CLINICAL ORAL IMPLANTS RESEARCH

Review Article

The risk of bias of animal experiments in implant dentistry: a methodological study

Clovis Mariano Faggion Jr
Karla Tatiana Diaz
Luisiana Aranda
Frank Gabel
Stefan Listl
Marco Antonio Alarcón

Authors' affiliations:
Clovis Mariano Faggion Jr Department of Periodontology and Operative Dentistry, Faculty of Dentistry, University of Münster, Münster, Germany
Karla Tatiana Diaz, Academic Department of Clinical Stomatology, Section of Integral Oral Implantology, Cayetano Heredia Peruvian University, Lima, Perú
Luisiana Aranda, Marco Antonio Alarcón, Academic Department of Clinical Stomatology, Section of Periodontology and Implants, Cayetano Heredia Peruvian University, Lima, Perú
Frank Gabel, Stefan Listl, Department of Conservative Dentistry, Heidelberg University, Heidelberg, Germany
Stefan Listl, Max Planck Institute for Social Law and Social Policy, Munich Center for the Economics of Aging, Munich, Germany

Corresponding author:
Clovis Mariano Faggion Jr
Faculty of Dentistry
University of Münster
Waldeyerstraße 30
48149 Münster
Germany
Tel.: +49 0 251/83 47061
Fax: +49 0 251/83 47134
e-mail: clovisfaggion@yahoo.com

Key words: animal experimentation, methods, risk of bias, SYRCLE's tool, treatment effect

Abstract

Objectives: To evaluate the risk of bias (ROB) in reports of randomised controlled trials (RCTs) of animal experiments published in implant dentistry, and to explore the association between animal experiment characteristics and ROB.

Material and methods: We searched the MEDLINE (via PubMed), SCOPUS and SciELO databases from 2010 to March 2015 for reports of RCTs of animal experiments published in implant dentistry. We evaluated independently and in duplicate the ROB of these experiments by the use of a tool specifically developed to evaluate ROB in animal studies, the SYRCLE's tool. ROB was judged as low, high or unclear (when there was not enough information to judge ROB). We used univariate and multivariate logistic regression analyses to evaluate the association of specific study characteristics and extent of ROB.

Results: We initially selected 850 publications and 161 reports of animal experiments were included. For a total of 1449 entries (records), 486 (34%) were rated as low ROB. High ROB was attributed to 80 (6%) of entries, and 883 (60%) entries were rated as unclear ROB. The characteristics "impact factor" (IF), reporting of standard error (SE) and reporting of confidence interval (CI) were significantly associated with low ROB in some SYRCLE domains.

Conclusions: A substantial number of items with unclear ROB were observed in this sample of animal experiments in implant dentistry. Furthermore, the present findings suggest that implant dentistry animal experiments published in journals with higher IF and better report of measures of precision; that is, CI and SE may have lower ROB than those not having these characteristics.

Introduction

Risk of bias (ROB) is an important concept that should be addressed in all types of studies. Knowing the level of ROB gives information about whether or not the results of a study are reliable and useful. For example, a clinician would likely be reluctant to apply a clinical technique if the information on effectiveness and safety is based on a randomized controlled trial (RCT) with high ROB. The high ROB in such RCT would mean that the treatment effect estimates may be over- or underestimated (Higgins et al. 2011). Thus, the technique may be ineffective or even cause some harm.

Basic research in the form of animal experiments is normally conducted to understand disease mechanisms and guide further research in the form of clinical trials (Faggion 2015a). For any of these objectives, the evaluation of ROB is pivotal to understand the validity of the findings. An animal experiment with high ROB may not adequately reflect the mechanism of the disease or may provide only inaccurate information on the effectiveness of interventions. Therefore, the evaluation of ROB in animal experiment is a pivotal step for the critical evaluation of reported effectiveness and safety in pre-clinical studies.

The objective of the present methodological study is to evaluate ROB in reports of RCTs of animal experiments published in implant dentistry by the use a specific tool developed to evaluate ROB in animal experiments, the SYstematic Review Centre for Laboratory animal Experimentation (SYRCLE) (Hooijmans et al. 2014). A second objective was to identify potential predictors of

Date:
Accepted 20 March 2016

To cite this article:
Faggion CM Jr, Diaz KT, Aranda L, Gabel F, Listl S, Alarcón MA. The risk of bias of animal experiments in implant dentistry: a methodological study.
Clin. Oral Impl. Res. **00**, 2016, 1–7
doi: 10.1111/clr.12852

low ROB in the different domains evaluated by the SYRCLE's tool.

Material and methods

Eligibility criteria

We included in the present project only RCT on animal experiments published in implant dentistry. We considered an animal experiment a RCT when authors performed at random any phase of the active therapy (i.e. during implant or prosthetic placements). Therefore, we included reports of studies on interventions. Furthermore, we also included studies of implant use for oral and maxillofacial purposes placed intra- or extra-orally.

We excluded non-RCT experiments or animal experiments where the randomization was performed after the therapies were applied. Furthermore, we excluded reports of *in vitro* or studies involving humans. We also excluded reports of animal RCT published in languages other than English and Spanish.

Database search

Two authors (KD, LA) independently searched the MEDLINE (via PubMed), SCOPUS and SciELO electronic databases to identify reports of RCT involving animal experiments in implant dentistry published from September 2010 to 31 March 2015. These search limits were applied due to the availability of the SYRCLE's tool which was published in 2010. Authors also searched for potential reports in the reference lists of included reports of RCT initially retrieved from the databases. Unpublished literature was also sought in the ProQuest Dissertations and Theses Database (pqdtopen.proquest.com) and in the OpenGrey (www.opengrey.eu) databases. The search strategy is presented in Table 1.

Reports selection

First, two authors (KD, LA) evaluated titles and abstracts of reports to determine eligibility for initial inclusion; then, they scrutinized full texts of articles to determine whether the studies met the inclusion criteria. The authors documented excluded articles with corresponding reasons for exclusion. Two authors (KD, LA) performed study selection independently and in duplicate. They discussed any disagreement regarding the inclusion or exclusion of articles until consensus was achieved.

Data extraction

Data were directly extracted from the reports to a standardized table including the SYRCLE's checklist items. Data were extracted independently by two authors (KD, LA). A third author (MAA) checked the precision of responses. Disagreements on data extraction were resolved by discussion until consensus was achieved.

The SYRCLE's risk of bias tool

The Cochrane ROB tool (Higgins et al. 2011) was used as a starting point for the development the SYstematic Review Centre for Laboratory animal Experimentation (SYRCLE's) RoB tool (Hooijmans et al. 2014). The SYRCLE's ROB tool consists of a domain-based instrument with 10 entries related to six types of bias: selection bias, performance bias, detection bias, attrition bias, reporting bias and other biases (Hooijmans et al. 2014). These 10 entries are organized in sub-items in the form of questions that support a Yes/No/Unclear answer.

A domain-based tool requires a careful analysis and interpretation of the data reported to understand whether the information (or lack of it) may have caused any potential bias to a specific study domain. Evaluation by means of such a tool is more realistic than by means of a scale or checklist because the former involves subjectivity in the assessment (Higgins et al. 2011). Scales, for example, require assigning "weights" to the different items to generate a summary score. This approach might not be feasible because it is challenging to understand the impact of each item on quality. Thus, a domain-based approach seems preferable to capture the true validity of a study.

In the present project, we applied nine items of the tool (item 4: *were the animals randomly housed during the experiment?* was not applied). The authors of the SYRCLE's tool suggest that the tool may be adjusted to the type of sample analysed. We understand that the omission of this item will not interfere with the validity of our findings. More detailed information on the SYRCLE's tool is reported elsewhere (Hooijmans et al. 2014). A list of main SYRCLE items used in this study is reported in the Supporting information.

Risk of bias evaluation

Two authors (KD, LA) independently answered the questions from the SYRCLE's ROB tool. Disagreements in the answers were resolved by discussion until consensus was achieved. In a second stage, the two authors independently rated the ROB of each of the nine entries by taking into account the sub-items answers. ROB was rated as low, high and unclear (Higgins et al. 2011). Disagreements in the ROB judgment were resolved by discussion and consensus.

Assessor training

One author (CMF) conducted a pilot test of the standardized form with a sample of randomly selected studies to ensure the accuracy and consistency of data extraction and ROB evaluation. The two assessors (KD, LA) then used the form to extract data from other randomly selected studies. Between rounds of data extraction, assessors comprehensively discussed the outcomes to improve homogeneity.

Data analysis

We analysed the possible associations between the ROB values (dependent variable) and independent predictors including the publication characteristics impact factor (IF), reporting of sample size *a priori* and *post hoc*, reporting of *P*-values, standard errors (SE) and confidence intervals (CI), the study design, number of therapies, the type of sponsorship, the type of animals used and the region the author originates from and if it was a pilot study.

Binary logistic regression (both univariate and multivariate) was used on SYRCLE-ROB

Table 1. **Search strategy used in Medline (via PubMed) to select animal experiments in implant dentistry**

Keywords
(1) "Non-humans primates"
(2) rat OR rats OR mice
(3) pig OR pigs OR swine
(4) dog OR dogs
(5) monkey OR monkeys OR baboon OR baboons OR cebidae
(6) rabbit OR rabbits OR hares
(7) cat OR cats OR felidae
(8) goat OR goats
(9) sheep
(10) #1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9
(11) implant OR implants OR fixture OR fixtures
(12) random*
(13) #10 AND #11 AND #12 (sort by: Author Filters: Publication date from 2010/09/01 to 2014/09/01; Other Animals) (Updated from 2014/09/02 to 2015/03/31)

1, 2 and 7 (with the two SYRCLE categories "low ROB" and "unclear"). Ordinal logistic regression was used on SYRCLE-ROB 8 and 10 (with the three ordinal SYRCLE categories "low ROB," "unclear" and "high ROB"). We checked potential outliers with the help of plotting leverage and Cook's distance as well as logistic regression assumptions by inspecting residual plots. To assess model fit, individual goodness-of-fit chi-square tests were conducted to compare the full models with the corresponding null models.

In the models SYRCLE 1, 2 and 7, odds ratios (OR) smaller than one imply that the probability for "low ROB" is higher than the probability for "unclear." Vice versa, odds ratios larger than one imply that the probability for "low ROB" is lower than the probability for "unclear." As for the models SYRCLE 8 and 10, the interpretation is quite similar: odds ratios smaller than one imply that the probability for a paper to be "low ROB" is highest and the probability to be "high ROB" is lowest (and vice versa). The implicit order for this ordinal logistic regression model is "low ROB," "unclear," "high ROB."

The models for SYRCLE 1 and SYRCLE 2 (containing all predictor variables) do not fit the data significantly better than the null model, whereas the models for SYRCLE 7, 8 and 10 fit the data better than the corresponding null models in a highly significant manner. To improve model fit, stepwise variable selection was conducted; however, no statistically significant improvements could be reached. The diagnostic checks for outliers (leverage, Cook's distance) did not show abnormalities, so that all data points could be kept in the model. To justify the use of ordinal logistic regression for SYRCLE 8 and 10, we checked the proportional odds assumption graphically.

The level of statistical significance for all tests was prespecified at 0.05. Statistical analyses were performed with R version 3.2.1, software (The R Foundation for Statistical Computing c/o Institute for Statistics and Mathematics Wirtschaftsuniversität Wien, Vienna, Austria).

Results

Number of reports

We initially identified 850 publications. After the assessment of titles and abstract, 632 publications were excluded. We evaluated further the full texts of the articles and more 59 publications were excluded. We found nine publications in the reference lists of articles retrieved from the electronic databases search. In addition, we updated the literature search from September 2014 to March 2015 which yielded further publications for inclusion. In total, the literature search resulted in 161 eligible publications (160 published in English). The literature search process is illustrated in Fig. 1, and publications included in the analysis as well as the list of excluded publications (with reasons for exclusion) are listed in the Supporting information.

Characteristics of animal experiments

One hundred and seven (66%) animal experiments had a split-mouth design. Fifty-four (46%) studies were supported by non-profit organizations, and the remainder received various levels of industry funding. The IF value of the journals where the animal experiments were published ranged from 0.200 to 8.312 (median 2.565, n = 158). Five species of animal were used in these experiments, most commonly dogs and rodents (42% and 24% of the sample, respectively). The number of animals used in the experiments ranged from 1 to 96 (median, n = 12). Table 2 shows the characteristics of the reports included in the study.

Risk of bias of animal experiments

For a total of 1449 entries, 486 (34%) were rated as low ROB. High ROB was rated in 80 (6%) entries, and 883 entries (60%) were rated as unclear ROB. The item with the greatest number of entries rated as low ROB (n = 161, 100%) was item 9 (*are reports of the study free of selective outcome reporting?*). The item with the greatest number of entries rated as high ROB (n = 60, 37%) was item 10 (*was the study apparently free of other problems that could result in high ROB?*). The item with the greatest number of entries rated as unclear ROB (n = 161, 100%) was item 6 (*were animals selected at random for outcome assessment?*). The SYRCLE items assessed with the total number of articles rated as low, high and unclear ROB are reported in Table 3.

Characteristics of reports and risk of bias

The univariate and multivariate binary logistic regression analyses are reported in Tables 4 and 5. From 161 articles included, only 155 were included in the regression analysis. Six articles were excluded because they could not provide complete cases for including in the logistic regression models. Logistic regression analysis was performed only in five SYRCLE domains (1, 2, 7, 8, 10). For other SYRCLE domains, regressions analysis was not applicable because of lack of variation in the dependent variable due to very high numbers of unclear items, or the vast majority of one type of ROB in the other domains.

Univariate analysis

In the univariate analysis, the following characteristics were significantly associated with low ROB: reporting of CI in SYRCLE domain 7 (OR 0.18, 95% CI 0.04, 0.87), IF in SYRCLE domain 10, although borderline (OR 0.77, CI 0.59, 0.99), and "the first author of the article being based in Asia" (OR 0.23, 95% CI 0.11, 0.52) (Table 4).

Multivariate analysis

In the multivariate analysis, the following characteristics were significantly associated with low ROB: The reporting of SE in SYRCLE domain 2 (OR 0.26, 95% CI 0.07, 0.93), and the reporting of IF in SYRCLE domain 8, although borderline (OR 0.72, 95% CI 0.52, 0.99) (Table 5). As a robustness check, we also re-ran our analyses using an alternative definition for the IF control variable (IF for the year of publication instead of IF for the year 2014) and obtained qualitatively similar results.

Discussion

Summary of findings

In this sample of 161 articles reporting animal experiments in implant dentistry, more than 50% of all entries evaluated across domains were at unclear ROB. When there is no information (or partial information only) to adequately address bias, the judgement of ROB is "unclear." This output clearly hinders a comprehensive evaluation of the data. It is important to mention that an unclear judgment means that a specific domain can be potentially at high or low ROB. Nevertheless, some domains were better reported such as the domain related to selective outcome reporting (item 9), which was judged at low ROB in all 161 animal experiments evaluated. These findings may be related to the methodology used: we considered no selective outcome reporting when authors consistently reported outcomes in the material and methods and results sections of the articles (Higgins et al. 2011). As far as we know, there is no public registry for animal research protocols, and therefore, the evaluation of these protocols to identify potential selective outcome reporting is not (yet) a reality in animal research.

Logistic regression analyses demonstrated that the report of some methodological issues in the article such as report of CI and SE might be associated with domains with lower ROB. These outcomes have nevertheless

wide CI, and therefore, some sort of uncertainty of these results may be expected. The present article is part of a project that evaluates ROB and measures of precision in animal experiments in implant dentistry. The article evaluating measures of precision in this sample of animal experiments will be published elsewhere (Faggion et al. in press). Interestingly, IF was significantly associated with low ROB in two SYRCLE domains. These findings suggest that animal experiments published in scientific journals with higher IF might have lower ROB than those experiments published in journals with lower IF. However, these findings need to be interpreted with caution due to the CI values that are borderlines.

The interpretation of the regression analysis is the following: a one-unit increase in the independent variable decreases the odds for a specific domain to be "low ROB" by the factor reported in the regression tables. Therefore, this implicates that coefficients >1 in the table decrease the probability for "low ROB" and coefficients <1 increase the probability for "low ROB." For example, in the multivariate analysis, a one-unit increase in the predictor "impact factor" decreases the odds for SYRCLE 1 to be "low ROB" by the factor 0.79. In other words, the greater the "impact factor," the higher the probability of this specific domain to be at low ROB (when compared to "unclear" ROB).

The results of the logistic regression analysis need to be interpreted with due care, because the sample size is rather low given the logistic regression framework (which requires a higher sample size than, e.g. linear regression) and the amount of covariates included in the model. This sometimes also led to serious separation issues (no observations corresponded to certain combinations of independent and dependent variables) and, accordingly, considerably widened confidence intervals. The high number of "unclear" items further weakens our model (for SYRCLE 1, 2, 7 as an independent variable, we had to model "low ROB" vs. "unclear" due to the mere missingness of "high ROB" items).

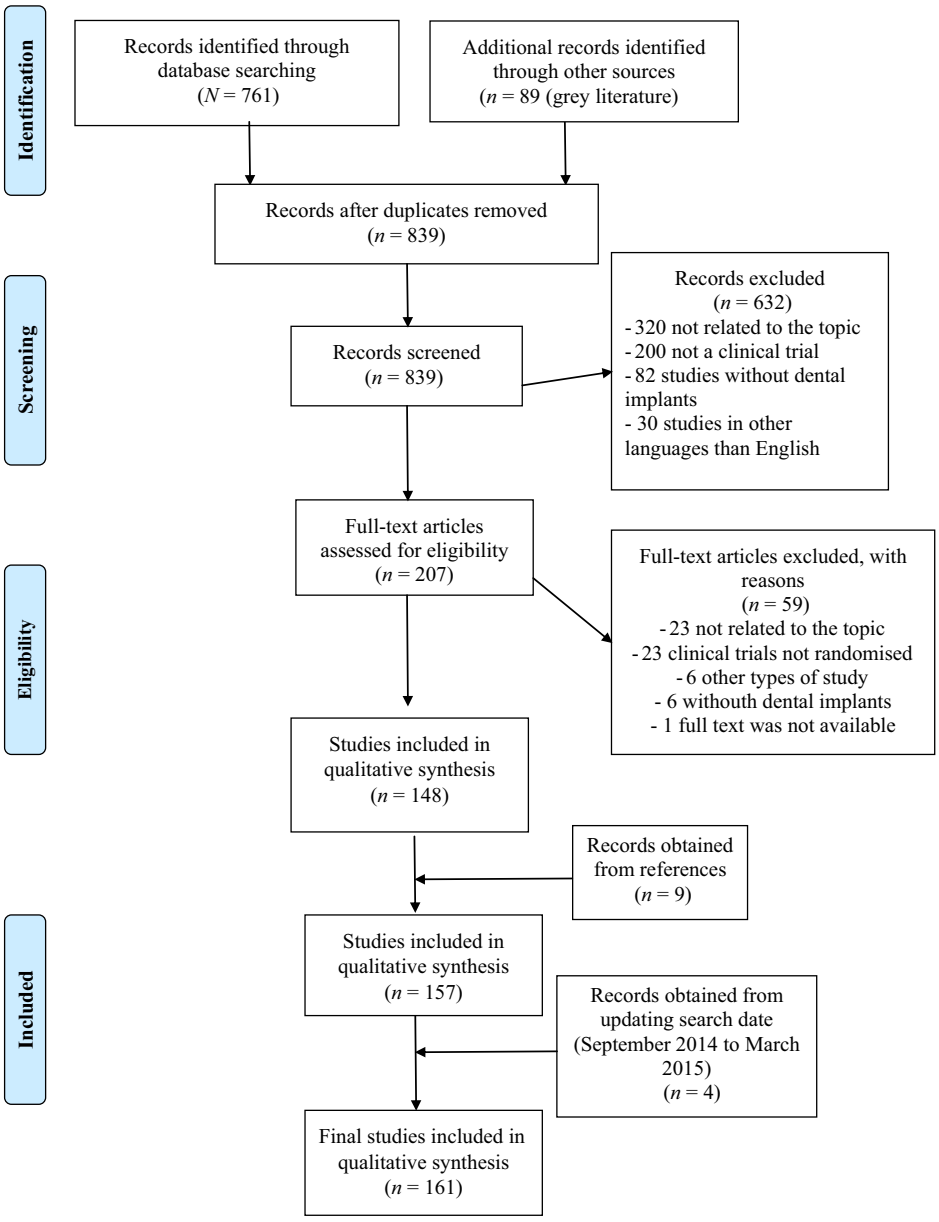

Fig. 1. PRISMA-style flow diagram of the literature search process.

Table 2. **Characteristics of the reports of animal experiments included in the study (*N* = 161)**

Variable	Categories	*N*	%
Impact factor	Numerical	158	2.45 (mean)
Sample size calculated *a priori*	Not reported (reference)	156	96.9
	Yes	5	3.1
Sample size calculated *post hoc*	No (reference)	161	100
	Yes	0	0
P-values reported	No (reference)	9	5.6
	Yes	152	94.4
Confidence interval reported	No (reference)	148	91.9
	Yes	7	4.3
	Graphically	6	3.7
Standard error reported	No (reference)	142	88.2
	Yes	19	11.8
Pilot study	No (reference)	150	93.2
	Yes	11	6.8
Study design	Parallel (reference)	54	33.5
	Split-mouth	107	66.5
	Cross-over	0	0
	Others	0	0
Number of therapies	Numerical	161	3.45 (mean)
Animals	Dogs (reference)	67	41.6
	Minipigs	12	7.5
	Pigs	8	5.0
	Rats	38	23.6
	Rabbits	34	21.1
	Mice	1	0.6
	Goats	1	0.6
Region	Europe (reference)	57	35.4
	Arabia	6	3.7
	Asia	54	33.5
	North America	19	11.8
	South America	24	14.9
Type of sponsorship	100% pharmaceutical/device company funded (reference)	21	13.0
	100% non-profit funded	74	46.0
	Mixed funding, for example non-profit and industrial funding	7	4.3
	Provision of drug or device only	29	18.0
	Undisclosed funding	26	16.1
	Author (s) of the review work for the company or declare to have COI such as stocks	4	2.5

Table 3. **SYRCLE main items evaluated in this study (with the respective number of articles rated as low, high and unclear risk of bias [ROB])**

SYRCLE item	Low ROB	High ROB	Unclear ROB
1. Was the allocation sequence adequately generated and applied?	31	0	130
2. Were the groups similar at baseline or were they adjusted for confounders in the analysis?	88	0	73
3. Was the allocation to the different groups adequately concealed during?	3	0	158
5. Were the caregivers and/or investigators blinded from knowledge which intervention each animal received during the experiment?	7	0	154
6. Were animals selected at random for outcome assessment?	0	0	161
7. Was the outcome assessor blinded?	34	0	127
8. Were incomplete outcome data adequately addressed?	89	20	52
9. Are reports of the study free of selective outcome reporting?	161	0	0
10. Was the study apparently free of other problems that could result in high ROB?	73	60	28

Strengths and limitations of the present study

The present study has clear strengths. Firstly, to our knowledge this is the first study in to evaluate ROB with the SYRCLE's tool in animal experiments published in implant dentistry. Secondly, this is the first study in dentistry aimed to identify potential predictors of level of ROB in different domains of these animal experiments. Thirdly, we also reported information from a representative sample of reports of animal experiments in the field of implant dentistry.

This study has, nevertheless, some limitations. It was not possible to adequately evaluate the level of ROB in many entries due to lack of information. Consequently, regression analyses could not be applied in four from the nine domains evaluated. This lack of information is not exclusive of animal research, but it is also prevalent in research with humans (Bialy et al. 2014), although several guidelines have been published in the last years to improve the quality of reporting in human trials (Simera et al. 2010), and several scientific journals already endorsed them. More recently, some guidelines to improve reporting of animal research have been published (Kilkenny et al. 2010), and they have been used in animal experiments published in implant dentistry (Schwarz et al. 2012; Stadlinger et al. 2012; Thoma et al. 2012; Vignoletti & Abrahamsson 2012; Delgado-Ruiz et al. 2015).

It is important to emphasize the differences from the previous projects evaluating the quality of reporting with ARRIVE guidelines and the present one, which evaluates the ROB with SYRCLE's tool. Guidelines such ARRIVE suggest how authors of experiments should report their research to improve the understanding of the reader about the methodology used. The approach used in SYRCLE refers to the evaluation of the internal validity of the experiment, and therefore, some sort of judgement is expected to adequately address ROB.

Therefore, both approaches (the evaluation of reporting and ROB) are inter-dependent. The adequate reporting of the experiment will allow the adequate evaluation of the ROB, that is, in the end, the important outcome for the judicious use of the animal research data.

It is important to emphasize the limitation of any methodological tool to evaluate whether the information reported in the article is true or false. For example, this applies to scenarios of scientific misconduct of researchers, by fabricating or falsifying the data (Fanelli 2009), or to fraudulent behaviour throughout the peer-review process (Haug 2015). Methodological tools for evaluating ROB are not necessarily fit for purpose to detect false assertion with respect to the reported content of a paper. So even if a publication may be judged to have low ROB according to the relevant tools, this may not necessarily mean that the underlying

Table 4. **Univariate binary logistic regression results with SYRCLE categories 1, 2, 7 as dependent variable and univariate ordinal logistic regression for categories 8, 10 for the included SR (*N* = 155) (coefficients are reported as odds ratios [OR])**

Predictor variables	Category or unit	SYRCLE1 OR (95% CI)	SYRCLE2 OR (95% CI)	SYRCLE7 OR (95% CI)	SYRCLE8 OR (95% CI)	SYRCLE10 OR (95% CI)
Impact factor	Numerical	1.01 (0.74, 1.38)	1.02 (0.80, 1.31)	0.79 (0.60, 1.06)	0.79 (0.60, 1.03)	0.77 (0.59, 0.99)
SS calc. *a priori*	No (reference)					
	Yes	1.00 (0.11, 9.28)	0.29 (0.03, 2.62)	0.96 (0.10, 8.90)	5.17 (0.96, 27.81)	0.59 (0.13, 2.78)
P-values reported	No (reference)					
	Yes	1.15 (0.23, 5.84)	0.66 (0.17, 2.56)	0.46 (0.06, 3.85)	Empty; separation	0.29 (0.07, 1.19)
CI reported	No (reference)					
	Yes	0.64 (0.12, 3.48)	0.45 (0.08, 2.38)	0.18 (0.04, 0.87)	0.99 (0.22, 4.43)	12.35 (1.47, 103.5)
	Graphically	Empty; separation	0.56 (0.10, 3.15)	Empty; separation	2.08 (0.48, 9.13)	3.06 (0.53, 17.53)
SE reported	No (reference)					
	Yes	5.09 (0.65, 39.73)	0.38 (0.13, 1.11)	0.39 (0.14, 1.08)	1.04 (0.43, 2.50)	0.94 (0.38, 2.32)
Pilot design	No (reference)					
	Yes	0.56 (0.14, 2.30)	0.28 (0.06, 1.34)	Empty; separation	2.40 (0.69, 8.33)	1.49 (0.43, 5.17)
Study design	Parallel (reference)					
	Split-mouth	0.51 (0.20, 1.28)	1.39 (0.71, 2.74)	0.88 (0.38, 2.02)	0.56 (0.30, 1.05)	1.18 (0.63, 2.20)
No. of therapies	Numerical	0.85 (0.68, 1.06)	0.93 (0.77, 1.14)	0.85 (0.68, 1.06)	1.14 (0.95, 1.37)	1.05 (0.88, 1.26)
Types of animals	Dogs (reference)					
	Minipigs	1.17 (0.28, 4.84)	0.33 (0.08, 1.35)	1.00 (0.24, 4.15)	3.80 (1.16, 12.45)	1.93 (0.57, 6.54)
	Pigs	1.17 (0.22, 6.37)	0.60 (0.13, 2.72)	2.33 (0.27, 20.44)	0.26 (0.03, 2.21)	1.37 (0.38, 4.96)
	Rabbits	2.19 (0.73, 6.56)	0.83 (0.36, 1.93)	3.33 (0.90, 12.41)	1.79 (0.78, 4.08)	0.61 (0.27, 1.38)
	Rats	4.57 (1.25, 16.73)	0.90 (0.40, 2.01)	1.07 (0.42, 2.74)	2.38 (1.09, 5.20)	0.88 (0.42, 1.89)
Region of research	Europe (reference)					
	Arabia	0.59 (0.10, 3.60)	0.48 (0.08, 2.84)	Empty; separation	2.69 (0.48, 15.15)	3.64 (0.67, 19.76)
	Asia	1.59 (0.60, 4.21)	0.79 (0.37, 1.69)	2.45 (1.44, 6.57)	1.14 (0.53, 2.44)	0.23 (0.11, 0.52)
	North America	2.36 (0.48, 11.65)	0.48 (0.16, 1.46)	1.37 (0.39, 4.78)	2.08 (0.72, 5.95)	5.05 (1.50, 17.00)
	South America	0.84 (0.27, 2.56)	0.89 (0.34, 2.33)	1.40 (0.45, 4.42)	3.65 (1.47, 9.06)	1.15 (0.46, 2.85)
Type of sponsorship	100% company funded (reference)					
	Non-profit	5.25 (1.73, 15.95)	1.22 (0.45, 3.33)	3.10 (1.00, 9.56)	5.26 (1.43, 19.38)	Empty; separation
	Mixed	1.50 (0.22, 10.08)	1.63 (0.26, 10.10)	Empty; separation	14.01 (1.93, 101.63)	Empty; separation
	Provision only	4.69 (1.20, 18.34)	1.52 (0.48, 4.76)	0.95 (0.29, 3.11)	6.50 (1.59, 26.56)	Empty; separation
	Undisclosed	2.13 (0.60, 7.57)	2.11 (0.63, 7.06)	1.80 (0.47, 6.90)	3.20 (0.72, 14.28)	Empty; separation
	Author inv. in comp.	1.25 (0.20, 25.37)	1.63 (0.19, 13.93)	Empty; separation	14.01 (1.14, 172.3)	Empty; separation

SS, sample size; CI, confidence interval; SE, standard error.

Table 5. **Multivariate binary logistic regression results with SYRCLE categories 1, 2, 7 as a dependent variable and multivariate ordinal logistic regression for categories 8, 10 for the included articles (*N* = 155 [odds ratios (OR)])**

Predictor variables	Category or unit	SYRCLE1 OR (95% CI)	SYRCLE2 OR (95% CI)	SYRCLE7 OR (95% CI)	SYRCLE8 OR (95% CI)	SYRCLE10 OR (95% CI)
Impact factor	Numerical	0.79 (0.51, 1.21)	1.10 (0.81, 1.49)	0.76 (0.52, 1.11)	0.72 (0.52, 0.99)	0.91 (0.42, 1.94)
SS calc. *a priori*	No (reference)					
	Yes	1.03 (0.06, 17.11)	0.31 (0.03, 3.64)	0.66 (0.04, 9.74)	7.57 (0.95, 60.61)	0.46 (0.01, 26.63)
P-values reported	No (reference)					
	Yes	0.29 (0.04, 2.35)	0.45 (0.08, 2.66)	0.09 (0.01, 1.29)	Empty; separation	0.76 (0.02, 30.72)
CI reported	No (reference)					
	Yes	0.74 (0.09, 6.46)	0.35 (0.05, 2.61)	0.49 (0.07, 1.29)	1.80 (0.22, 14.90)	Empty; separation
	Graphically	Empty; separation	0.48 (0.07, 3.16)	Empty; separation	0.92 (0.17, 5.08)	Empty; separation
SE reported	No (reference)					
	Yes	3.43 (0.35, 33.51)	0.26 (0.07, 0.93)	0.56 (0.14, 2.26)	0.74 (0.24, 2.26)	0.33 (0.03, 3.38)
Pilot design	No (reference)					
	Yes	0.36 (0.05, 2.51)	0.15 (0.02, 1.00)	Empty; separation	6.21 (1.15, 33.60)	1.64 (0.06, 44.48)
Study design	Parallel (reference)					
	Split-mouth	0.76 (0.19, 3.04)	2.09 (0.81, 5.41)	1.71 (0.44, 6.67)	0.81 (0.33, 1.97)	0.74 (0.10, 5.29)
No. of therapies	Numerical	0.78 (0.57, 1.05)	0.93 (0.73, 1.19)	0.82 (0.61, 1.10)	1.25 (0.98, 1.60)	1.08 (0.71, 1.64)
Types of animals	Dogs (reference)					
	Minipigs	3.23 (0.50, 20.69)	0.33 (0.07, 1.64)	3.52 (0.45, 27.28)	4.06 (0.92, 17.97)	0.47 (0.01, 22.00)
	Pigs	3.40 (0.39, 29.64)	0.29 (0.05, 1.72)	6.93 (0.54, 88.46)	0.18 (0.01, 4.52)	1.11 (0.04, 28.46)
	Rabbits	3.64 (0.78, 16.93)	1.20 (0.43, 3.34)	4.25 (0.82, 22.03)	1.02 (0.37, 2.80)	0.67 (0.09, 4.93)
	Rats	7.22 (1.05, 49.54)	1.23 (0.40, 3.84)	2.49 (0.56, 11.18)	1.87 (0.63, 5.59)	0.50 (0.05, 5.29)
Region of research	Europe (reference)					
	Arabia	0.52 (0.06, 4.70)	0.15 (0.02, 1.16)	Empty; separation	9.66 (0.98, 95.16)	1.55 (0.05, 44.85)
	Asia	0.70 (0.18, 2.75)	0.51 (0.19, 1.36)	1.80 (0.49, 6.59)	0.63 (0.24, 1.68)	0.61 (0.02, 17.74)
	North America	2.69 (0.36, 20.31)	0.47 (0.12, 1.93)	1.51 (0.26, 8.80)	2.40 (0.63, 9.08)	13.61 (0.21, 885.0)
	South America	0.31 (0.06, 1.52)	0.61 (0.18, 2.05)	1.41 (0.30, 6.55)	2.34 (0.76, 7.26)	3.54 (0.10, 122.5)
Type of sponsorship	100% company funded (reference)					
	Non-profit	9.68 (2.09, 8.63)	1.16 (0.33, 4.07)	4.45 (0.99, 20.08)	11.09 (2.01, 61.08)	Empty; separation
	Mixed	0.81 (0.07, 25.22)	2.70 (0.33, 21.89)	Empty; separation	11.00 (1.18, 102.7)	Empty; separation
	Provision only	5.01 (0.99, 18.54)	2.06 (0.50, 8.60)	0.60 (0.13, 2.77)	6.20 (1.10, 35.09)	Empty; separation
	Undisclosed	3.78 (0.77, 15.18)	3.43 (0.80, 14.72)	2.23 (0.42, 11.87)	2.68 (0.39, 18.27)	Empty; separation
	Author inv. in comp.	0.79 (0.04, 1.21)	2.48 (0.15, 40.31)	Empty; separation	8.86 (0.40, 196.2)	Empty; separation

SS, sample size; CI, confidence interval; SE, standard error.

reporting of ROB of the respective publication was distortion-free.

Comparison with other studies

The ROB of animal experiments in periodontology and implant dentistry has been evaluated in the past (Faggion et al. 2011) with the assessment of some domains sensitive to bias reported in the Cochrane handbook for systematic reviews (Higgins et al. 2011). In Faggion et al. (2011) reports of animal experiments (sample published up to 2010) had 78% of items assessed at unclear ROB. Our more recent sample of reports (from 2010) demonstrated an improvement in the percentage of items with unclear ROB (60%), although the methodologies used in the two projects cannot be compared directly.

Implications of the present findings

The lack of information clearly hinders the adequate evaluation of the data, and this was demonstrated by the high percentage of "unclear" items evaluated across the reports. Therefore, it can be difficult to separate quality of reporting from the quality of the study, because the information necessary to interpret and judge this quality is not always available. One strategy that could potentially reduce the number of unclear items is the direct contact with authors of the animal experiments to clarify issues. Nevertheless, obtaining data from authors of studies is not an easy task. For example, the response rate from authors is low (Young & Hopewell 2011). Moreover, one can argue that there is no sound approach to obtain reliable information from the authors. Thus, ideally, the evaluation of ROB should be limited to the scientific article reported (Faggion 2015b).

Future developments in animal research

Future developments in the field of evaluating the ROB in animal experiments should be focused on shedding more light on the unclear issues. Thus, the strict fulfilment of guidelines such as ARRIVE should be a requirement for submitting reports of animal experiments. In other words, to have the paper sent to peer-review, editors should carefully evaluate whether these items were reported in an adequate level that allows the evaluation of bias without the need of contacting authors of the studies to get the "extra" information. The idea is to provide more comprehensive information for ROB analysis and, consequently, to reduce the prevalence of "unclear" judgments. Eventually, higher standards for reporting ROB should result in better design, planning and conduct of animal studies, thereby also optimizing their informative value for the design of clinical studies and reducing wasteful research (e.g. Kleinert & Horton 2014).

Conclusions

Reports of animal experiments are not comprehensively reported to allow an adequate ROB assessment. The present findings suggest that reports of animal experiments reporting measures of precision and published in journals with higher IF might have lower ROB in some domains than those reports not having these characteristics.

Acknowledgements: This study was self-funded by the authors.

Conflict of interest

The authors declare they have no conflict of interests.

References

Bialy, L., Vandermeer, B., Lacaze-Masmonteil, T., Dryden, D.M. & Hartling, L. (2014) A meta-epidemiological study to examine the association between bias and treatment effects in neonatal trials. *Evidence-Based Child Health* **9**: 1052–1059.

Delgado-Ruiz, R.A., Calvo Guirado, J.L. & Romanos, G.E. (2015) Bone grafting materials in critical defects in rabbit calvariae. A systematic review and quality evaluation using ARRIVE guidelines. *Clinical Oral Implants Research* 1–5. doi: 10.1111/clr.12614.

Faggion, C.M. Jr., Aranda, L., Diaz, K., Shih, M.C., Tu, Y.K. & Alarcón, M. (in press) The quality of reporting of measures of precision in animal experiments in implant dentistry: a methodological study. *The International Journal of Oral and Maxillofacial Implants*.

Faggion, C.M. Jr. (2015a) Animal research as a basis for clinical trials. *European Journal of Oral Sciences* **123**: 61–64.

Faggion, C.M. Jr. (2015b) Risk of bias assessment should not go beyond reporting assessment. *Journal of Clinical Epidemiology* **72**: 126–127.

Faggion, C.M. Jr. Giannakopoulos, N.N. & Listl, S. (2011) Risk of bias of animal studies on regenerative procedures for periodontal and peri-implant bone defects – a systematic review. *Journal of Clinical Periodontology* **38**: 1154–1160.

Fanelli, D. (2009) How many scientists fabricate and falsify research? A systematic review and meta-analysis of survey data. *PLoS ONE* **4**: e5738.

Haug, C.J. (2015) Peer-review fraud – hacking the scientific publication process. *New England Journal of Medicine* **373**: 2393–2395.

Higgins, J.P.T., Altman, D.G. & Sterne, J.A.C., eds. (2011) Chapter 8: Assessing risk of bias in included studies. In: Higgins, J.P.T. & Green, S., eds. Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0 [updated March 2011]. The Cochrane Collaboration. Available from www.cochrane-handbook.org.

Hooijmans, C.R., Rovers, M.M., de Vries, R.B., Leenaars, M., Ritskes-Hoitinga, M. & Langendam, M.W. (2014) SYRCLE's risk of bias tool for animal studies. *BMC Medical Research Methodology* **14**: 43.

Kilkenny, C., Brown, W.J., Cuthill, I.C., Emerson, M. & Altman, D.G. (2010) Improving bioscience research reporting: the ARRIVE guidelines for reporting animal research. *PLoS Biology* **8**: e1000412.

Kleinert, S. & Horton, R. (2014) How should medical science change? *Lancet* **383**: 97–198.

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. *Journal of Clinical Periodontology* **39**(Suppl. 12): 63–72.

Simera, I., Moher, D., Hirst, A., Hoey, J., Schulz, K.F. & Altman, D.G. (2010) Transparent and accurate reporting increases reliability, utility, and impact of your research: reporting guidelines and the EQUATOR Network. *BMC Medicine* **8**: 24.

Stadlinger, B., Pourmand, P., Locher, M.C. & Schulz, M.C. (2012) Systematic review of animal models for the study of implant integration, assessing the influence of material, surface and design. *Journal of Clinical Periodontology* **39** (Suppl. 12): 28–36.

Thoma, D.S., Martin, I.S., Mühlemann, S. & Jung, R.E. (2012) Systematic review of pre-clinical models assessing implant integration in locally compromised sites and/or systemically compromised animals. *Journal of Clinical Periodontology* **39** (Suppl. 12): 37–62.

Vignoletti, F. & Abrahamsson, I. (2012) Quality of reporting of experimental research in implant dentistry. Critical aspects in design, outcome assessment and model validation. *Journal of Clinical Periodontology* **39**(Suppl. 12): 6–27.

Young, T. & Hopewell, S. (2011) Methods for obtaining unpublished data. *The Cochrane Database of Systematic Reviews* **11**: MR000027.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Data S1. List of included and excluded articles.