IN BRIEF

1 Describes how quality assessment of randomised clinical trial (RCT) reports can be used for locating sources of bias.
1 Shows that most RCTs in dental and medical research were inadequately reported.
1 Demonstrates a large variation in the quality of RCT reports in dental and medical research.
1 Shows that the journal impact factor was not correlated with the quality of RCT reports..

Quality of reporting randomised clinical trials in dental and medical research

P. Sjögren[1] and A. Halling[2]

Objective To assess 1) the quality of reporting randomised clinical trials in dental (RCT-Ds) and medical research (RCT-Ms), 2) the quality of RCT reports in relation to the journal impact factor, 3) the source of funding, and 4) the quality of RCT-Ds in different areas of dental research.
Design Random samples of 100 RCT-Ds and 100 RCT-Ms published in 1999 were evaluated for quality of reporting under blinded conditions with the Jadad quality assessment scale. In addition, correlation between the quality scores and journal impact factor or source of funding, as well as area of dental research were analysed.
Results The quality of RCT-Ds and RCT-Ms published in 1999 was generally inadequate. The quality was largely equivalent in RCT-Ds and RCT-Ms. There was no correlation between the quality scores and the journal impact factor or the source of funding. Some differences were found in the quality scores between different areas of dental research.
Conclusions The results from these RCT-Ds and RCT-Ms show that most of them were imperfect in the reporting of methodology and trial conduct. There is a clear need to improve the quality of trial reporting in dental and medical research.

Randomised clinical trials (RCTs) are regarded as the most reliable method of evaluating the effects of interventions in healthcare.[1–3] RCTs are also considered the 'golden standard' for providing research evidence for interventions in evidence-based healthcare.[3] The validity and reliability of trial results are, however, largely dependent on the study design and the methodology in its conduct. Reporting of RCTs has been a subject of concern that has led to an international agreement on how trials ideally should been reported in medical journals, such as the CONSORT statement.[4,5] This includes a checklist of 21 items related to different aspects of a trial report that are considered important in the publication of an RCT.[4, 5]

Jadad et al. defined the quality of a trial, with emphasis on the methodological quality, as 'the confidence that the trial design, conduct, and analysis have minimised or avoided biases in its treatment comparisons'.[6] In addition, the quality of a trial report was given a somewhat different and more suitable definition 'providing information about the design, conduct, and analysis of the trial'.[6] Both definitions relate to attempts to eliminate bias.

There are several scales and checklists available for quality assessment of RCTs.[1] One widely used, reliable, and to our knowledge, only validated quality scale has been developed by Jadad et al.[1,6] It focuses on the methods for random allocation, double-blinding and withdrawals and drop-outs. The total scores range from 0 to 5 points, where trials with 0–2 points are considered to be of poor quality, and those with 3–5 points represent high quality RCTs.[6] The Jadad scale has been widely used in different areas of medicine,[7-12] but has not yet found its way to dental research. Quality assessments of RCTs in dental research (RCT-Ds) are scarce.[13,14] Thus, we conducted this study with the primary aim of evaluating and comparing the quality of trial reports in a random sample of RCT-Ds and RCTs in medical research (RCT-Ms). Secondary aims were to assess the reporting of allocation concealment, the quality of RCTs in relation to the journal impact factor and the source of funding, but also to evaluate the quality of RCT-Ds in different areas of dental research.

MATERIAL AND METHODS
Medline searches
The Medline database (Entrez PubMed, www.ncbi.nlm.nih.gov) was searched for RCT-Ds and RCT-Ms (September 2000), using the MeSH terms: 'dentistry' and 'medicine', respectively, limited to publication year '1999', publication type 'randomized controlled trial', and language 'English'.

Inclusion and exclusion criteria
The RCTs were scrutinised for relevance to dental and medical research, respectively. RCTs covering other research areas were excluded from the study. Moreover, articles that were found during the quality assessments to be non-randomised trials or in vitro experimental studies were excluded from further analysis. The proportion of false inclusions in the Medline searches was calculated.

Outcome measures
The primary outcome measures were the Jadad RCT quality scores.[1,6] Secondary outcome measures were the reporting of allocation concealment,[15] the relation of quality scores to journal impact factor or source of funding, and the quality of RCT-Ds in different areas of dental research (the main topic covered by the trial).

[1]Postgraduate student, [2]*Associate Professor
*Correspondence to: Arne Halling, Department of Health Sciences, Kristianstad University, SE-291 88 Kristianstad, Sweden.
Email:arne.halling@hv.hkr.se

Sample size

Power calculations with 95% power (p=0.05) in order to detect differences of one point (or equivalence) in the quality scores between RCT-Ds and RCT-Ms gave a minimum sample size of 67 RCTs per group. We chose to randomise 100 RCT-Ds and 100 RCT-Ms, using a computer-generated random number table.

Randomisation and blinding

The RCTs were randomised and coded by an independent laboratory technician. The laboratory technician also obtained, photocopied in duplicate and masked one copy of each RCT with an opaque black marker pen, deleting the authors names, titles and affiliations, the journal title, country of origin, source of funding and acknowledgements.[6] The technician coded the masked RCTs (computer-generated random number table) and kept the code and the unmasked copies concealed from the authors during the quality assessments.

Quality assessments

A validated quality scale (Jadad scale) was used for quality assessments of the RCTs.[1,6] The scale includes three items directly related to the validity of RCTs, giving a total score of 0 to 5 points (0–2 points=poor quality, 3–5 points=high quality), described in detail elsewhere.[6] Briefly, when it was stated that random allocation was used, one point was given. One additional point was given if the method for random allocation was appropriate. If an inappropriate method was used, one point was withdrawn. In addition, one point was given if it was stated that the study was double-blinded. If the method for double-blinding was appropriate, an additional point was given; inappropriate one point was deducted. Reporting of withdrawals and drop-outs was given one point if the study stated the number and reasons of subjects who were included in the study, but who discontinued, in each group.[6]

The rater (PS) was trained by assessing RCTs that were previously rated with the Jadad scale.[7, 11] In order to ensure the consistency of the assessments throughout the study, all RCTs were rated at least twice.[15] Furthermore, twenty randomly selected RCTs (computer-generated random table) were assessed by both authors to verify the reliability of the scorings, and doubtful cases were solved by consensus. In addition, reporting of allocation concealment was assessed. When all ratings were completed, the random code for masking was broken. No quality scores were changed after this point.

Data analyses and statistics

After the quality assessments, the scores were paired with the unmasked RCTs for analysis of the secondary outcome measures. Journal impact factors for 1999 were obtained from the Journal Citation Reports® database (Institute for Scientific Information). The source of funding was, when stated, registered as 'research or community funded' or 'industry funded', or as a combination of both. Statistics were analysed with SPSS 9.0 for Windows, with non-parametric methods using Kruskal-Wallis test, Mann-Whitney U test, and Spearman's rank correlation coefficient (r_s). P levels below 5% were considered significant.

RESULTS

Medline searches

The Medline searches gave a total of 251 search hits for RCT-Ds and 181 for RCT-Ms. Of these, 100 RCT-Ds and 100 RCT-Ms were randomly allocated to this study. Reports covering other research areas than dentistry (n=5) or medicine (n=1) were excluded, as were reports that were non-RCT-Ds (n=3) or non-RCT-Ms (n=10), or were *in vitro* studies in dental (n=9) or medical (n=0) research.

A total of 83 RCT-Ds and 89 RCT-Ms remained throughout the study for analysis. Thus, the proportion of false inclusions in the Medline searches were 17% for RCT-Ds and 11% for RCT-Ms. The RCTs were published in 101 different journals and originated from 26 countries worldwide.

Randomisation

All RCTs were reported as randomised, but an adequate method of random allocation was reported in less than one third of the RCT-Ds and RCT-Ms. An inadequate method was reported in 6% of the RCT-Ds and RCT-Ms. The remaining two thirds of the RCT-Ds and RCT-Ms did not report the method of random allocation (Table 1).

Table 1. The number and percentage (within parentheses) of randomised clinical trials in dental (RCT-Ds) and medical research (RCT-Ms), in 1999, that reported randomisation, double–blinding, withdrawals and dropouts, and allocation concealment

Variable	RCT-Ds (n=83)	RCT-Ms (n=89)
Randomisation		
Reported,	83 (100)	89 (100)
Adequately	22 (27)	25 (28)
Inadequately	5 (6)	5 (6)
Not reported	0 (0)	0 (0)
Double-blinding		
Reported,	23 (28)	4 (4)
Adequately	15 (18)	2 (2)
Inadequately	1 (1)	0 (0)
Not reported	60 (72)	85 (96)
Withdrawals /dropouts		
Reported	29 (35)	39 (44)
Not reported	54 (65)	50 (56)
Allocation concealment		
Reported	11 (13)	8 (9)
Not reported	72 (87)	81 (91)

Double–blinding

Double-blinding was reported in 28% of the RCT-Ds and in 4% of the RCT-Ms. An appropriate method of double blinding was found in 18% of the RCT-Ds and in 2% of the RCT-Ms. Inadequate double blinding was reported in 1% of the RCT-Ds and in 0% of the RCT-Ms (Table 1).

Withdrawals and drop-outs

A complete reporting of withdrawals and drop-outs was given in 35% of the RCT-Ds and in 44% of the RCT-Ms (Table 1).

Allocation concealment

Allocation concealment was reported in 13% of the RCT-Ds and in 9% of the RCT-Ms (Table 1).

Quality scores

High quality scores (3–5 points on the Jadad scale) were assigned for 27% of the RCT-Ds and for 18% of the RCT-Ms. The difference between RCT-Ds and RCT-Ms was not significant (Table 2).

The median Jadad quality score was 2.0 for RCT-Ds and for RCT-Ms. There were no significant differences between the RCT-D and RCT-M scores, indicating equal quality (Table 3).

Journal impact factor

There was no correlation between the Jadad quality scores and the journal impact factor for RCT-Ds (r_s=0.140) or for RCT-Ms (r_s=0.271).

Table 2. Frequency distribution of randomised controlled trials in dental (RCT–Ds) and medical research (RCT–Ms) per Jadad quality score (0–5 points), in 1999

	RCT–Ds			RCT–Ms		
Jadad Score	Number RCT–Ds	Cumulative Number	Cumulative (%)	Number RCT–Ms	Cumulative Number	Cumulative (%)
0	1	1	1.2	4	4	4.5
1	32	33	39.8	33	37	41.6
2	28	61	73.5	36	73	82.0
3	12	73	88.0	15	88	98.9
4	8	81	97.6	1	89	100.0
5	2	83	100.0	0	89	100.0

Jadad scores: 0–2 points indicate poor quality and 3–5 points indicate high quality. There were no significant differences between the number of RCT–Ds and RCT–Ms with poor and high quality (Mann-Whitney U, p=0.179). The frequency distribution of Jadad scores was positively skewed.

Table 3. Total Jadad scores for quality reporting of randomised controlled trials in dental (RCT–Ds) and medical research (RCT–Ms)

	RCT–Ds (n=83)				RCT–Ms (n=89)		
	N	Median	1st quart.	3rd quart.	N	Median	1st quart.	3rd quart.
Jadad Score	83	2.0	1.0	3.0	89	2.0	1.0	2.0

There was no detectable difference between RCT–Ds and RCT–Ms (Mann-Whitney U, p=0.247).

Source of funding

There were no significant differences in the Jadad quality scores between RCTs that were 'research or community funded' (median Jadad score 2.0, n=63) or 'industry funded' (median score 2.0, n=28). The remaining RCTs were funded by a combination of 'research or community funds' and 'industry' (median score 2.0, n=17), or the source of funding was not stated (median score 2.0, n=64).

Area of dental research

The quality scores for RCT–Ds differed between the areas of dental research. RCT–Ds about oral surgery had significantly higher median quality scores, compared with all other areas grouped together. In addition, oral surgery had significantly higher quality scores than periodontics and restorative treatment. No other significant differences were found. The number of RCT–Ds in many areas was inadequate for subgroup analyses (Table 4.)

Table 4. The median Jadad scores for quality of randomised controlled trials (RCT–Ds) in different areas of dental research

Main area covered by the RCT–Ds	Number RCT–Ds	Median score	Range
Endodontics	3	2.0	1-2
Local anesthetics	6	1.5	1-3
Oral surgery*	17	3.0	1-5
Orthodontics	6	1.0	1-3
Periodontics†	20	2.0	1-5
Prosthodontics	4	1.5	1-3
Restorative treatment‡	11	1.0	1-2
Others	16	2.0	0-4
Total	83	2.0	0-5

*Significantly higher scores compared to all other areas together (p=0.001).
† Significant difference to scores of oral surgery (p=0.028) and ‡ (p=0.001), respectively, (Mann Whitney U). 'Others' covered; anxiety control (n=1), apnoea treatment (n=1), caries prevention (n=2), mercury (n=1), oral microbiology (n=1), patient records (n=1), radiology (n=1), sedation (n=2), sedation and local anesthetics (n=1), stain removal (n=1), stomathognatic physiology (n=2), tooth bleaching (n=1), traumatology (n=1).

DISCUSSION

The median quality scores of RCT–Ds and RCT–Ms published in 1999 indicated an overall inadequate quality. The quality was equal for RCT–Ds and RCT–Ms.

The random sample of RCTs in this study contributed 46% of the total number of RCT–Ds and RCT–Ms published in 1999 that were available on Medline. The power calculations indicated adequate sample size for analyses of equivalence in the quality of RCT-D and RCT-M reports.

The quality of an RCT is dependent on all aspects of the study design and trial conduct. A few key features have been shown to have a discriminating effect in assessments of the scientific quality of a trial report.[6,15] The internal validity of an RCT is strongly related to reporting of adequate methodology for random allocation, double-blinding, patient follow-up and allocation concealment.[1,15] It has been shown that trials with poor or inadequately reported methodology tend to exaggerate the treatment effects.[9,16] Empirical evidence supports the view that inadequate random allocation leads to systematic errors in estimates of intervention effects due to selection bias.[15,16] So far, proper randomisation is the only known way to eliminate selection bias from the trial.[15] In line with Dickinson et al[17] we found that less than a third of the RCTs reported an adequate method of randomisation.

Double-blinding is related to ascertainment bias and when lacking, is associated with overestimation of treatment effects.[15,16] Our results indicated that 28% of the RCT–Ds and 4% of the RCT-Ms were double-blinded. This record may prove difficult to improve, since in many trials double-blinding is difficult to achieve. Moreover, the smaller proportion of double-blinded RCT-Ms in comparison to RCT–Ds may be a reflection of differences between these disciplines.

We found complete reporting of withdrawals and drop-outs in about four out of ten RCTs. Reporting of withdrawals and drop-outs is mainly related to the number and the reasons for discontinuing the trial,[6] concerning the possible relation of drop-outs to the intervention itself. Ideally all subjects that were incorporated to the trial should be analysed with intention to treat.[2]

Trials without allocation concealment tend to overestimate the treatment effects.[15] We found appropriately reported allocation concealment in about one out of ten RCTs. This is a low proportion compared with earlier findings, where about one third of the RCT-Ms that had been included in various meta-analyses were adequately concealed.[15] This difference is largely explained by the fact that RCTs included in meta-analyses have already passed a rigorous quality control, selecting high quality RCTs.

The journal impact factor is often viewed as a measure of scientific quality,[18,19] which is why we chose to analyse its relation to the quality scores of RCTs. Our findings indicated no correlation between the journal impact factor and the quality of RCTs. The journal impact factor represents a quota of the number of citations of a journal and the number of citable items during the previous two years.[18,19] Thus, journals with many review articles or methodology papers have relatively high impact factors and journals in smaller or less active research areas have relatively low impact factors. It should nevertheless be recognised that many journals with high impact factors have a rigorous review process before accepting an article. Although it has been shown that industrial funding may influence the outcomes of a trial,[20] we found no differences in the quality of reporting between industry funded and research or community funded RCTs.

Differences in the quality of RCT–Ds were found between different areas of dental research. These results should, however, be regarded with caution as the variations were wide and the sample size was calculated with respect to the primary outcome measures. Thus, the number of RCT–Ds in each area was too small for meaningful subgroup analyses.

CONCLUSION

Considering our findings, the quality of RCTs was generally inadequate. The journal impact factor may not be regarded as a measure of the quality of RCTs. Furthermore, it seems that the quality of RCTs is not influenced by the source of funding. Differences in the quality of RCT-Ds between different areas of dental research seem to exist, but larger samples are needed in order to draw meaningful conclusions. There is a possibility that a large part of the RCTs with poor quality scores are, in fact, high quality trials. However, it was not possible to evaluate this, due to a lack of information in the trial reports. The importance of adequate reporting of the methodology in the RCT conduct cannot be sufficiently emphasised and the quality of RCTs needs to be improved if we are to practice evidence-based healthcare.

1. Moher D, Jadad A R, Nichol G, Penman M, Tugwell P, Walsh S. Assessing the quality of randomised controlled trials: An annotated bibliography of scales and checklists. *Control Clin Trials* 1995; **16**: 62–73.
2. Jadad A R. *Randomised controlled trials.* pp 28-36. London: BMJ Books, 1998.
3. Sacket D L, Rosenberg W M C, Gray J A M, Haynes R B, Richardsson W S. Evidence based medicine: what it is and what it isn't. *Br Med J* 1996; **312**: 71–72.
4. Begg C, Cho M, Eastwood S, *et al.* Improving the quality of reporting of randomised controlled trials: the CONSORT statement. *J Am Med Assoc* 1996; **276**: 637–639.
5. Altman D G. Better reporting of controlled trials: the CONSORT statement. *Br Med J* 1996; **313**: 570–571.
6. Jadad A R, Moore A, Carrol D, *et al.* Assessing the quality of reports of randomised clinical trials: Is blinding necessary? *Control Clin Trials* 1996; **17**: 1–12.
7. Linde K, Clausius N, Ramirez G, *et al.* Are the clinical effects of homeopathy placebo effects? A meta-analysis of placebo-controlled trials. *Lancet* 1997; **350**: 834–843.
8. Moher D, Pham B, Jones A, *et al.* Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses. *Lancet* 1998; **352**: 609–613.
9. Khan K S, Daya S, Jadad A R. The importance of quality of primary studies in producing unbiased systematic reviews. *Arch Intern Med* 1996; **156**: 661–666.
10. Moher D, Fortin P, Jadad A R, *et al.* Completeness of reporting trials published in languages other than English: implications for conduct and reporting of systematic reviews. *Lancet* 1996; **347**: 363–366.
11. McQuay H, Carrol D, Jadad A R, Wiffen P, Moore A. Anticonvulsant drugs for management of pain: a systematic review. *Br Med J* 1995; **311**: 1047–1052.
12. Egger M, Zellweger-Zähner T, Schneider M, Junker C, Lengeler C, Antes G. Language bias in randomised controlled trials published in English and German. *Lancet* 1997; **350**: 326–329.
13. Antczak A A, Tang J, Chalmers T C. Quality assessment of randomised control trials in dental research I. Methods. *J Periodontol Res* 1986; **21**: 305–314.
14. Antczak A A, Tang J, Chalmers T C. Quality assessment of randomised control trials in dental research II. Results. *J Periodontol Res* 1986; **21**: 315–321.
15. Schultz K F, Chalmers A, Haynes R J, Altman D G. Empirical evidence of bias. *J Am Med Assoc* 1995; **273**: 408–412.
16. Chalmers T C, Celano P, Sacks H S, Smith H. Bias in treatment assignment in controlled clinical trials. *New Engl J Med* 1983; **309**: 1358–1361.
17. Dickinson K, Bunn F, Wentz R, Edwards P, Roberts I. Size and quality of randomised controlled trials in head injury: review of published studies. *Br Med J* 2000; **320**: 1308–1311.
18. Garfield E. Citation analyses as a tool in journal evaluation. *Sci* 1972; **178**: 471–479.
19. Garfield E. Significant journals of science. *Nature* 1976; **264**: 609–615.
20. Cho M K, Bero L A. The quality of drug studies published in symposium proceedings *Ann Intern Med* 1996; **124**: 485–489.