cell and tissue types. The field of histology has been as affected as others by the lack of reproducibility of data across labs. In a recent

report, Dukkipati et al. made the observation that conflicting data on the presence of pathological changes in cholinergic synaptic inputs (C-boutons) exists in the field of amyotrophic lateral sclerosis (ALS), thus making it difficult to assess roles of these synaptic inputs in the pathophysiology of the disease (Dukkipati et al. 2017). The authors sought to determine whether or not the reported changes described in the scientific literature are indeed truly statistically and biologically significant and to evaluate the possible reasons for why reproducibility has proven problematic. Thus, histological analyses were conducted using several variations on experimental design and data analysis and indeed, it was shown that factors including the grouping unit, sampling strategy and lack of blinding could all be contributors to the failure in replication of results. Furthermore, the lack of power analysis and effect size made the assessment of biological significance difficult. Experimental design has also been the focus of a report by Torlakovic et al. who have highlighted the importance of inclusion of appropriate and standardised controls in immunohistochemistry studies so that data can be reproduced from one test to another and indeed from one lab to another (Torlakovic et al. 2015). Lai et al. point to the difficulty in standardising complex methods in their report of the development of the OPTIClear method using fresh and archived human brain tissue (Lai et al. 2018).

A comparison of different quantification methods has been described by Wang et al. to determine hippocampal damage after cerebral ischemia (Wang et al. 2015). The authors start with the comment that multiple techniques are used to evaluate histological damage following ischemic insult although the sensitivity and reproducibility of these techniques is poorly characterised. Nonetheless, their output has a pivotal impact on results and conclusions drawn therefrom. In this study, two factors emerged as being important methodological aspects. Firstly, since neuronal cell death does not occur homogeneously within the CA1 region of the hippocampus, it is critical that the time post ischemic insult is accurately reported. Secondly, in terms of analysis regarding counting strategy, window size and position were both shown to have a major impact on study results and should therefore be clearly reported. Ward et al. make the point that in order to reproduce histopathological results from, e.g. the mouse, the pathology protocol, including necropsy methods and slide preparation, should be followed by interpretation of the slides by a pathologist familiar with reading mouse slides and familiar with the consensus medical nomenclature used in mouse pathology (Ward et al. 2017). Additionally, for the peer review of manuscripts where histopathology is a key part of the investigation, pathologists should be consulted.

The importance of such studies to the field is further acknowledged by the existence of numerous initiatives to improve reproducibility. For in situ hybridisation (ISH) and immunohistochemistry (IHC) biomarkers, the minimum information specification for ISH and IHC experiments (MISFISHIE) guidelines has been developed by the Stem Cell Genome Anatomy Projects consortium, and it is anticipated that compliance should enable researchers at different laboratories to fully evaluate data and reproduce experiments (Deutsch et al. 2008). The MISFISHIE checklist includes six aspects of information to be provided in the reporting of experiments ranging from experimental design, biomaterials and

treatments, reporter (probe or antibody) information, staining protocols and parameters, imaging data and parameters and imaging characterisations. The use of statistics and any guidance on interpretation of results is, however, not included. The authors stress that the implementation of MISFISHIE should not remove variability in data but rather facilitate the identification of specific sources of variability. A similarly intended study describes a checklist of 20 items that are recommended to be included when reporting histopathology studies (Knijn et al. 2015). Thus, while reproducibility in histological analyses has been a problem and has perhaps hindered scientific progress, the field has adapted and adherence to new tools and guidelines that are now available offer hope that we are moving rapidly in a positive direction.

7 Ex Vivo Biochemical Analysis

Biochemical assessments can be performed in numerous ex vivo biological materials ranging from CSF to organoids and are routinely used to assess mRNA and proteins such as hormones.

Flow cytometry of ventricular myocytes is an emerging technology in cardiac research. Cellular variability and cytometer flow cell size are known to affect cytometer performance, and these two factors of variance are considered to limit assay validity and reproducibility across laboratories. In a study by Lopez et al., the authors hypothesised that washing and filtering create a bias towards sampling smaller cells than actually exist in the adult heart and they performed a study to test this (Lopez et al. 2017). The study results revealed that there was indeed a significant impact of washing and filtering on the experimental outcome and thus proposed a no-wash step in the protocol that could become part of a standard experimental design to minimise variability across labs.

Deckardt et al. have investigated the effect of a range of commonly used anaesthetics on clinical pathology measures including glucose, serum proteins, hormones and cholinesterase (Deckardt et al. 2007). The authors demonstrated differential effects of the different anaesthetics with regard to some of the measured parameters and differences across the sex and species used, thus demonstrating the importance of understanding the impact that an anaesthetic can have – even on ex vivo readouts – and to include appropriate controls. A similar study was conducted by Lelovas et al. which further highlights the importance of concise and accurate reporting of the use of anaesthetics in the collection of biological samples for biochemical readouts since their use can have a significant impact on outcome (Lelovas et al. 2017). Watters and Goodman published a comparison of basic methods in clinical studies and in in vitro tissue and cell culture studies reported in three anaesthesia journals (Watters and Goodman 1999). The authors identified 16 in vitro articles, and although they were not able to identify anything inherently wrong with the studies, they noted the small sample sizes and the lack of reporting on failures (only 2 of 53) and describe anecdotal evidence of experimenters only reporting on the experiments that work. The authors conclude with a call for all

investigators to give reasons for sample size, to use randomisation and blinding wherever possible and to report exclusions and withdrawals, thus enabling an improvement in robustness and general applicability of the data published.

Antibodies are commonly used tools in research, particularly in *ex vivo* analyses. A common cause for the lack of reproducibility of data using antibodies could be due to the lack of thorough validation (Drucker 2016). The importance of standardised reagents has been highlighted by Venkataraman et al. who have described the establishment of a toolbox of immunoprecipitation-grade monoclonal antibodies to human transcription factors with the aim of improving quality and reproducibility across labs (Venkataraman et al. 2018). This work was conducted as part of the NIH protein capture reagents programme (PCRP) which has generated over 1,500 reagents that can be used by the scientific community.

8 Perspective

An improvement in quality in preclinical research and particularly where animals are used is urgently needed. To achieve this, it is of fundamental importance to change the way experimental results are reported in the scientific literature so that data can be more easily reproduced across labs. This should enable more rapid scientific progress and reduce waste. Scientists are encouraged to adopt the existing guidelines by defining all relevant information that has to be included in publications and study reports, with the aim of enhancing the transparency, reproducibility and reliability of scientific work. Ensuring that preclinical research proceeds along structured guidelines will strengthen the robustness, rigour and validity of scientific data and ultimately the suitability of animal studies for translation into clinical trials.

We have described several important factors relating to behavioural experiments that may influence the outcomes of some selected behavioural animal studies. Obviously, this represents only a small part of the various possible variations of the laboratory environment, equipment and methodological procedures that can affect animal behaviour. However, we have indicated the importance of considering and reporting all relevant details regarding behavioural experiments, which will help to resolve the common problem of poor reproducibility of certain findings across different laboratories and to ensure high quality of behaviour animal studies.

We have highlighted the use of anaesthesia and analgesia as factors that can have a significant impact on experimental data, and it is therefore of utmost importance that their use is reported comprehensively. High animal welfare standards require the use of anaesthetics and analgesics when performing painful and stress-inducing experiments. However, since these drugs may severely influence research outcomes, it is necessary to carefully select the most suitable procedures for the scientific question under investigation and to evaluate the importance of the scientific needs in the context of animal wellbeing and existing guidelines for the description of experimental animal research should be applied. The complete reporting of anaesthesia procedures as well as pain management could significantly improve the quality and reproducibility of preclinical animal studies and enhance animal welfare.

Ex vivo measures including histological analysis and biochemical readouts are seemingly just as prone to poor reproducibility as in vivo experiments. Clearly, the precise details of the in-life part of the study should not be overlooked in the reporting of such studies since this aspect can have a significant impact on overall experimental outcome and conclusions.

The field has reached a point where something needs to be done to improve standards, and indeed, to this end, numerous initiatives are ongoing. One such initiative is the Innovative Medicines Initiative consortium project “European Quality in Preclinical Research” (EQIPD). The EQIPD project aims to identify ways to enable a smoother, faster and safer transition from preclinical to clinical testing by establishing common guidelines to strengthen the robustness, rigour and validity of research data. Numerous academic and industrial partners are involved in this initiative, which should have a significant and positive impact in the next few years. Nevertheless, the output of EQIPD and similar efforts need to be embraced and for that, the entire scientific community has an important role to play.

References

- Al-Hashimi M, Scott SW, Thompson JP, Lambert DG (2013) Opioids and immune modulation: more questions than answers. *Br J Anaesth* 111:80–88
- Alstrup AK, Smith DF (2013) Anaesthesia for positron emission tomography scanning of animal brains. *Lab Anim* 47:12–18
- Avey MT, Moher D, Sullivan KJ, Fergusson D, Griffin G, Grimshaw JM, Hutton B, Lalu MM, Macleod M, Marshall J, Mei SH, Rudnicki M, Stewart DJ, Turgeon AF, McIntyre L, Group CCCTB (2016) The devil is in the details: incomplete reporting in preclinical animal research. *PLoS One* 11:e0166733
- Baker D, Lidster K, Sottomayor A, Amor S (2014) Two years later: journals are not yet enforcing the ARRIVE guidelines on reporting standards for pre-clinical animal studies. *PLoS Biol* 12:e1001756
- Balcombe JP, Barnard ND, Sandusky C (2004) Laboratory routines cause animal stress. *Contemp Top Lab Anim Sci* 43:42–51
- Baldwin AL, Schwartz GE, Hopp DH (2007) Are investigators aware of environmental noise in animal facilities and that this noise may affect experimental data? *J Am Assoc Lab Anim Sci* 46:45–51
- Baumans V, Van Loo P, Pham TM (2010) Standardisation of environmental enrichment for laboratory mice and rats: utilisation, practicality and variation in experimental results. *Scand J Lab Anim Sci* 37:101–114
- Bebarta V, Luyten D, Heard K (2003) Emergency medicine animal research: does use of randomization and blinding affect the results? *Acad Emerg Med* 10:684–687
- Begley CG, Ellis LM (2012) Drug development: raise standards for preclinical cancer research. *Nature* 483:531–533
- Begley CG, Ioannidis JP (2015) Reproducibility in science: improving the standard for basic and preclinical research. *Circ Res* 116:116–126
- Bello S, Krogsboll LT, Gruber J, Zhao ZJ, Fischer D, Hrobjartsson A (2014) Lack of blinding of outcome assessors in animal model experiments implies risk of observer bias. *J Clin Epidemiol* 67:973–983
- Bessa JM, Oliveira M, Cerqueira JJ, Almeida OF, Sousa N (2005) Age-related qualitative shift in emotional behaviour: paradoxical findings after re-exposure of rats in the elevated-plus maze. *Behav Brain Res* 162:135–142

- Brunner D, Balci B, Kabitzke P, Hill H (2016) Consensus preclinical checklist (PRECHECK): experimental conditions – rodent disclosure checklist. *Int J Comp Psychol* 29(1):1–5
- Button KS, Ioannidis JP, Mokrysz C, Nosek BA, Flint J, Robinson ES, Munafò MR (2013) Power failure: why small sample size undermines the reliability of neuroscience. *Nat Rev Neurosci* 14:365–376
- Carbone L (2011) Pain in laboratory animals: the ethical and regulatory imperatives. *PLoS One* 6:e21578
- Carbone L, Austin J (2016) Pain and laboratory animals: publication practices for better data reproducibility and better animal welfare. *PLoS One* 11:e0155001
- Carlsson HE, Hagelin J, Hau J (2004) Implementation of the ‘three Rs’ in biomedical research. *Vet Rec* 154:467–470
- Castelhano-Carlos MJ, Baumans V (2009) The impact of light, noise, cage cleaning and in-house transport on welfare and stress of laboratory rats. *Lab Anim* 43:311–327
- Collins FS, Tabak LA (2014) Policy: NIH plans to enhance reproducibility. *Nature* 505:612–613
- Crabbe JC, Wahlsten D, Dudek BC (1999) Genetics of mouse behavior: interactions with laboratory environment. *Science* 284:1670–1672
- Curtis MJ, Abernethy DR (2015) Revision of instructions to authors for pharmacology research and perspectives: enhancing the quality and transparency of published work. *Pharmacol Res Perspect* 3:e00106
- Curtis MJ, Bond RA, Spina D, Ahluwalia A, Alexander SP, Giembycz MA, Gilchrist A, Hoyer D, Insel PA, Izzo AA, Lawrence AJ, MacEwan DJ, Moon LD, Wonnacott S, Weston AH, McGrath JC (2015) Experimental design and analysis and their reporting: new guidance for publication in *BJP*. *Br J Pharmacol* 172:3461–3471
- Davis JA (2008) Mouse and rat anesthesia and analgesia. *Curr Protoc Neurosci*. Appendix 4: Appendix 4B. <https://doi.org/10.1002/0471142301.nsa04bs42>
- Deckardt K, Weber I, Kaspers U, Hellwig J, Tennekens H, van Ravenzwaay B (2007) The effects of inhalation anaesthetics on common clinical pathology parameters in laboratory rats. *Food Chem Toxicol* 45:1709–1718
- Delgado-Ruiz RA, Calvo-Guirado JL, Romanos GE (2015) Critical size defects for bone regeneration experiments in rabbit calvariae: systematic review and quality evaluation using ARRIVE guidelines. *Clin Oral Implants Res* 26:915–930
- Deutsch EW, Ball CA, Berman JJ, Bova GS, Brazma A, Bumgarner RE, Campbell D, Causton HC, Christiansen JH, Daian F, Dauga D, Davidson DR, Gimenez G, Goo YA, Grimmond S, Henrich T, Herrmann BG, Johnson MH, Korb M, Mills JC, Oudes AJ, Parkinson HE, Pascal LE, Pollet N, Quackenbush J, Ramialison M, Ringwald M, Salgado D, Sansone SA, Sherlock G, Stoeckert CJ Jr, Swedlow J, Taylor RC, Walashek L, Warford A, Wilkinson DG, Zhou Y, Zou LI, Liu AY, True LD (2008) Minimum information specification for in situ hybridization and immunohistochemistry experiments (MISFISHIE). *Nat Biotechnol* 26:305–312
- Drucker DJ (2016) Never waste a good crisis: confronting reproducibility in translational research. *Cell Metab* 24:348–360
- Duke JL, Zammit TG, Lawson DM (2001) The effects of routine cage-changing on cardiovascular and behavioral parameters in male Sprague-Dawley rats. *Contemp Top Lab Anim Sci* 40:17–20
- Dukkipati SS, Chihai A, Wang Y, Elbasiouny SM (2017) Experimental design and data analysis issues contribute to inconsistent results of C-bouton changes in amyotrophic lateral sclerosis. *eNeuro* 4:0281
- File SE (2001) Factors controlling measures of anxiety and responses to novelty in the mouse. *Behav Brain Res* 125:151–157
- Flecknell P (2018) Rodent analgesia: assessment and therapeutics. *Vet J* 232:70–77
- Freedman LP, Gibson MC (2015) The impact of preclinical irreproducibility on drug development. *Clin Pharmacol Ther* 97:16–18
- Fueger BJ, Czernin J, Hildebrandt I, Tran C, Halpern BS, Stout D, Phelps ME, Weber WA (2006) Impact of animal handling on the results of 18F-FDG PET studies in mice. *J Nucl Med* 47:999–1006

- Fuentes JM, Talamini MA, Fulton WB, Hanly EJ, Aurora AR, De Maio A (2006) General anesthesia delays the inflammatory response and increases survival for mice with endotoxic shock. *Clin Vaccine Immunol* 13:281–288
- Gaertner DJ, Hallman TM, Hankenson FC, Batchelder MA (2008) *Anesthesia and analgesia in laboratory animals*, 2nd edn. Academic Press, London
- Galley HF, DiMatteo MA, Webster NR (2000) Immunomodulation by anaesthetic, sedative and analgesic agents: does it matter? *Intensive Care Med* 26:267–274
- Gargiulo S, Greco A, Gramanzini M, Esposito S, Affuso A, Brunetti A, Vesce G (2012) Mice anesthesia, analgesia, and care, part I: anesthetic considerations in preclinical research. *ILAR J* 53:E55–E69
- Garner JP (2005) Stereotypies and other abnormal repetitive behaviors: potential impact on validity, reliability, and replicability of scientific outcomes. *ILAR J* 46:106–117
- Gaspani L, Bianchi M, Limiroli E, Panerai AE, Sacerdote P (2002) The analgesic drug tramadol prevents the effect of surgery on natural killer cell activity and metastatic colonization in rats. *J Neuroimmunol* 129:18–24
- Gerdin AK, Igosheva N, Roberson LA, Ismail O, Karp N, Sanderson M, Cambridge E, Shannon C, Sunter D, Ramirez-Solis R, Bussell J, White JK (2012) Experimental and husbandry procedures as potential modifiers of the results of phenotyping tests. *Physiol Behav* 106:602–611
- Glaziou P, Altman DG, Bossuyt P, Boutron I, Clarke M, Julious S, Michie S, Moher D, Wager E (2014) Reducing waste from incomplete or unusable reports of biomedical research. *Lancet* 383:267–276
- GV-SOLAS (1985) Guidelines for specification of animals and husbandry methods when reporting the results of animal experiments. Working Committee for the Biological Characterization of Laboratory Animals/GV-SOLAS. *Lab Anim* 19:106–108
- GV-SOLAS (1999) Implications of infectious agents on results of animal experiments. Report of the Working Group on Hygiene of the Gesellschaft für Versuchstierkunde--Society for Laboratory Animal Science (GV-SOLAS). *Lab Anim* 33(Suppl 1):S39–S87
- Han S, Olonisakin TF, Pribis JP, Zupetic J, Yoon JH, Holleran KM, Jeong K, Shaikh N, Rubio DM, Lee JS (2017) A checklist is associated with increased quality of reporting preclinical biomedical research: a systematic review. *PLoS One* 12:e0183591
- Hildebrandt IJ, Su H, Weber WA (2008) Anesthesia and other considerations for in vivo imaging of small animals. *ILAR J* 49:17–26
- Hirst JA, Howick J, Aronson JK, Roberts N, Perera R, Koshiaris C, Heneghan C (2014) The need for randomization in animal trials: an overview of systematic reviews. *PLoS One* 9:e98856
- Hogan MC, Norton JN, Reynolds RP (2018) Environmental factors: macroenvironment versus microenvironment. In: Weichbrod RH, GAH T, Norton JN (eds) *Management of animal care and use programs in research, education, and testing*. CRC Press, Boca Raton, pp 461–478
- Hogg S (1996) A review of the validity and variability of the elevated plus-maze as an animal model of anxiety. *Pharmacol Biochem Behav* 54:21–30
- Hooijmans CR, Ritskes-Hoitinga M (2013) Progress in using systematic reviews of animal studies to improve translational research. *PLoS Med* 10:e1001482
- Hooijmans CR, Leenaars M, Ritskes-Hoitinga M (2010) A gold standard publication checklist to improve the quality of animal studies, to fully integrate the Three Rs, and to make systematic reviews more feasible. *Altern Lab Anim* 38:167–182
- Hooijmans CR, de Vries R, Leenaars M, Curfs J, Ritskes-Hoitinga M (2011) Improving planning, design, reporting and scientific quality of animal experiments by using the Gold Standard Publication Checklist, in addition to the ARRIVE guidelines. *Br J Pharmacol* 162:1259–1260
- Hurst JL, West RS (2010) Taming anxiety in laboratory mice. *Nat Methods* 7:825–826
- Ioannidis JP (2005) Why most published research findings are false. *PLoS Med* 2:e124
- Ioannidis JP, Greenland S, Hlatky MA, Khoury MJ, Macleod MR, Moher D, Schulz KF, Tibshirani R (2014) Increasing value and reducing waste in research design, conduct, and analysis. *Lancet* 383:166–175

- Jarvis MF, Williams M (2016) Irreproducibility in preclinical biomedical research: perceptions, uncertainties, and knowledge gaps. *Trends Pharmacol Sci* 37:290–302
- Johnson PD, Besselsen DG (2002) Practical aspects of experimental design in animal research. *ILAR J* 43:202–206
- Jones-Bolin S (2012) Guidelines for the care and use of laboratory animals in biomedical research. *Curr Protoc Pharmacol*. Appendix 4: Appendix 4B. <https://doi.org/10.1002/0471141755.pha04bs59>
- Kafkafi N, Agassi J, Chesler EJ, Crabbe JC, Crusio WE, Eilam D, Gerlai R, Golani I, Gomez-Marin A, Heller R, Iraqi F, Jaljuli I, Karp NA, Morgan H, Nicholson G, Pfaff DW, Richter SH, Stark PB, Stiedl O, Stodden V, Tarantino LM, Tucci V, Valdar W, Williams RW, Wuerbel H, Benjamini Y (2018) Reproducibility and replicability of rodent phenotyping in preclinical studies. *Neurosci Biobehav Rev* 87:218–232
- Kilkenny C, Parsons N, Kadoszewski E, Festing MF, Cuthill IC, Fry D, Hutton J, Altman DG (2009) Survey of the quality of experimental design, statistical analysis and reporting of research using animals. *PLoS One* 4:e7824
- Kilkenny C, Browne W, Cuthill IC, Emerson M, Altman DG, Group NCRGW (2010a) Animal research: reporting in vivo experiments: the ARRIVE guidelines. *Br J Pharmacol* 160:1577–1579
- Kilkenny C, Browne WJ, Cuthill IC, Emerson M, Altman DG (2010b) Improving bioscience research reporting: the ARRIVE guidelines for reporting animal research. *PLoS Biol* 8: e1000412
- Knijff N, Simmer F, Nagtegaal ID (2015) Recommendations for reporting histopathology studies: a proposal. *Virchows Arch* 466:611–615
- Kobayashi K, Takemori K, Sakamoto A (2007) Circadian gene expression is suppressed during sevoflurane anesthesia and the suppression persists after awakening. *Brain Res* 1185:1–7
- Korevaar DA, Hooft L, ter Riet G (2011) Systematic reviews and meta-analyses of preclinical studies: publication bias in laboratory animal experiments. *Lab Anim* 45:225–230
- Lai HM, Liu AKL, Ng HHM, Goldfinger MH, Chau TW, DeFelice J, Tilley BS, Wong WM, Wu W, Gentleman SM (2018) Next generation histology methods for three-dimensional imaging of fresh and archival human brain tissues. *Nat Commun* 9:1066
- Landis SC, Amara SG, Asadullah K, Austin CP, Blumenstein R, Bradley EW, Crystal RG, Darnell RB, Ferrante RJ, Fillit H, Finkelstein R, Fisher M, Gendelman HE, Golub RM, Goudreau JL, Gross RA, Gubitza AK, Hesterlee SE, Howells DW, Huguenard J, Kelner K, Koroshetz W, Krainc D, Lasic SE, Levine MS, Macleod MR, McCall JM, Moxley RT 3rd, Narasimhan K, Noble LJ, Perrin S, Porter JD, Steward O, Unger E, Utz U, Silberberg SD (2012) A call for transparent reporting to optimize the predictive value of preclinical research. *Nature* 490:187–191
- Leach MC, Klaus K, Miller AL, Scotto di Perrotolo M, Sotocinal SG, Flecknell PA (2012) The assessment of post-vasectomy pain in mice using behaviour and the Mouse Grimace Scale. *PLoS One* 7:e35656
- Lelovas PP, Stasinopoulou MS, Balafas EG, Nikita MA, Sikos NT, Kostomitsopoulos NG (2017) Valuation of three different anaesthetic protocols on complete blood count and biochemical parameters on Wistar rats. *J Hellenic Vet Med Soc* 68:587–598
- Leung V, Rousseau-Blass F, Beauchamp G, Pang DSJ (2018) ARRIVE has not ARRIVED: support for the ARRIVE (animal research: reporting of in vivo experiments) guidelines does not improve the reporting quality of papers in animal welfare, analgesia or anesthesia. *PLoS One* 13:e0197882
- Liles JH, Flecknell PA (1993) The effects of surgical stimulus on the rat and the influence of analgesic treatment. *Br Vet J* 149:515–525
- Lister RG (1987) The use of a plus-maze to measure anxiety in the mouse. *Psychopharmacology* 92:180–185

- Liu Y, Zhao X, Mai Y, Li X, Wang J, Chen L, Mu J, Jin G, Gou H, Sun W, Feng Y (2016) Adherence to ARRIVE guidelines in Chinese journal reports on neoplasms in animals. *PLoS One* 11:e0154657
- Lopez JE, Jaradeh K, Silva E, Aminololama-Shakeri S, Simpson PC (2017) A method to increase reproducibility in adult ventricular myocyte sizing and flow cytometry: avoiding cell size bias in single cell preparations. *PLoS One* 12:e0186792
- Macleod M (2017) Findings of a retrospective, controlled cohort study of the impact of a change in Nature journals' editorial policy for life sciences research on the completeness of reporting study design and execution. *bioRxiv*. <https://doi.org/10.1101/187245>
- Macleod MR, Lawson McLean A, Kyriakopoulou A, Serghiou S, de Wilde A, Sherratt N, Hirst T, Hemblade R, Bahor Z, Nunes-Fonseca C, Potluru A, Thomson A, Baginskaite J, Egan K, Vesterinen H, Currie GL, Churilov L, Howells DW, Sena ES (2015) Risk of bias in reports of in vivo research: a focus for improvement. *PLoS Biol* 13:e1002273
- Marino MJ (2014) The use and misuse of statistical methodologies in pharmacology research. *Biochem Pharmacol* 87:78–92
- Martucci C, Panerai AE, Sacerdote P (2004) Chronic fentanyl or buprenorphine infusion in the mouse: similar analgesic profile but different effects on immune responses. *Pain* 110:385–392
- McGrath JC, Lilley E (2015) Implementing guidelines on reporting research using animals (ARRIVE etc.): new requirements for publication in *BJP*. *Br J Pharmacol* 172:3189–3193
- McNutt M (2014a) Journals unite for reproducibility. *Science* 346:679
- McNutt M (2014b) Reproducibility. *Science* 343:229
- Moser P (2019) Out of control? Managing baseline variability in experimental studies with control groups. In: *Handbook of experimental pharmacology, good research practice in pharmacology/experimental life sciences* (in press)
- National Institutes of Health (NIH) Principles and guidelines for reporting preclinical research [Internet]. Available: <https://www.nih.gov/research-training/rigor-reproducibility/principles-guidelines-reporting-preclinical-research>. Accessed 18 Aug 2018
- Nature (2013) Reducing our irreproducibility. *Nature* 496:398
- NC3Rs ARRIVE: animal research: reporting in vivo experiments [Internet]. Available: <https://www.nc3rs.org.uk/arrive-animal-research-reporting-vivo-experiments>. Accessed 23 Aug 2018
- Nicholson A, Malcolm RD, Russ PL, Cough K, Touma C, Palme R, Wiles MV (2009) The response of C57BL/6J and BALB/cJ mice to increased housing density. *J Am Assoc Lab Anim Sci* 48:740–753
- Nicklas W, Baneux P, Boot R, Decelle T, Deeny AA, Fumanelli M, Illgen-Wilcke B, FELASA (2002) Recommendations for the health monitoring of rodent and rabbit colonies in breeding and experimental units. *Lab Anim* 36:20–42
- Oebrink KJ, Reh binder C (2000) Animal definition: a necessity for the validity of animal experiments? *Lab Anim* 34:121–130
- Pasquarelli N, Voehringer P, Henke J, Ferger B (2017) Effect of a change in housing conditions on body weight, behavior and brain neurotransmitters in male C57BL/6J mice. *Behav Brain Res* 333:35–42
- Peers IS, Ceuppens PR, Harbron C (2012) In search of preclinical robustness. *Nat Rev Drug Discov* 11:733–734
- Pellow S, Chopin P, File SE, Briley M (1985) Validation of open:closed arm entries in an elevated plus-maze as a measure of anxiety in the rat. *J Neurosci Methods* 14:149–167
- Pereira LO, da Cunha IC, Neto JM, Paschoalini MA, Faria MS (2005) The gradient of luminosity between open/enclosed arms, and not the absolute level of Lux, predicts the behaviour of rats in the plus maze. *Behav Brain Res* 159:55–61
- 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
- Prinz F, Schlange T, Asadullah K (2011) Believe it or not: how much can we rely on published data on potential drug targets? *Nat Rev Drug Discov* 10:712
- Reardon S (2016) A mouse's house may ruin experiments. *Nature* 530:264

- Richter SH, Garner JP, Auer C, Kunert J, Wuerbel H (2010) Systematic variation improves reproducibility of animal experiments. *Nat Methods* 7:167–168
- Rock FM, Landi MS, Hughes HC, Gagnon RC (1997) Effects of caging type and group size on selected physiologic variables in rats. *Contemp Top Lab Anim Sci* 36:69–72
- Schwarz F, Iglhaut G, Becker J (2012) Quality assessment of reporting of animal studies on pathogenesis and treatment of peri-implant mucositis and peri-implantitis. A systematic review using the ARRIVE guidelines. *J Clin Periodontol* 39(Suppl 12):63–72
- Simpson J, Kelly JP (2011) The impact of environmental enrichment in laboratory rats—behavioural and neurochemical aspects. *Behav Brain Res* 222:246–264
- Sivula CP, Suckow MA (2018) Euthanasia. In: Weichbrod RH, Thompson GAH, Norton JN (eds) *Management of animal care and use programs in research, education, and testing*. CRC Press, Boca Raton, pp 827–840
- Sousa N, Almeida OF, Wotjak CT (2006) A hitchhiker's guide to behavioral analysis in laboratory rodents. *Genes Brain Behav* 5(Suppl 2):5–24
- Speakman JR, Keijer J (2012) Not so hot: optimal housing temperatures for mice to mimic the thermal environment of humans. *Mol Metab* 2:5–9
- Stokes EL, Flecknell PA, Richardson CA (2009) Reported analgesic and anaesthetic administration to rodents undergoing experimental surgical procedures. *Lab Anim* 43:149–154
- Sukoff Rizzo SJ, Silverman JL (2016) Methodological considerations for optimizing and validating behavioral assays. *Curr Protoc Mouse Biol* 6:364–379
- Swoap SJ, Overton JM, Garber G (2004) Effect of ambient temperature on cardiovascular parameters in rats and mice: a comparative approach. *Am J Physiol Regul Integr Comp Physiol* 287:R391–R396
- Tannenbaum J, Bennett BT (2015) Russell and Burch's 3Rs then and now: the need for clarity in definition and purpose. *J Am Assoc Lab Anim Sci* 54:120–132
- Taylor K (2010) Reporting the implementation of the Three Rs in European primate and mouse research papers: are we making progress? *Altern Lab Anim* 38:495–517
- Ting KH, Hill CL, Whittle SL (2015) Quality of reporting of interventional animal studies in rheumatology: a systematic review using the ARRIVE guidelines. *Int J Rheum Dis* 18:488–494
- Torlakovic EE, Nielsen S, Vyberg M, Taylor CR (2015) Getting controls under control: the time is now for immunohistochemistry. *J Clin Pathol* 68:879–882
- Tsilidis KK, Panagiotou OA, Sena ES, Aretouli E, Evangelou E, Howells DW, Al-Shahi Salman R, Macleod MR, Ioannidis JP (2013) Evaluation of excess significance bias in animal studies of neurological diseases. *PLoS Biol* 11:e1001609
- Uhlig C, Krause H, Koch T, Gama de Abreu M, Spieth PM (2015) Anesthesia and monitoring in small laboratory mammals used in anesthesiology, respiratory and critical care research: a systematic review on the current reporting in top-10 impact factor ranked journals. *PLoS One* 10:e0134205
- Van der Meer E, Van Loo PL, Baumans V (2004) Short-term effects of a disturbed light-dark cycle and environmental enrichment on aggression and stress-related parameters in male mice. *Lab Anim* 38:376–383
- van Praag H, Kempermann G, Gage FH (2000) Neural consequences of environmental enrichment. *Nat Rev Neurosci* 1:191–198
- Venkataraman A, Yang K, Irizarry J, Mackiewicz M, Mita P, Kuang Z, Xue L, Ghosh D, Liu S, Ramos P, Hu S, Bayron Kain D, Keegan S, Saul R, Colantonio S, Zhang H, Behn FP, Song G, Albino E, Asencio L, Ramos L, Lugo L, Morell G, Rivera J, Ruiz K, Almodovar R, Nazario L, Murphy K, Vargas I, Rivera-Pacheco ZA, Rosa C, Vargas M, McDade J, Clark BS, Yoo S, Khambadkone SG, de Melo J, Stevanovic M, Jiang L, Li Y, Yap WY, Jones B, Tandon A, Campbell E, Montelione GT, Anderson S, Myers RM, Boeke JD, Fenyo D, Whiteley G, Bader JS, Pino I, Eichinger DJ, Zhu H, Blackshaw S (2018) A toolbox of immunoprecipitation-grade monoclonal antibodies to human transcription factors. *Nat Methods* 15:330–338
- Vesce G, Micieli F, Chiavaccini L (2017) Preclinical imaging anesthesia in rodents. *Q J Nucl Med Mol Imaging* 61:1–18

- Voelkl B, Vogt L, Sena ES, Wuerbel H (2018) Reproducibility of preclinical animal research improves with heterogeneity of study samples. *PLoS Biol* 16:e2003693
- Wahlsten D (2001) Standardizing tests of mouse behavior: reasons, recommendations, and reality. *Physiol Behav* 73:695–704
- Wahlsten D, Bachmanov A, Finn DA, Crabbe JC (2006) Stability of inbred mouse strain differences in behavior and brain size between laboratories and across decades. *Proc Natl Acad Sci U S A* 103:16364–16369
- Wang J, Jahn-Eimermacher A, Bruckner M, Werner C, Engelhard K, Thal SC (2015) Comparison of different quantification methods to determine hippocampal damage after cerebral ischemia. *J Neurosci Methods* 240:67–76
- Ward JM, Schofield PN, Sundberg JP (2017) Reproducibility of histopathological findings in experimental pathology of the mouse: a sorry tail. *Lab Anim* 46:146–151
- Watters MP, Goodman NW (1999) Comparison of basic methods in clinical studies and in vitro tissue and cell culture studies reported in three anaesthesia journals. *Br J Anaesth* 82:295–298
- Wolfer DP, Litvin O, Morf S, Nitsch RM, Lipp HP, Wuerbel H (2004) Laboratory animal welfare: cage enrichment and mouse behaviour. *Nature* 432:821–822
- Wuerbel H (2000) Behaviour and the standardization fallacy. *Nat Genet* 26:263
- Wuerbel H (2001) Ideal homes? Housing effects on rodent brain and behaviour. *Trends Neurosci* 24:207–211
- Wuerbel H (2017) More than 3Rs: the importance of scientific validity for harm-benefit analysis of animal research. *Lab Anim* 46:164–166

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

