

ABSTRACT

Objective: To evaluate the methodological quality of published randomized controlled trials (RCT) in burn care treatment and management.

Methods: Using a predetermined search strategy we searched Ovid MEDLINE (1950 to January 2008) database to identify all English RCT related to burn care. Full text studies identified were reviewed for key demographic and methodological characteristic. Methodological trial quality was assessed using the Jadad Scale.

Results: A total of 257 studies involving 14535 patients met inclusion criteria. The median Jadad score was 2 (out of a possible score of 5). Information was given in the introduction and discussion sections of most RCT, although insufficient detail was provided on randomization, allocation concealment, and blinding. The number of RCT increased between 1950 and 2008 (Spearman's $\rho = 0.6129$, $P < 0.001$), although the reporting quality did not improve over the same time period ($P = 0.1896$), and was better in RCT with larger sample sizes (median Jadad score, 4 vs. 2 points, $P < 0.0001$). Methodological quality did not correlate with journal impact factor ($P = 0.2371$).

Conclusions: The reporting standards of RCT is highly variable and less than optimal in most cases. The advent of evidence-based medicine heralds a new approach to burns care and systematic steps are needed to improve the quality of RCT in this field. Identifying and reviewing the existing number of RCT not only highlights the need for burn clinicians to conduct more trials, but may also encourage burn health clinicians to consider the importance of conducting trials that follow appropriate, evidence-based standards.

BACKGROUND

A well-designed and properly executed randomized controlled trial (RCT) is considered the gold standard for assessing the relative effectiveness of various treatment regimes in clinical research and practice [1, 2].

Due to their ability to eliminate spurious causality and bias, the results of RCT are considered the most reliable form of scientific evidence [2-4]. However, the methodological quality of a specific RCT may be compromised by weaknesses in one or more of the design elements. It is difficult for consumers of research to judge the methodological quality of research trials without adequate reporting of design features such as randomisation techniques, description of outcome assessments or handling of drop-outs and losses to follow-up. Hence the completeness of methodological reporting is critical in assisting readers to judge whether or not the results of a trial justify a change in clinical practice [2, 5].

In recent years, more than 25 scales have been developed to evaluate trial quality [6]. The Jadad scale [2, 5, 7] is one such scale that emphasises the quality of reporting on the description and appropriate use of randomisation, masking (or blinding) and patient attrition [8].

Although several studies have evaluated the quality of RCT using checklists and scales in areas of colorectal, urological and plastic surgery [7, 9, 10, 11], there have been no published reports on the quality of RCT in the field of burn care. More so, no attempt has been made to quantify the proportion of RCT in the adult or paediatrics burns literature.

The reason for this noted gap are various, but have been related to the much broader issues of research feasibility, acceptability, methodology, and ethics in the acute care surgical setting. In particular, burns research has been hampered by a series of challenges in terms of trial design: the heterogeneity of the patient population due to various burns sizes and depths; the volume and variety of numerous burn care interventions and associated confounding variables; the timing and selection of various outcome measures; and the resource constraints of undertaking clinical research [12, 13].

Therefore, the aim of this study was to appraise and evaluate the quality of reporting of existing RCT using the Jadad scale in the field of burn care according to several variables including focus of care (resuscitation, usual burn care, rehabilitation), intervention types (e.g. resuscitation, nutrition, wound, surgery) and general characteristics (e.g. calculation of sample size, funding source, and country of origin).

METHODS

Literature search strategy

A pre-defined search strategy was designed by the authors in conjunction with The Cochrane Collaboration's highly sensitive search strategy to identify RCT in the field of burn care. The RCT search strategy filter was combined with key words and Medical Subject Headings (MeSH) related to burns and thermal injury (Figure 1). The database searched was Ovid MEDLINE from 1950 to January 2008.

INSERT FIGURE ONE

Study selection criteria

Included studies were RCT using human subjects that were allocated at random either to an intervention group undergoing a particular procedure, whether for diagnostic, preventive, or therapeutic purposes or to a control group, which reported on clinical or laboratory outcomes of burn care interventions. We excluded pseudo- or non randomized approaches (i.e. randomizing by odd or even date of birth, rules based on date (or day) of admission, or hospital or clinical record number).

Data extraction

Four authors independently reviewed all RCT and extracted information related to specific design characteristics and analytical approaches. Key design characteristics included sample size, participant age and gender; types of outcomes considered; presence of parallel groups and the number of intervention groups. In order to evaluate the analytical approaches, we observed for *a priori* outcomes, sample size calculations; and whether randomisation was stated. The journal's impact factor -viewed as a measure of scientific quality - was also reported.

Key focus interventions were arbitrarily selected and included interventions relevant to resuscitation therapies, standard burn care practices, rehabilitation techniques including sequelae management.

Intervention types were subdivided into the following: wound care (standard dressing practices of the burn wound), analgesia (use of pharmacological and non-pharmacological agents to minimise the impact of pain in any phase of care), steroid and metabolic therapies (use of pharmacological agents to help reverse the catabolic burn state), nutrition (enteral or parenteral feeding), infection control (regimes to prevent or treat infections), surgery (comparison of different surgical techniques or bioengineered tissues or specific medical versus surgical practices), critical and

intensive care (use of advanced technologies such as mechanical ventilation, invasive haemodynamic monitoring), physical therapy (activities comparing physical or occupational therapies), resuscitation (use of fluid therapy at any phase of care), scar management (measures designated to improve the cosmetic and functional appearance of the scar) and patient education (interventions comparing educational protocols regarding burn care treatments).

Assessment of trial and methodological quality

Studies were objectively assessed using the previously validated Jadad score, which scores RCT on a 5-point scale according to performance in three categories: randomization, blinding and handling of losses to follow-up (both withdrawals and drop-outs). Each question entailed a yes or no response option. Total scores range from 0 to 5 points with a score of 3 or more indicating superior quality [8]. In particular, each item is given one point, if described, or zero points, if not described. For the first two items (randomization and double blinding), if these methods are described and are appropriate, one additional point is given for each item. Conversely, if the methods used to generate the randomization sequence or create blinded conditions are described but inappropriate, the relevant item is given zero points. A trial could be judged as having poor quality if it is awarded two points or less. All studies were independently assessed by four reviewers (SD, CP, SS and CA) using a standardised form, with disagreements resolved by a fifth reviewer by discussion until consensus was reached. To evaluate the reliability of reviewer assessment, an independent reviewer was asked to assess a random sample of 15 studies, using the same evaluation form. The degree of agreement was evaluated using the Cohen's kappa statistic (κ statistic).

Data analysis

All data were managed and analysed using Microsoft Excel and STATA 9.2 (Statacorp Corporation, TX, USA) software respectively. Initial presentation of data was descriptive, which included the median (interquartile range [IQR]) Jadad scores, percentages of the RCT with adequate reporting and the specific characteristics of the studies. Statistical associations between study characteristics and Jadad scores were examined using multiple tests with multivariate logistic modeling employed to determine if any of the study characteristics were capable of predicting a Jadad score greater than 3. An alpha value of 0.05 was considered statistically significant for all tests and confidence intervals (CI) were reported at the 95% level.

RESULTS

Over 7,000 references in Ovid MEDLINE (1950 through to January 2008.) were identified by the search strategy. Independent scrutiny of the titles and abstracts identified 302 potentially relevant articles of which 257 satisfied the inclusion criteria. The remaining 45 studies were

excluded for the following reasons: 24 failed to meet the methodological criteria for an RCT, 15 could not be retrieved in full-text, 4 were performed in non- burned patients, 1 was an animal study and 1 did not clearly state any identifiable intervention.

Study design characteristics

The study characteristics of the included RCT are shown in Table 1. Randomized controlled trials were conducted in 31 different countries, although more than half came from the United States of America (146, 56.8%) followed by United Kingdom (14, 5.5%) and Canada (13, 5.1%). In regards to the studies, 148 (57.6%) included adult patients only, 71 (27.6%) included paediatric patients only and 38 (14.8%) studies included mixed adult and paediatric populations. The mean total burn surface (%TBSA) was 34.9% (range %TBSA: 0.3 to 99).

The most common study design used parallel groups (226, 87.9%) with the remaining studies using a cross-over feature. Among the parallel studies, 195 (86.3%) had two intervention arms while the remaining had three or more intervention arms. Wound care was by far the most prevalent topic of RCT cited (77, 30.0%), followed by analgesia (39, 15.2%), steroid therapies (32, 12.5%), nutrition (26, 10.1%) and infection control (23, 9.0%).

INSERT TABLE ONE

Jadad Scale analysis

The median overall Jadad score was 2 (IQR: 1-3). Only 10 (3.9%) of the included RCT had a Jadad score of 5, whilst 45 (17.51%) studies scored the standard 3 points, and 166 studies (64.6%) scored less than 3 points. Eighteen RCT (7.0%) scored zero points on the Jadad score. For a breakdown of individual Jadad scale scores, see Table 2. The Kappa statistic (κ statistic) used to rate observer inter-reliability showed near perfect agreement beyond chance on overall Jadad scores ($\kappa = 0.90$, 95% CI 0.62-1.00), for randomization $\kappa = 1.00$ (95% CI 0.61 – 1.00), for blinding $\kappa = 0.89$ (95% CI 0.52 – 1.00) and withdrawals $\kappa = 0.81$ (95% CI 0.30 – 1.00).

Although there was an increase in the number of published RCT from 1975 (Spearman's $\rho = 0.6129$, $P < 0.001$, see Figure 2), Jadad scores did not increase over time ($P = 0.1896$) nor did they differ significantly from year to year ($P = 0.0911$); before and after the development and uptake of the Consolidated Standards for Reporting Trials (CONSORT) statement in 1996 ($P = 0.26$, Mann-Whitney). However, the re-development of the CONSORT statement in 2001 showed significant favourable changes in reporting when mean Jadad scores were compared in 1996 to 2001 (1.84

vs. 2.24 , $P= 0.03$, Mann Whitney). There was no association between the journal's impact factor and the Jadad score (Spearman's $\rho=0.0764$, $P=0.2371$).

INSERT FIGURE TWO

Although there were no associations or significant differences in Jadad scores by country of origin ($P=0.06$, Kruskal-Wallis), continent ($P=0.07$, Kruskal-Wallis), funding source ($P=0.09$, Kruskal-Wallis), study design ($P=0.65$, Wilcoxon), intervention type ($P=0.7001$, Kruskal-Wallis) or number of study arms (Spearman's $\rho=0.0212$, $P=0.74$), there was a positive correlation between sample size and Jadad scores (Spearman's $\rho=0.2089$, $P< 0.001$). More so, studies that stated the method for calculating the sample size (15, 5.84%) had better Jadad scores than studies not calculating the sample size (median Jadad score: 4 vs. 2, $P<0.001$, Wilcoxon).

INSERT TABLE TWO

Methodological quality analysis

Adequate randomisation methods and allocation concealment were explicitly cited in only 24 (9.3%) RCT published after the development and uptake of the CONSORT statement in 1996. Double blinding was cited in 36 studies (14.0 %). A total of 172 studies stated their main outcome, among them 131 (76.2%) were clinically specific and 41 (23.8%) biochemical in nature. Jadad scores did not differ ($P=0.18$, Wilcoxon) between studies that stated the main outcome and those that did not. However clinical specific outcomes were associated with higher Jadad scores than biochemical outcomes (median Jadad scores 2 vs. 1, $P=0.04$, Wilcoxon).

A total of 102 (39.7%) studies rejected the null hypothesis of the main outcome. There were no significant differences in Jadad scores between studies rejecting or not rejecting the null hypothesis ($P=0.97$, Wilcoxon). The source of RCT funding was also not associated with the rejection of the null hypothesis ($p=0.3$, Fisher). However, studies that rejected the secondary outcome had worse median Jadad scores than their counterpart (IQR 1-2 vs 1-3, $P=0.03$, Wilcoxon), with stratified analysis showing that this was independent of the rejection of the main hypothesis.

Multivariate logistic modelling was conducted to determine which of the methodological variables had more weight predicting a superior quality study (measured as a Jadad score greater than 3). The results indicated that studies with more than 100 patients (OR 3.61, 95%CI 1.60 to 8.16, $P=0.002$) or RCT stating the method of calculation for a sample size (OR 7.79, 95%CI 2.3 to 25.88,) had higher Jadad scores than those who did not.

DISCUSSION

This systematic analysis of the reporting quality of RCT in burn care research has revealed that the reporting standards in many publications were less than optimal and highly variable in providing information on the design, execution, and interpretation of the study. One in ten RCT scored adequately on the Jadad Scale for reporting. Since we only included RCT published in peer-review journals, it is likely that quality of all RCT in burn care published elsewhere would be of a similar or lesser standard.

As researchers and methodologists, we chose the Jadad Scale because of its use in multiple settings, and more so, it was known to be useful when providing investigators with a numeric indicator of quality of reporting trends over time [5]. However, we were cognizant that the use of this scale did not always reflect the actual methodological quality of a trial [13, 14]. Overall, there was no difference in Jadad scores based on trial design, focus or intervention type. Only sample size and the method for calculating the sample size were factors found to be significantly associated with superior reporting.

As a matter of design and reporting, our results indicated that the majority of studies (84.%) did not state the method of randomization (i.e. reporting of the sequence generation) and the concealment of allocation was adequately explained in only 9.3% of the studies. We did not attempt to contact investigators for precise detail on randomization or allocation concealment, which in turn could have clarified if the RCT were either well designed and poorly reported or just simply inadequate in order to conceal a methodological deficiency. In the latter scenario, it is anticipated that the RCT with poor or unexplained concealment were more likely to exaggerate the treatment effects [15, 16].

This study also found that adequate blinding of the investigator or assessor was reported in only 36 (14.0 %) studies. This poor number may have been secondary to the ethical and logistical problems of randomizing those with burn injury to a specific therapeutic intervention. However, this hurdle, according to Balasubramanian et al 2006^[11], should not have stopped any RCT from receiving a higher Jadad score if they included randomization, allocation concealment, and a detailed account of withdrawals and dropouts. Alternatively, these findings may have also suggested that there was simply an absence in reporting by those undertaking RCT following the release of the CONSORT statement in 1996.

Sample size estimates and calculations may have increased the probability of detecting a clinically significant difference [7], although it was reported in only 15 RCT (5.8%). More so, our results also revealed that the single most important variable predicting a good quality trial was the statement of sample size calculation method (i.e. if a study reported the sample size calculation method, it had a 97.8% specificity of being a good quality trial, regardless of the sample size number. The low sample size numbers also suggested that the remainder of the studies may have been underpowered and prone to Type II errors. However, it was more likely that the variability and impact of burn injuries (i.e. injuries from the very minor, when no or self treatment is sufficient, through to the most severe requiring the highest levels of intensive care and surgery) was a factor in the recruitment of study participants and overall sample size estimates. More so, it may have been suggestive that smaller sample sizes were not necessarily of poorer quality as other aspects of methodological quality may also have been achieved, but not reported [17].

It has been documented that pharmaceutical sponsored trials are prone to bias on methodology and reporting. In our sample, we found that most trials (56.4%) did not report the funding source. Interestingly, studies sponsored by pharmacological industry did not differ on quality or 'positive' findings from other studies. This could be true in nature or biased by the small number of studies sponsored by industry and the large number of studies not reporting the funding source.

Along with other serial accounts of methodological reporting, we had a number of limitations that affected the study. Although the current study had multiple investigators, including qualified information search specialists for the search strategy, a research methodologist to detail the Jadad Scale and burn care experts, we were restricted to one medical database which may not have been representative of all indexed RCT. Other medical databases, burn specific textbooks, conference proceedings, national registries, and non-published RCT were not systematically searched.

Secondly, we were presented with a common challenge that faces bibliometric researchers of separating methodological assessment with methodological reporting [16]. In this instance, there was a potential that the design of our study was an assessment of study methodology using a quality reporting tool rather than a pure methodological assessment of the burns literature. As a result of this approach, there was a potential lapse in the appreciation and recording of the relevance and impact of the RCT because of unsatisfactory reporting and design.

Thirdly, several criteria in the Jadad scale such as "double blinding" were prone to misinterpretation because of the variations in validity and definition of this term by burn care experts and textbooks. No studies showed that the term "double blind" was used with any

statement about who was actually blinded. This highlighted the need to have explicit statements about the blinding status of specific groups involved in RCT development rather than being stuck with the current ambiguous terminology.

Although there appears to be a number of concerns with the burns RCT literature, it is important to acknowledge that RCT may not be able to answer all clinical questions, and the uptake of alternative study designs that are “self matched” in nature e.g. parallel or cross-over designs, maybe the contributing factor to a stronger evidence-base from which clinical practice guidelines or something similar can be derived. More importantly, it is encouraging to see the output of burn RCT literature steadily increasing over time. Although there is ongoing debate and discussion on the merits of evidence-based surgery and the need for surgically-based randomized studies, solving these challenges can only happen when, and if, training of clinicians about the importance and benefit of clinical research can take place.

CONCLUSION

Most randomized controlled trials are of less than optimal and highly variable quality. The advent of evidence-based medicine heralds a new approach to burns care. Systematic steps are needed to improve the quality of RCT in the field of surgery and in particular, in the field of burns care. Identifying the number of RCT can only highlight the need for burn clinicians to conduct more trials, but also encourage burn health clinicians to consider the importance of conducting trials that follow appropriate, evidence-based standards.

REFERENCES

1. Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D, Gøtzsche PC, Lang T; CONSORT GROUP (Consolidated Standards of Reporting Trials). The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med.* 2001 Apr 17; 134(8):663-94.
2. Gummesson C, Atroshi I, Ekdahl C. The quality of reporting and outcome measures in randomized clinical trials related to upper-extremity disorders. *J Hand Surg [Am].* 2004 Jul;29(4):727-34; discussion 735-7.
3. Lachin JM, Matts JP, Wei LJ. Randomization in Clinical Trials: Conclusions and Recommendations. *Controlled Clinical Trials*, 1988; 9(4): 365-74.
4. Sacket D L, Rosenberg W M C, Gray J A M, Haynes R B, Richardson W S. Evidence based medicine: what it is and what it isn't. *Br Med J* 1996; 312: 71-72
5. Latronico N, Botteri M, Minelli C, Zanotti C, Bertolini G, Candiani A. Quality of reporting of randomised controlled trials in the intensive care literature. A systematic analysis of papers

- published in *Intensive Care Medicine* over 26 years. *Intensive Care Med.* 2002 Sep;28(9):1316-23. Epub 2002 Jun 13
6. 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 Clinical Trials* 1995; 16: 62-73.
 7. Karri V. Randomised clinical trials in plastic surgery: survey of output and quality of reporting. *J Plast Reconstr Aesthet Surg.* 2006;59(8):787-96. Epub 2006 Feb 21.
 8. Jadad A, Moore D, Carroll D et al., Assessing the quality of reports of randomized clinical trials: Is blinding necessary. *Control Clinical Trials* 1996; 17: 1–12.
 9. Schwenk W, Haase O, Günther N, Neudecker J. Methodological quality of randomised controlled trials comparing short-term results of laparoscopic and conventional colorectal resection. *Int J Colorectal Dis.* 2007 Nov;22(11):1369-76. Epub 2007 May 26.
 10. Welk B, Afshar K, MacNeily AE. Randomized controlled trials in pediatric urology: room for improvement. *J Urol.* 2006 Jul;176(1):306-9; discussion 309-10.
 11. Balasubramanian SP, Wiener M, Alshameeri Z, Tiruvoipati R, Elbourne D, Reed MW. Standards of reporting of randomized controlled trials in general surgery: can we do better? *Ann Surg.* 2006 Nov;244(5):663-7.
 12. Wasiaik J, Cleland H. How relevant and useful is the Cochrane Library as a resource for evidence in burn care, practice, and management? *J Burn Care Res.* 2006 May-Jun;27(3):386-93.
 13. Potokar TS, Prowse S, Whitaker IS, Ali S, Chamania S. A global overview of burns research highlights the need for forming networks with the developing world. *Burns.* 2008 Feb;34(1):3-5. Epub 2007 Dec 26.
 14. Huwiler-Müntener K, Juni P, Junker C, Egger M. Quality of reporting of randomized trials as a measure of methodologic quality. *JAMA.* 2002 Jun 5;287(21):2801-4.
 15. Juni P, Altman DG, Egger M. Assessing the quality of randomized controlled trials In: Egger M, Smith GD, Altman DG, editors. *Systematic reviews in health care. Meta-analysis in context.* London: BMJ Publishing Group; 2001:87-109.
 16. Schulz KF. Unbiased research and the human spirit: the challenges of randomized controlled trials. *CMAJ.* 1995 Sep 15;153(6):783-6.
 17. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA.* 1995 Feb 1;273(5):408-12.

Figure 1. Medline search strategy

MEDLINE (Ovid) 1950 to January 2008

1 RANDOMIZED CONTROLLED TRIAL.pt.
2 CONTROLLED CLINICAL TRIAL.pt.
3 RANDOMIZED CONTROLLED TRIALS.sh.
4 RANDOM ALLOCATION.sh.
5 DOUBLE BLIND METHOD.sh.
6 SINGLE-BLIND METHOD.sh.
7 or/1-6
8 ANIMALS.sh. not HUMAN.sh.
9 7 not 8
10 CLINICAL TRIAL.pt.
11 exp CLINICAL TRIALS
12 (clin\$ adj25 trial\$).ti,ab.
13 ((singl\$ or doubl\$ or trebl\$ or tripl\$) adj25 (blind\$ or mask\$)).ti,ab.
14 PLACEBOS.sh.
15 placebo\$.ti,ab.
16 random\$.ti,ab.
17 RESEARCH DESIGN.sh.
18 or/10-17
19 18 not 8
20 19 not 9
21 COMPARATIVE STUDY.sh.
22 exp EVALUATION STUDIES
23 FOLLOW UP STUDIES.sh.
24 PROSPECTIVE STUDIES.sh.
25 (control\$ or prospectiv\$ or volunteer\$).ti,ab.
26 or/21-25
27 26 not 8
28 27 not (9 or 20)
29 9 or 20 or 28
30 exp Burns/
31. burn\$.ti, ab
32. thermal adj\$ injur\$. Ti, ab
33. or/31-33
34. 29 and 33

Legend:

exp = retrieves all items with the subject heading

ti = retrieves items with words contained in the title

ab= retrieves items with words contained in the abstract

tw = looks for the term/s in the abstract, title, MeSH headings and Subheadings

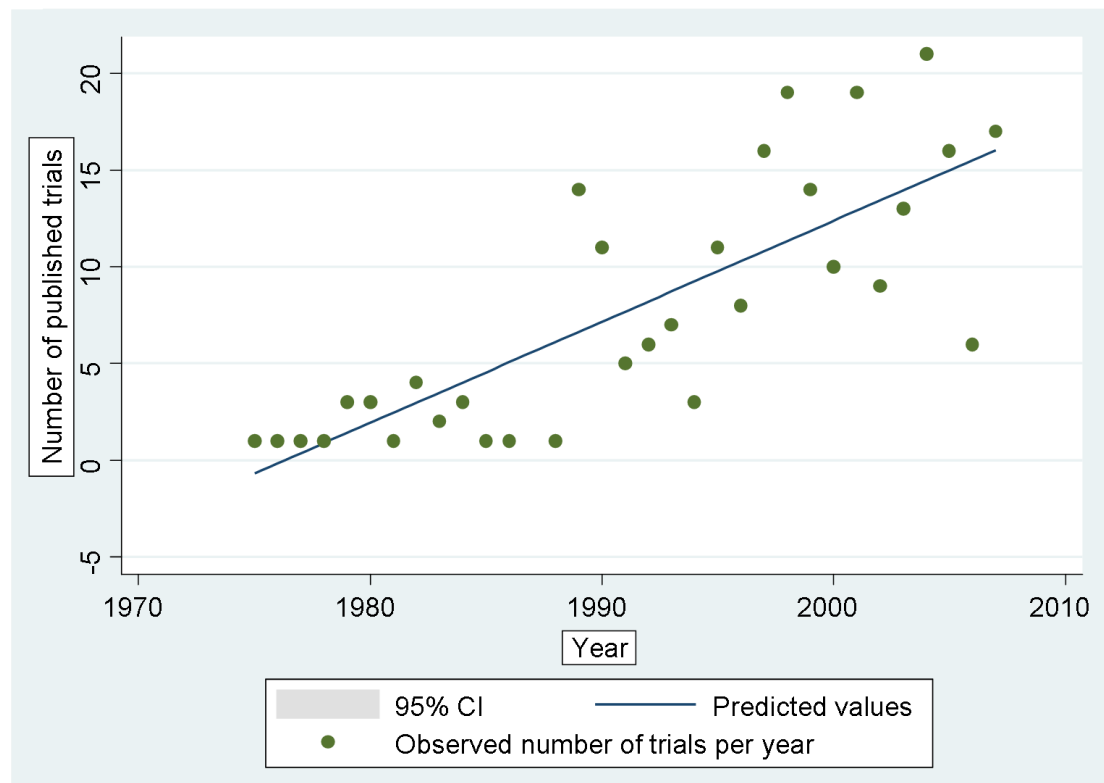
\$ = looks for variations in spelling i.e. pre-operative, preoperative, pre operative

or = retrieves documents that contain at least one of the specified search terms

and = retrieves a set in which each citation contains all the search terms

adj= retrieves items with query terms on either side in the specified order

Figure 2. Number of published RCTs in burn care by year, 1975 - 2008



* Observed values are represented by dots, predicted values by Prais-Winsten regression fit of number of trials per year is displayed with the 95% confidence interval in the ascending line (Coef. 0.5, $p < 0.001$, $R^2 0.6310$).