

Quality of reporting of otorhinolaryngology articles using animal models with the ARRIVE statement

Aren Bezdjian^{1,2,3}, Sjaak F L Klis^{1,2}, Jeroen P M Peters^{1,2}, Wilko Grolman^{1,2} and Inge Stegeman^{1,2}

Laboratory Animals
2018, Vol. 52(1) 79–87
© The Author(s) 2017
Reprints and permissions:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/0023677217718862
journals.sagepub.com/home/lan




Abstract

Research involving animal models is crucial for the advancement of science, provided that experiments are designed, performed, interpreted, and reported well. In order to investigate the quality of reporting of articles in otorhinolaryngology research using animal models, a PubMed database search was conducted to retrieve eligible articles. The checklist of the ARRIVE (Animal Research: Reporting of In Vivo Experiments) guidelines was used to assess the quality of reporting of articles published in ear, nose and throat (ENT) and multidisciplinary journals. Two authors screened titles, abstracts, and full texts to select articles reporting otorhinolaryngology research using in vivo animal models. ENT journals ($n = 35$) reported a mean of 57.1% adequately scored ARRIVE items (median: 58.3%; 95% confidence interval [CI; 53.4–60.9%]), while articles published in multidisciplinary journals ($n = 36$) reported a mean of 49.1% adequately scored items (median: 50.0; 95% CI [46.2–52.0%]). Articles published in ENT journals showed better quality of reporting of animal studies based on the ARRIVE guidelines ($P < 0.05$). However, adherence to the ARRIVE guidelines is generally poor in otorhinolaryngology research using in vivo animal models. The endorsement of the ARRIVE guidelines by authors, research and academic institutes, editorial offices and funding agencies is recommended for improved reporting of scientific research using animal models.

Keywords

quality of reporting, animal studies, publication, otorhinolaryngology

Date received: 13 February 2017; accepted: 12 June 2017

Research involving animal models plays an important role in scientific innovation provided that the experiments are designed, performed, interpreted, and reported well. Recently, a lack of transparency in scientific publications using animal models has been identified.¹ Crucial experimental design elements are frequently ignored in scientific publications, rendering experiments irreproducible.

Kilkenny et al. reported that 59% of investigated articles using animal models failed to adequately report a research hypothesis, objective, or the number and characteristics of the animals used (i.e. species/strain, sex, and age/weight).² The vast majority of these articles did not report randomization of animals (87%) or blinding of researchers (86%). A similar study revealed that 72% of articles using animal models did not report randomization of animals and 98% did not report blinding of researchers.³ Others have reported deficiencies in

reporting important methodological parameters that render experiments irreproducible.⁴ Inadequate reporting greatly hinders the conclusions drawn in scientific publications, especially when they cannot be reproduced due to non-transparent methodology.⁵ Studies may in fact be well-designed and well-conducted. However,

¹Department of Otorhinolaryngology and Head & Neck Surgery, University Medical Center Utrecht, Utrecht, The Netherlands

²Brain Center Rudolf Magnus, University Medical Center Utrecht, Utrecht, The Netherlands

³Department of Experimental Surgery, Faculty of Medicine, McGill University, Montreal, Quebec, Canada

Corresponding author:

J P M Peters, MD, Department of Otorhinolaryngology, University Medical Center Utrecht, Heidelberglaan 100, 3584 CX Utrecht, The Netherlands.
Email: h.p.m.peters@umcutrecht.nl

when poorly reported they may lead to inaccurate interpretation and translation.^{6,7}

In an attempt to optimize the overall quality of reporting of scientific research using animal models, the ARRIVE (Animal Research: Reporting of In Vivo Experiments) guidelines were developed.⁸ These guidelines aim to improve the experimental design, analysis, and reporting of research using animal models.

Using the checklist of the ARRIVE guidelines as a scoring tool, our primary aim was to evaluate the quality of reporting of articles using animal models in otorhinolaryngology research. Furthermore, the quality of reporting of articles published in ear, nose and throat (ENT) journals was compared with those published in multidisciplinary journals.

Methods

Journal selection

The quality of reporting of articles which describe animal experiments in otorhinolaryngology research was compared between two journal categories: ENT journals and multidisciplinary journals. Based on ISI Web of Knowledge impact factors (www.webofknowledge.com, date inspected: 12 June, 2015), the five ENT journals with the highest impact factors in 2012 were selected: Ear & Hearing (*Ear Hear*), Journal of the Association for Research in Otolaryngology (*JARO*), Head & Neck – Journal for the Sciences and Specialties of the Head and Neck (*Head Neck*), Hearing Research (*Hear Res*), and Audiology & Neurotology (*Audiol Neurotol*). None of these journals implemented the ARRIVE guidelines in the 'Instructions to Authors' (date inspected: 12 June 2015). The top five multidisciplinary journals in 2013 were *Nature*, *Science*, *Nature Communications (Nat Commun)*, *Proceedings of the National Academy of Sciences of the United States of America (PNAS)* and *Scientific Reports (Sci Rep)*. Two journals (*Nature* and *Nat Commun*) recommended the ARRIVE guidelines when documenting animal studies (date inspected: 12 June 2015). The included journals and their impact factors are summarized in Table 1.

Search strategy

A PubMed database search was conducted on June 12, 2015 using four predefined filters. First, an adapted version of the ENT filter developed by the Cochrane ENT group was used to retrieve articles conducting research in otorhinolaryngology.⁹ Second, a filter was applied to only retrieve research using animal models.¹⁰ Subsequently, date restrictions were applied per journal

Table 1. Impact factors of the top five ENT and multidisciplinary journals.

Journal	Impact factor*
ENT journals	
Ear & Hearing (<i>Ear Hear</i>)	3.262
Journal of the Association for Research in Otolaryngology (<i>JARO</i>)	2.952
Head & Neck (<i>Head Neck</i>)	2.833
Hearing Research (<i>Hear Res</i>)	2.537
Audiology & Neurotology (<i>Audiol Neurotol</i>)	2.318
Multidisciplinary journals	
Nature	42.531
Science	31.477
Nature Communications (<i>Nat Commun</i>)	10.700
Proceedings of the National Academy of Sciences of the United States of America (<i>PNAS</i>)	9.809
Scientific Reports (<i>Sci Rep</i>)	5.078

*Source: ISI Web of Knowledge 2012, Journal Citations Reports (JCR) via www.webofknowledge.com (accessed on 12 June 2015). ENT: ear, nose and throat.

category to limit the amount of retrieved articles. We searched PubMed for articles published in ENT journals in the year 2014. Since less otorhinolaryngology related articles are published in multidisciplinary journals, we searched for articles conducting animal experiments in otorhinolaryngology research published in multidisciplinary journals from 2010 to 2014. It is important to note that the ARRIVE guidelines were first published in 2010. Thus, studies published in multidisciplinary journals in 2010 might have been written prior to the publication of these guidelines. An analysis was performed to investigate correlations between year of publication and quality of reporting.

The complete search syntax with specific filters is outlined in Supplemental digital content 1 (see <http://journals.sagepub.com/doi/full/10.1177/0023677217718862> for all supplementary materials in this article).

Study selection

Two authors (SFLK and JPMP) independently screened titles, abstracts and full texts of the retrieved articles and selected those reporting in vivo animal experiments. To be considered for inclusion, studies must have assessed preclinical phases of diseases or disorders commonly treated by otorhinolaryngologists. Discrepancies between the two reviewers were discussed until consensus was reached.

Scoring articles

To assess the quality of reporting of articles, two authors (AB and SFLK) independently scored articles using the checklist from the ARRIVE guidelines. The checklist contains 20 points, some with subsections (a, b, c or d). Subsections were considered as separate items for scoring, yielding a total of 38 items. Two items on the ARRIVE checklist (10c and 15b) were optional and were rarely applicable. Therefore, to standardize our assessment of quality of reporting in all articles, these two items were excluded from the analysis. In total, 36 items were scored for each article.

Supplemental digital content 2 summarizes the scoring criteria per item. Articles were reviewed in order to extract all provided information. This included supplementary information; available online or in appendices. No more than five articles per journal category were scored consecutively to distribute possible learning effects evenly across the two journal categories.

Inter-observer agreement

Cohen's kappa value for inter-observer agreement was evaluated to analyze discrepancies among the scorers. Cohen's kappa was calculated for the complete dataset, and per item.

Data analysis

Descriptive statistics including median and mean scores of adequately reported ARRIVE items were calculated. A two-tailed Mann–Whitney *U*-test for two independent samples was used to evaluate significant differences between the two journal categories. Chi-square analysis was used to evaluate each ARRIVE item between the two journal categories. For the articles published in multidisciplinary journals, a correlation between year of publication and quality of reporting was investigated using Spearman's rank correlation coefficient (Spearman's rho).

Statistical tests were performed using the SPSS v20 statistics package (IBM, Armonk, NY, USA). Statistical significance was set at 5%.

Results

Search and study selection

The combined search syntaxes (Supplemental digital content 1) yielded 51 articles published in ENT journals, and 63 articles in multidisciplinary journals. Figure 1 summarizes the search and study selection process.

Of the 51 articles retrieved from ENT journals, 11 were not primary research articles and five did not involve in vivo animal experiments. Therefore, 35 articles

from ENT journals were included in the analysis (*Hear Res*: $n = 15$, *JARO*: $n = 11$, *Head Neck*: $n = 9$).

Of the 63 articles retrieved from multidisciplinary journals, 18 were not related to otorhinolaryngology research, five did not report on primary research and four did not include in vivo animal experiments. Therefore, 36 articles were included in the analysis (*PNAS*: $n = 24$, *Nature*: $n = 4$, *Nat Commun*: $n = 4$, *Sci Rep*: $n = 3$, *Science*: $n = 1$). Eight articles were published in multidisciplinary journals that endorse the ARRIVE guidelines. Six articles were published in multidisciplinary journals in 2010, eight in 2011, eight in 2012, nine in 2013 and five in 2014.

The numbers of retrieved and selected articles per journal are summarized in Table 2.

Overall quality of reporting scores

The 35 articles published in ENT journals reported a mean of 57.1% adequately scored items (95% confidence interval [CI]: 53.4–60.9%; median: 58.3%). The 36 articles published in multidisciplinary journals reported a mean of 49.1% adequately scored items (95% CI: 46.2–52.0%; median: 50.0%). The overall difference between the journal categories was statistically significant (Mann–Whitney *U*-test, $P = 0.001$), suggesting that ENT journals adhered better to the ARRIVE guidelines.

For the articles published in multidisciplinary journals, there was no statistically significant correlation between the year of publication and the number of adequately reported ARRIVE items ($P = 0.083$). Moreover, there was no significant difference in the quality of reporting between the eight articles published in multidisciplinary journals that endorsed the ARRIVE guidelines (*Nature*, *Nat Commun*), and the 28 articles published in journals that did not endorse the ARRIVE guidelines: 51.4% (95% CI: 45.6–57.2%, median: 54.2%) compared with 48.4% (95% CI: 45.1–51.7%, median: 50.0%), respectively.

Quality of reporting for specific items

When examining ARRIVE items separately, five items (6a, 7a, 9c, 14, 18b) were scored significantly higher in the articles published in ENT journals (Figure 2, Table 3). These items assessed if the study reported the number of experimental and control groups (6a), information on the drug dose, site and route of administration, and surgical procedure and equipment used (7a), welfare-related assessments and interventions carried out prior, during, or after experiments (9c), information on health status of animals prior to treatment or testing (14) and study limitations (18b).

Several items were not adequately reported in both journal categories: 10 items were reported less

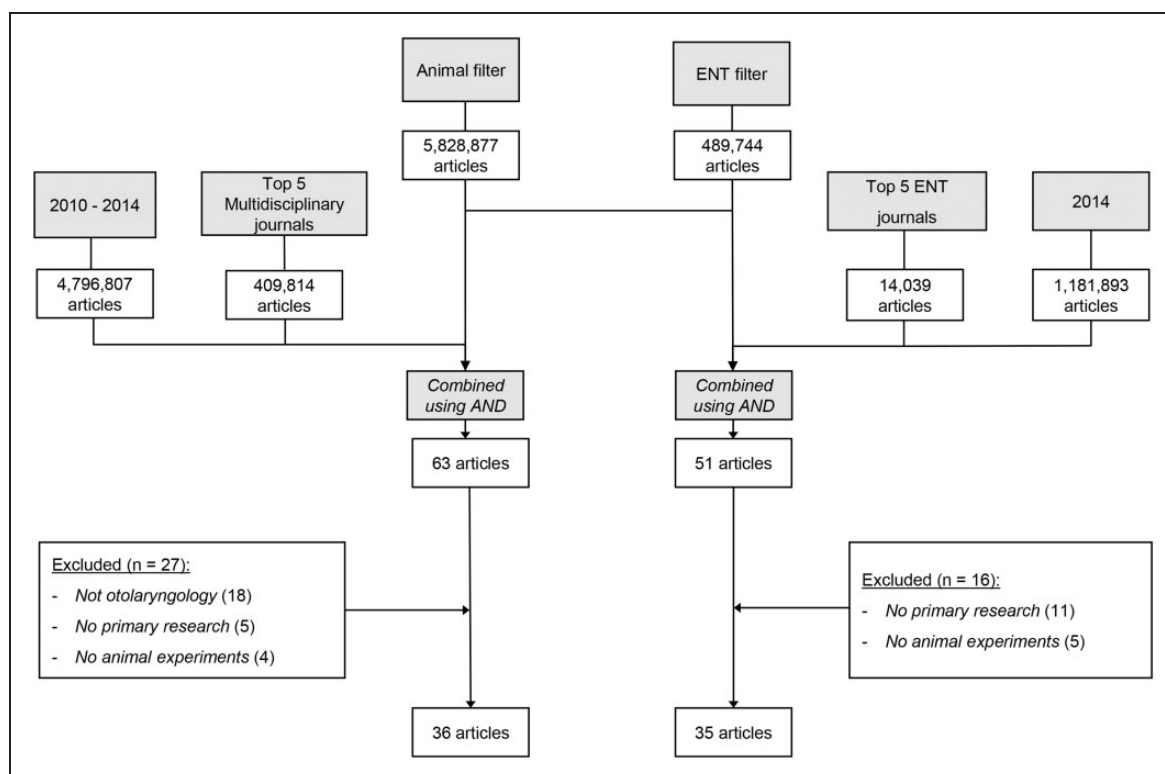


Figure 1. Flow chart demonstrating study selection.
ENT: ear, nose and throat.

Table 2. Retrieved articles by search and included articles following study selection.

	Retrieved	Included
ENT journals		
<i>Hear Res</i>	26	15
<i>JARO</i>	13	11
<i>Head Neck</i>	11	9
<i>Ear Hear</i>	1	0
<i>Audiol Neurotol</i>	0	0
Total	51	35
Multidisciplinary journals		
<i>PNAS</i>	39	24
<i>Nature</i>	13	4
<i>Nat Commun</i>	6	4
<i>Sci Rep</i>	3	3
<i>Science</i>	2	1
Total	63	36

ENT: ear, nose and throat.

frequently than 20% in both journal categories (Figure 2). These items include the time of day when experiments were carried out (7b), the rationale behind the choice of the specific anesthetic, its dose and route of administration opted for (7d), information regarding

housing of animals (9a), sample size calculation (10b), allocation of the animals to groups (11a,b), methods used to assess whether the data met the assumptions of the statistical approach (13c), reporting of adverse events (17a,b), and implications of the experimental methods or findings for the replacement, refinement or reduction of the use of animals in research (18c).

Inter-observer agreement

Out of a total number of 2556 scored items, 158 (6.1%) were scored differently. Cohen's kappa value for inter-observer agreement was 0.87 (standard error = 0.10). A Cohen's kappa score between 0.61 and 0.80 suggests a good agreement between independent scorers.¹¹ Cohen's kappa value for inter-observer agreement per item is presented in Supplemental digital content 3. The inter-observer agreement was high for most items, and there were no Cohen's kappa values lower than 0.3.¹¹

Discussion

The present study evaluated the quality of reporting of scientific publications using animal models in otorhinolaryngology research. Articles published in ENT journals adhered better to the ARRIVE guidelines than articles published in multidisciplinary journals.

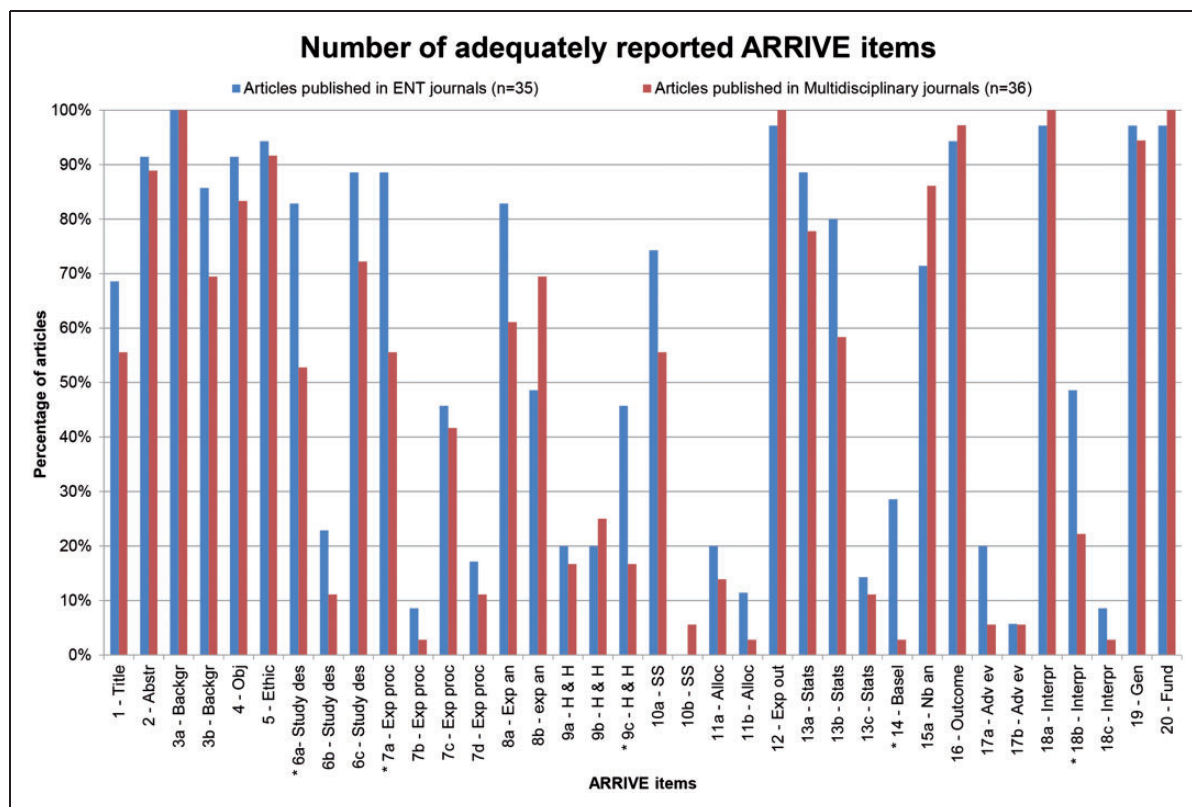


Figure 2. Adequately reported scores per item according to the ARRIVE guidelines.

*Significant difference (χ^2) between journal type.

Abstr: abstract; Adv ev: adverse events; Alloc: allocation of animals; Backgr: background; Exp an: experimental animals; Basel: baseline data; Exp out: experimental outcomes; Exp proc: experimental procedures; Fund: funding; Gen: generalisability; H & H: housing and husbandry; Interpr: interpretation; Nb an: number of animals; Obj: objectives; Study des: study design; SS: sample size; Stats: statistical method.

Therefore, articles published in multidisciplinary journals with high impact factors do not have a superior overall quality of reporting in otorhinolaryngology research using animal models. Similarly, MacLeod et al. have identified significantly fewer reporting of randomization in articles published in journals with high impact factors.¹² Our findings are contrary to reports investigating the quality of reporting of randomized controlled trials¹³ and systematic reviews¹⁴ in otorhinolaryngology research, where ENT journals underperformed.

Interpretation of results

Although ENT journals showed better quality of reporting, adherence to the ARRIVE guidelines is generally poor in otorhinolaryngology research for both journal categories. Items such as choice of the specific anesthetic, dose and route of administration (7d) and information regarding the housing of animals (9a) were rarely (<20% of all studies) reported. This information is essential for accurate replication of animal

experiments, as it may influence study outcomes.¹⁵ Prager et al. reported that housing and husbandry information of animals have the potential to influence responses of rodents, and thus alter study outcomes.¹⁶ Our findings also revealed that sample size calculation for the number of animals chosen per group (10b) and allocation of the animals to groups (11a,b) were rarely reported (<20% of all studies). These two items are essential for optimizing statistical design, and for fulfilling ethical obligations, as they aim to reduce potential bias and the number of animals used in research.^{1,17}

Articles published in multidisciplinary journals often described additional experiments alongside the animal model. As such, the animal experiment could have not been the primary focus of the study. Nevertheless, all multidisciplinary journals included had a methodology section containing information relating to the animal experiments. These sections do not have word limits that may have justified the missing information.

Similar outcomes are found in other disciplines. Gulin et al. performed a quality assessment review of animal studies for Chagas disease by comparing studies

Table 3. Proportion of adequately reported ARRIVE items per journal category.

Item:	1: Title	2: Abstract	3a: Background	3b: Background	4: Objective	5: Ethic	*6a: Study design	6b: Study design	6c: Study design	*7a: Exp proc	7b: Exp proc	7c: Exp proc	7d: Exp proc	8a: Exp an	8b: Exp an	9a: H & H	9b: H & H	*9c: H & H
Articles published in ENT journals (n=35)	69%	91%	100%	86%	91%	94%	83%	23%	89%	89%	9%	46%	17%	83%	49%	20%	20%	46%
Articles published in multidisciplinary journals (n=36)	56%	89%	100%	69%	83%	92%	53%	11%	72%	56%	3%	42%	11%	61%	69%	17%	25%	17%
P value chi ² (2-tailed)	0.330	1.000	NA	0.155	0.478	1.000	0.011*	0.220	0.135	0.003*	0.357	0.813	0.514	0.064	0.094	0.767	0.778	0.011*
Item:	10a: SS	10b: SS	11a: Allocation	11b: Allocation	12: Exp outcome	13a: Stats	13b: Stats	13c: Stats	*14: Basel	15a: Nb an	16: Outcome	17a: Adv ev	17b: Adv ev	18a: Interpretation	18b: Interpretation	18c: Interpretation	19: Gen	20: Fund
Articles published in ENT journals (n=35)	74%	0%	20%	11%	97%	89%	80%	14%	29%	71%	94%	20%	6%	97%	49%	9%	97%	97%
Articles published in multidisciplinary journals (n=36)	56%	6%	14%	3%	100%	78%	58%	11%	3%	86%	97%	6%	6%	100%	22%	3%	94%	100%
P value chi ² (2-tailed)	0.137	0.493	0.541	0.199	0.493	0.343	0.072	0.735	0.003*	0.155	0.614	0.085	1.000	0.493	0.026*	0.357	1.000	0.493

*Significant difference [chi²] between journal categories.

Adv ev: adverse events; Basel: generalisability; Fund: funding; Exp an: experimental animals; Exp outcome: experimental outcome; Exp proc: experimental procedures; H & H: housing and husbandry; n: number; Nb an: number of animals; SS: sample size; Stats: statistical methods.

published before and after the publication of the ARRIVE guidelines. In line with our findings, their study revealed that items such as randomization (16%) and sample size calculations (7%) were rarely reported.¹⁸ Ting et al. investigated interventional animal studies in rheumatology and reported missing information such as randomization (17%), sample size calculation (0%), allocation (0%), housing, husbandry and welfare-related information (5%), and implications for replacement, refinement or reduction of the use of animal assessments (0%).¹⁹ These items are essential to reduce bias in scientific research, and to make experiments transparent and replicable.²⁰ Furthermore, Schwarz et al. reviewed publications on preclinical research for the treatment of mucositis/peri-implantitis, Freshwater et al. conducted reviews on animal research published in plastic surgery journals, and Tsilidis et al. investigated the reporting of animal models for neurological diseases.^{21–23} All studies concluded that there is an urgent need for improving the quality of reporting when using animal models.

Methodological considerations

Strengths of our study include a search strategy that could be reproduced to evaluate the quality of reporting of animal studies in other disciplines. To account for learning effects, the two authors who independently scored 2556 items did not score more than five articles consecutively per journal category. The limitations of the study include firstly a subjective assessment by the two independent scorers. The scorers were also not blinded to which journal category the paper belonged. However, the high inter-observer agreement demonstrated that both reviewers had fairly similar judgment (Supplemental digital content 3). Second, in order to obtain a sufficient amount of articles, we included articles published in multidisciplinary journals from 2010–2014, whereas we included articles published in ENT journals in 2014 only. Since the ARRIVE guidelines were first published in 2010, articles published that year could not have had access to these guidelines. Nevertheless, a subanalysis revealed no correlation between the year of publication and the quality of reporting. A third limitation is that *Nature* and *Nat Commun* have recommended that authors use the ARRIVE checklist (date inspected: 12 June 2015). However, no statistical difference was found in the quality of reporting between articles published in multidisciplinary journals that endorsed the ARRIVE guidelines and those that did not. Finally, our findings were only based on eight journals (ENT journals: *Hear Res*: $n = 15$, *JARO*: $n = 11$, *Head Neck*: $n = 9$; multidisciplinary journals: *PNAS*: $n = 24$, *Nature*: $n = 4$, *Nature Comm*: $n = 4$, *Science Rep*: $n = 3$, *Science*: $n = 1$).

Reporting guidelines

Evidence that clinical trials lacked crucial methodological information led to the development of the Consolidated Standards for Reporting Trials (CONSORT) statement, which is now implemented by many journals and funding agencies. Implementing the CONSORT statement has been shown to drastically improve the quality of reporting of clinical trials.^{24–27} By contrast, the development of the ARRIVE guidelines did not enhance quality of reporting when comparing articles appearing before and after the ARRIVE guidelines were published.²⁸ Baker et al. showed that reporting of animal research in *PLoS* journals, which have been early proponents of the ARRIVE guidelines, still remained low.²⁸ In our sample, we also showed that there was no improvement in the quality of reporting with increasing year of publication (2010–2014). Therefore, we recommend a stronger endorsement of the ARRIVE guidelines from authors, journal editors and funding agencies.

Conclusion

Although articles using animal models published in ENT journals have better quality of reporting scores than those published in multidisciplinary journals, adherence to the ARRIVE guidelines is generally poor in otorhinolaryngology research. There is an urgent need to improve the quality of reporting in otorhinolaryngology research using animal models. Editorial endorsement of the ARRIVE guidelines from authors, research and academic institutes, editorial offices, and funding agencies is warranted to optimize quality of reporting.

Declaration of Conflicting Interests

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

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

1. Landis SC, Amara SG, Asadullah K, et al. A call for transparent reporting to optimize the predictive value of pre-clinical research. *Nature* 2012; 490: 187–191.
2. Kilkeny C, Parsons N, Kadyszewski E, et al. Survey of the quality of experimental design, statistical analysis and reporting of research using animals. *PLoS One* 2009; 4: e7824.
3. Hess KR. Statistical design considerations in animal studies published recently in cancer research. *Cancer Res* 2011; 71: 625.

4. Sena E, van der Worp HB, Howells D and Macleod M. How can we improve the pre-clinical development of drugs for stroke? *Trends Neurosci* 2007; 30: 433–439.
5. van der Worp HB, Howells DW, Sena ES, et al. Can animal models of disease reliably inform human studies? *PLoS Med* 2010; 7: e1000245.
6. Crossley NA, Sena E, Goehler J, et al. Empirical evidence of bias in the design of experimental stroke studies – a metaepidemiologic approach. *Stroke* 2008; 39: 929–934.
7. Macleod MR, van der Worp HB, Sena ES, et al. Evidence for the efficacy of NXY-059 in experimental focal cerebral ischaemia is confounded by study quality. *Stroke* 2008; 39: 2824–2829.
8. Kilkenny C, Browne WJ, Cuthill IC, et al. Improving bioscience research reporting: the ARRIVE guidelines for reporting animal research. *PLoS Biol* 2010; 8: e1000412.
9. The Editorial Team, Cochrane ENT Group. *About the Cochrane Collaboration (Cochrane Reviews Group (CRGs))*, <http://onlinelibrary.wiley.com/o/cochrane/clabout/articles/ENT/sect0-meta.html> (2017, accessed 12 June 2015).
10. Hooijmans A, Leenaars M and Ritskes-Hoitinga M. Enhancing search efficiency by means of a search filter for finding all studies on animal experimentation in PubMed. *Lab Anim* 2010; 44: 170–175.
11. Altman DG. *Practical statistics for medical research*. London: Chapman and Hall/CRC, 1991.
12. Macleod MR, McLean AL, Kyriakopoulou A, Serghiou S, de Wilde A and Sherratt N. Risk of bias in reports of in vivo research: a focus for improvement. *PLoS Biol* 2015; 13: e1002301.
13. Peters JPM, Hooft L, Grolman W and Stegeman I. Assessment of the quality of reporting of randomized controlled trials in otorhinolaryngologic literature – adherence to the CONSORT statement. *PLoS One* 2015; 10: e0122328.
14. Peters JPM, Hooft L, Grolman W and Stegeman I. Reporting quality of systematic reviews and meta-analyses of otorhinolaryngologic articles based on the PRISMA statement. *PLoS One* 2015; 10: e0136540.
15. Von der Behrens W. Animal models of subjective tinnitus. *Neural Plast* 2014; 2014: 741452.
16. Prager EM, Bergstrom HC, Grunberg NE and Johnson LR. The importance of reporting housing and husbandry in rat research. *Front Behav Neurosci* 2011; 5: 1–4.
17. Festing MF and Altman DG. Guidelines for the design and statistical analysis of experiments using laboratory animals. *ILAR J* 2002; 43: 320.
18. Gulin JE, Rocco DM and Garcia-Bournissen F. Quality of reporting and adherence to ARRIVE guidelines in animal studies for Chagas disease preclinical drug research: a systematic review. *PLoS Negl Trop Dis* 2015; 9: e0004194.
19. Ting KH, Hill CL and Whittle SL. Quality of reporting of interventional animal studies in rheumatology: a systematic review using the ARRIVE guidelines. *Int J Rheum Dis* 2015; 18: 488–494.
20. Hirst JA, Howick J, Aronson JK, Roberts N, Perera R and Koshiaris C. The need for randomization in animal trials: an overview of systematic reviews. *PLoS One* 2014; 9: e98856.
21. Schwarz F, Iglhaut G and Becker J. Quality of 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* 2012; 39: 63–72.
22. Freshwater MF. Laboratory animal research published in plastic surgery journals in 2014 has extensive waste: a systematic review. *J Plast Reconstr Aesthet Surg* 2015; 68: 1485–1490.
23. Tsilidis KK, Panagiotou OA, Sena ES, et al. Evaluation of excess significance bias in animal studies of neurological diseases. *PLoS Biol* 2013; 11: e1001609.
24. Plint AC, Moher D, Morrison A, et al. Does the CONSORT checklist improve the quality of reports of randomised controlled trials? A systematic review. *Med J Aust* 2006; 185: 263–267.
25. Kane RL, Wang J and Garrard J. Reporting in randomized clinical trials improved after the adoption of the CONSORT statement. *J Clin Epidemiol* 2007; 60: 241–249.
26. Altman DG. Endorsement of the CONSORT statement by high impact medical journals: survey of instructions for authors. *BMJ* 2005; 330: 1056–1057.
27. Hopewell S, Altman DG, Moher D and Schulz KF. Endorsement of the CONSORT statement by high impact factor medical journals: a survey of journal editors and journal ‘Instructions to Authors’. *Trials* 2008; 9: 20.
28. Baker D, Lidster K, Sottomayor A and 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: e1001756.

Résumé

La recherche impliquant des modèles animaux est cruciale pour l'avancement de la science à condition que les expériences soient bien conçues, réalisées, interprétées et rapportées. Une recherche dans la base de données Pubmed a été effectuée afin de récupérer les articles admissibles pour enquêter sur la qualité des informations rapportées par les articles de recherche otorhinolaryngologique utilisant des modèles animaux. La liste de vérification des lignes directrices ARRIVE [de l'anglais Animal Research : Reporting of In Vivo Experiments [Recherche animale : communication des expériences in vivo]] a été utilisée pour évaluer la qualité des informations rapportées par les articles publiés dans les journaux et revues multidisciplinaires d'ORL. Deux auteurs ont examiné les titres, résumés et textes complets pour sélectionner les

articles rapportant la recherche menée en otorhinolaryngologie sur des modèles animaux in vivo. Les revues d'ORL (n=35) montraient une moyenne de 57,1 % d'éléments obtenant un score ARRIVE adéquat (médiane : 58,3 % ; intervalle de confiance à 95 % (IC) [53,4 - 60,9 %]), tandis que les articles publiés dans les revues multidisciplinaires (n=36), montraient une moyenne de 49,1 % d'éléments obtenant un score adéquat (médiane : 50,0 ; IC à 95 % [46,2 - 52,0 %]). Les articles publiés dans des revues d'ORL démontraient une meilleure qualité des rapports d'études animales sur la base des lignes directrices ARRIVE (p<0,05). Cependant, l'adhésion aux lignes directrices ARRIVE est généralement pauvre dans le secteur de la recherche en otorhinolaryngologie utilisant des modèles animaux in vivo. L'utilisation systématique des lignes directrices ARRIVE par les auteurs, les instituts universitaires et de recherche, les bureaux de rédaction et les organismes de financement est recommandée afin d'améliorer les rapports de recherche scientifique menée sur des modèles animaux.

Abstract

Forschung unter Einsatz von Tiermodellen ist für den wissenschaftlichen Fortschritt von wesentlicher Bedeutung, vorausgesetzt, die Experimente werden gut konzipiert, durchgeführt, interpretiert und berichtet. Zur Untersuchung der Berichts-Qualität von Artikeln der HNO-Heilkunde-Forschung mittels Tiermodellen wurde eine PubMed-Datenbank-Suche zur Ermittlung relevanter Artikel durchgeführt. Anhand der Checkliste der ARRIVE (Animal Research: Reporting of In Vivo Experiments)-Richtlinien wurde die Berichts-Qualität von in HNO-Fachzeitschriften und multidisziplinären Zeitschriften veröffentlichten Artikeln bewertet. Zwei Autoren sichteten Titel, Abstracts und Volltext zwecks Auswahl von Artikeln über HNO-Forschung mittels In-vivo-Tiermodellen. HNO-Fachzeitschriften (n=35) berichteten ein Mittel von 57,1 % adäquat bewertete ARRIVE-Elemente (Medianwert: 58,3 %; 95 % Konfidenzintervall (CI) [53,4 - 60,9 %]), während in multidisziplinären Zeitschriften erschienene Artikel (n=36) ein Mittel von 49,1 % adäquat bewertete ARRIVE-Elemente berichteten (Medianwert: 50,0; 95 % CI [46,2 - 52,0 %]). In HNO-Fachzeitschriften veröffentlichte Artikel zeigten eine bessere Berichts-Qualität von Tierstudien basierend auf den ARRIVE-Richtlinien (p<0,05). Dennoch ist die Einhaltung der ARRIVE-Richtlinien in der HNO-Heilkunde-Forschung mittels In-vivo-Tiermodellen generell dürftig. Die Unterstützung der ARRIVE-Richtlinien durch Autoren, Forschungseinrichtungen und wissenschaftliche Institute, Redaktionen und Finanzierungsstellen wird für eine optimierte Berichterstattung über wissenschaftliche Forschung unter Einsatz von Tiermodellen empfohlen.

Resumen

La investigación con modelos animales es crucial para el avance de la ciencia siempre que los experimentos estén bien diseñados, realizados, interpretados y registrados. Para poder investigar la calidad del registro de artículos en la investigación de otorrinolaringología utilizando modelos animales, se realizó una búsqueda en la base de datos Pubmed para recopilar artículos elegibles. Se utilizó la lista de comprobación de las directrices ARRIVE (Investigación Animal: Registro de Experimentos In Vivo, por sus siglas en inglés) para evaluar la calidad del registro de artículos publicado en revistas ENT y publicaciones multidisciplinarias. Dos autores analizaron títulos, resúmenes y textos completos para seleccionar artículos que registraran investigaciones de otorrinolaringología utilizando modelos animales in vivo. Revistas ENT (n=35) arrojaron un promedio de 57,1 % de artículos ARRIVE adecuadamente puntuados (media: 58,3 %; 95 % intervalo de confianza (CI) [53,4 - 60,9 %]), mientras que los artículos publicados en revistas multidisciplinarias (n=36) registraron un promedio de 49,1 % de artículos adecuadamente puntuados (media: 50,0; 95 % CI [46,2 - 52,0 %]). Los artículos publicados en revistas ENT mostraron una mayor calidad del registro de estudios animales basado en las directrices ARRIVE (p<0,05). No obstante, la adherencia a las directrices ARRIVE es por lo general baja en la investigación de otorrinolaringología utilizando modelos animales in vivo. Se recomienda un respaldo a las directrices ARRIVE por parte de autores, institutos académicos y de investigación, oficinas editoriales y agencias de financiación para un registro mejorado de la investigación científica utilizando modelos animales.