BMJ Open Science

BMJ Open Science is committed to open peer review. As part of this commitment we make the peer review history of every article we publish publicly available.

When an article is published we post the peer reviewers' comments and the authors' responses online. We also post the versions of the paper that were used during peer review. These are the versions that the peer review comments apply to.

The versions of the paper that follow are the versions that were submitted during the peer review process. They are not the versions of record or the final published versions. They should not be cited or distributed as the published version of this manuscript.

BMJ Open Science is an open access journal and the full, final, typeset and author-corrected version of record of
the manuscript is available on our site with no access controls, subscription charges or pay-per-view fees (http://openscience.bmj.com).

If you have any questions on BMJ Open's open peer review process please email info.bmjos@bmj.com

Revision of the ARRIVE guidelines: Rationale and scope

Nathalie Percie du Sert[1], Viki Hurst[1], Amrita Ahluwalia[2], Sabina Alam[3], Douglas G. Altman[4], Marc T. Avey[5], Monya Baker[6], William Browne[7], Alejandra Clark[8], Innes C. Cuthill[7], Ulrich Dirnagl[9], Michael Emerson[10], Paul Garner[11], David W. Howells[12], Natasha A. Karp[13], Catriona J. MacCallum[14], Malcolm Macleod[15], Ole Petersen[16], Frances Rawle[17], Penny Reynolds[18], Kieron Rooney[19], Emily S. Sena[15], Shai D. Silberberg[20], Thomas Steckler[21], Hanno Würbel[22], Stephen T Holgate[23]

1   NC3Rs, UK
2   Queen Mary University of London, UK
3   F1000 Research, UK
4   University of Oxford, UK
5   ICF, Canada
6   Nature, USA
7   University of Bristol, UK
8   PLOS ONE, UK
9   Universitätsmedizin Berlin, Germany
10   Imperial College London, UK
11   Liverpool School of Tropical Medicine, UK
12   University of Tasmania, Australia
13   Quantitative Biology, Discovery Science, IMED Biotech Unit, Astrazeneca, Cambridge, UK
14   Hindawi, UK
15   Centre for Clinical Brain Sciences, University of Edinburgh, UK
16   Cardiff University, UK
17   Medical Research Council, UK
18   University of Florida, USA
19   University of Sydney, Australia
20   National Institute of Neurological Disorders and Stroke, USA
21   Janssen Pharmaceutica NV / Johnson & Johnson, Belgium
22   University of Bern, Switzerland
23   University of Southampton, UK

Abstract

In 2010, the NC3Rs published the ARRIVE guidelines to improve the reporting of animal research. Despite considerable levels of support from the scientific community, the impact on the quality of reporting in animal research publications has been limited. This position paper highlights the strategy of an expert working group established to revise the guidelines and facilitate their uptake. The group's initial work will focus on three main areas: prioritisation of the ARRIVE items into a tiered system, development of an explanation and elaboration document, and revision of specific items.

Scientists, funders and the public are increasingly concerned about the reproducibility of preclinical research, including studies that use animals[1]. While the reasons for failing to reproduce the methods and findings of a study are complex and wide ranging, a lack of transparency stemming from poor reporting clearly contributes to the problem[2]. The NC3Rs coordinated the development of the ARRIVE reporting guidelines in 2010[3]. The guidelines consist of a 20-item checklist that covers the key information that should be described in a scientific publication. The goal is to ensure that the reader can assess the methodological rigour of the experiment, and other scientists can evaluate and reproduce the methods.

To date, over a thousand journals, funders and research institutes support and endorse ARRIVE[4]. The guidelines have contributed to the understanding of the issues that compromise reproducibility, and prompted actions to improve the situation. For example, major UK funders now explicitly mandate a comprehensive description of the study design, including plans to minimise experimental bias. Grant applicants must also explain how the number of animals to be used was decided and provide detailed statistical analysis plans to ensure that peer reviewers and panel members can fully assess the rigour and validity of the proposed research[5].

Have the guidelines improved reporting? In the seven years since the ARRIVE guidelines were published, researchers have sought to measure the impact of the guidelines on the quality of reporting[6 7], with mixed results. A recent randomised controlled trial in PLOS ONE[8] showed that mandating the completion of an ARRIVE checklist with manuscript submission, with no additional emphasis on reporting during the editorial process, did not improve adherence to the guidelines in published papers. While these results are disappointing to those seeking immediate change, this study provides an evidence base to improve the guidelines, and ultimately the rigour and reproducibility of animal studies.

In the light of methodological advances in science and experience with the guidelines since their introduction, the NC3Rs convened an international working group to revise them. The authors here are members of the working group, and include funders, journal editors, statisticians, methodologists and animal researchers from academia and industry. The aim of this report is to provide readers and stakeholders with information about areas that we are currently working on to improve the ARRIVE guidelines during 2018. This work includes:

Prioritising the items of the ARRIVE guidelines

Each of the 20 items of the ARRIVE guidelines are important for various reasons. For example, a description of study design (item 6), how the animals were allocated to groups (item 11) and how the sample size was chosen (item 10) are crucial to understand how reliable and robust the findings are. Similarly, items such as the experimental procedures (item 7) or animal characteristics (item 8), are important to ensure that papers contain enough information for others to replicate and build upon the study. Other ARRIVE items such as the scientific background (item 3) and relevance to other species

(item 19) provide information about the context of the study[3]. In their current form, the guidelines do not lend themselves easily to retrospective evaluation; assessing whether a manuscript includes all 20 ARRIVE items necessitates operationalising the checklist into over a hundred separate elements[9]. To enable a more manageable approach for assessing the quality of reporting in manuscripts, we plan to organise the items in the ARRIVE guidelines into tiers reflecting different levels of priority; tier one items will include the most important items on which initial efforts from authors, reviewers and journals should focus. We will carry out a Delphi exercise[10], to structure communications within the working group and with external stakeholders, and reach a consensus on the criteria defining the tiers, and on the most appropriate tier for each item. Importantly, the tiers will also enable a step-wise approach for journals and others to improve reporting standards, the objective being that ultimately all manuscripts will include all elements of the guidelines.

Prioritising subsets of items in this way will provide straightforward measures for journals, institutions and researchers. We anticipate that journals will continue to recommend that authors follow the ARRIVE guidelines in their entirety, to encourage comprehensive reporting. At the same time, focusing editorial efforts on a smaller number of key pieces of information that can be particularly scrutinised by editors and reviewers will enable a more rapid assessment of both individual manuscript quality, and the overall impact of their improvement strategies. This is an approach already used by some journals[11 12]. With the use of text mining and machine learning technologies, automating many of these checks is possible[13], and coordinating work on top tier items will accelerate the development of tools to facilitate this.

Developing an Explanation and Elaboration document

Understanding the rationale for a set of guidelines is essential for securing support from the scientific community. The CONSORT statement, for example, has been accompanied by an Explanation and Elaboration document since its second iteration in 2001[14]. This summarises the evidence behind each item of the guidelines and explains why each item is important to include in a manuscript.

A recent survey of *in vivo* researchers carried out by the NC3Rs[15] highlighted that the main reason for authors not including an ARRIVE item in a manuscript was because they did not think it was necessary to disclose that information. Incomplete reporting is also exacerbated by the fact that some of the concepts included in the guidelines, such as measures to reduce bias, are not well understood by researchers or not considered relevant for their own research[16]. To address this, we are now developing an Explanation and Elaboration document for the ARRIVE guidelines. This document will provide explanations and definitions for technical terms, empirical evidence in support of each ARRIVE item, as well as examples from the published literature on how authors might report items. Following publication of the document, the information will be made readily accessible via a dedicated ARRIVE website.

Revising the guidelines

We are reviewing specific ARRIVE items to ensure that the guidance provided is in line with the current best evidence. Where evidence is lacking we will seek to develop it. The revision is an opportunity to improve the clarity of individual items, ensure their relevance across the breadth of *in vivo* research and enhance the logical flow of information within the guidelines. The publishing landscape has also evolved over the last decade and the revised guidelines will reflect these changes by providing advice on emerging best practice.

The scope of the ARRIVE guidelines is broad; they are designed to be flexible and accommodate the reporting of comparative studies in a wide range of research areas. As such, the existing guidelines were formulated to provide general advice for heterogeneous study types. However, recent calls have been made to encourage researchers to explicitly distinguish between exploratory and hypothesis-testing studies[17 18]. For hypothesis-testing studies that are using inferential statistics, the manuscript would be expected to describe the primary and secondary outcome measures, the parameters used in the sample size calculation, whether the study protocol was preregistered, and if so, where it can be found. Exploratory studies on the other hand are designed to generate hypotheses; they might confer the same importance to all outcomes measured, might justify the sample size based on feasibility or experience, and would report only descriptive statistics. Thus the reporting requirements for exploratory and hypothesis-testing studies can differ, and the revision will ensure that the updated guidelines provide adequate advice for both.

Final remarks

Revising the guidelines is just the first step; their primary goal is to improve transparency and the standards of reporting, but transparent reporting can be used to address common weaknesses in the design and conduct of animal research and encourage researchers to adopt more rigorous scientific practices. Ultimately the ARRIVE guidelines will form the basis for a powerful suite of tools and resources that provide optimal support for researchers to improve the design, conduct and reporting of *in vivo* research; this will also benefit research users and stakeholders tasked with assessing the quality and translational value of preclinical research.

Improving reporting should be a community-wide effort, and the working group recognises the importance of engaging others in the evolution of the ARRIVE guidelines. It is essential that this endeavour includes scientists from a range of research fields and countries, and we encourage the community to share their experience and views.

References

1. Collins FS, Tabak LA. Policy: NIH plans to enhance reproducibility. *Nature* 2014;505(7485):612-3.
2. Glasziou P, Altman DG, Bossuyt P, Boutron I, Clarke M, Julious S, et al. Reducing waste from incomplete or unusable reports of biomedical research. *Lancet* 2014;383(9913):267-76.
3. Kilkenny C, Browne WJ, Cuthill IC, Emerson M, Altman DG. Improving bioscience research reporting: the ARRIVE guidelines for reporting animal research. *PLoS Biol* 2010;8(6):e1000412.
4. NC3Rs. ARRIVE: Animal Research Reporting In Vivo Experiments. https://www.nc3rs.org.uk/arrive-animal-research-reporting-vivo-experiments, 2017.
5. Cressey D. UK funders demand strong statistics for animal studies. *Nature* 2015;520(7547):271-2.
6. Baker D, Lidster K, Sottomayor A, Amor S. Two years later: journals are not yet enforcing the ARRIVE guidelines on reporting standards for pre-clinical animal studies. *PLoS Biol* 2014;12(1):e1001756.
7. Florez-Vargas O, Brass A, Karystianis G, Bramhall M, Stevens R, Cruickshank S, et al. Bias in the reporting of sex and age in biomedical research on mouse models. *eLife* 2016;5.
8. Sena E. Impact of an Intervention to Improve Compliance with the ARRIVE Guidelines (IICARus) for the reporting of in vivo animal research. *Eighth International Congress on Peer Review and Scientific Publication*. Chicago, 2017:https://www.pscp.tv/w/1YqJDRBOYWwKV.
9. Avey MT, Moher D, Sullivan KJ, Fergusson D, Griffin G, Grimshaw JM, et al. The devil is in the details: Incomplete reporting in preclinical animal research. *PLoS One* 2016;11(11):e0166733.
10. Moher D, Schulz KF, Simera I, Altman DG. Guidance for developers of health research reporting guidelines. *PLoS medicine* 2010;7(2):e1000217.
11. Eisenach JC, Warner DS, Houle TT. Reporting of preclinical research in anesthesiology: Transparency and enforcement. *Anesthesiology* 2016;124(4):763-5.
12. McGrath JC, Curtis MJ. BJP is changing its requirements for scientific papers to increase transparency. *British Journal of Pharmacology* 2015;172(11):2671-74.
13. Bahor Z, Liao J, Macleod MR, Bannach-Brown A, McCann SK, Wever KE, et al. Risk of bias reporting in the recent animal focal cerebral ischaemia literature. *Clin Sci (Lond)* 2017;131(20):2525-32.
14. Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D, et al. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med* 2001;134(8):663-94.
15. Hurst V, Percie du Sert N. The ARRIVE guidelines survey. *Open Science Framework* 2017:https://osf.io/g8t5q/register/565fb3678c5e4a66b5582f67.
16. Reichlin TS, Vogt L, Wurbel H. The researchers' view of scientific rigor-survey on the conduct and reporting of in vivo research. *PLoS One* 2016;11(12):e0165999.
17. Mogil JS, Macleod MR. No publication without confirmation. *Nature* 2017;542(7642):409-11.
18. Kimmelman J, Mogil JS, Dirnagl U. Distinguishing between exploratory and confirmatory preclinical research will improve translation. *PLoS Biol* 2014;12(5):e1001863.