Intensive Care Med (2002) 28:1316–1323
DOI 10.1007/s00134-002-1339-x

ORIGINAL

Nicola Latronico
Marco Botteri
Cosetta Minelli
Cinzia Zanotti
Guido Bertolini
Andrea Candiani

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

Received: 25 September 2001
Accepted: 1 February 2002
Published online: 13 June 2002

An editorial regarding this article can be found in the same issue (http://dx.doi.org/10.1007/s00134-002-1400-9)

N. Latronico (✉) · M. Botteri · C. Zanotti
A. Candiani
Istituto di Anestesia e Rianimazione,
University of Brescia,
Piazzale Spedali Civili 1, 25125 Brescia,
Italy
e-mail: latronic@master.cci.unibs.it
Tel.: +39-030-3995561
Fax: +39-030-2525 507

C. Minelli · G. Bertolini
Laboratorio di Epidemiologia Clinica,
Istituto di Ricerche Farmacologiche Mario
Negri, Centro per lo Studio delle Malattie
Rare Aldo e Cele Dacco', Villa Camozzi,
Via Camozzi 3, 24020 Ranica (Bergamo),
Italy

Abstract *Objective:* To assess the number and quality of the reporting of randomised controlled trials (RCTs) published in *Intensive Care Medicine*. *Design:* Systematic revision. *Setting:* Randomised controlled trials published in *Intensive Care Medicine*. *Study selection:* All RCTs published in this journal from its birth to December 2000 identified by MEDLINE and our own research. *Measurements and results:* The Jadad scale and the individual assessment of key methodological components, namely the randomisation process, blinding and reporting and handling of loss to follow-up, were used to evaluate the quality of reporting. Other information was extracted regarding the design characteristics and the analytical approach. 173 RCTs, 63% of which were from European countries, were analysed. Adequately reported RCTs according to a Jadad scale score of more than 2 were 44 (25.4%). Analysis of individual methodological components revealed a variable percentage of adequate reporting ranging from 3.5% for randomisation to 10.4% for blinding and to 49.1% for loss to follow-up. Sample sizes were small with a median of 30 patients and rationale for its estimation was reported in 7.5%. Despite this, 81.5% of RCTs reported statistically significant results, suggesting that the treatment effects were strong or that a publication bias existed or that the uncertainty principle was not fulfilled. *Conclusions:* Randomised controlled trials offer the best evidence of the efficacy of medical interventions, provided that high standards of transparent reporting are used. More resolute attention to the methodological quality of reporting and adherence to recently published guidelines (CONSORT II) may help to achieve this result.

Keywords Systematic revision · Randomised controlled trial · Reporting quality · Blinding · Withdrawals · Sample size

Introduction

Evidence-based medicine involves the systematic collection, analysis, synthesis and application of the best clinical evidence to integrate individual clinical expertise [1]. Randomised controlled trials (RCT) produce one of the highest levels of evidence available to evaluate the efficacy of health care interventions [2]. As such, they are usually consulted by clinicians as a guide to their daily practice. The patients and the scientific community can derive maximum benefit from trials only if they are rigorously designed, performed and completely reported. In fact, unclear or incomplete data reporting makes the interpretation of RCTs difficult, if not impossible, and may jeopardise an otherwise well planned and conducted work [3, 4].

Several features are considered relevant to the definition of the quality of RCTs. Proper randomisation is a

crucial component of high quality trials, since it balances the groups for prognostic factors, either known or unknown, provided large samples are used. Furthermore, physicians responsible for entering the patients into the study are unaware of which treatment the next patient will receive ("allocation concealment"), thus eliminating selection bias (provided patients are kept in the assigned group, *see below*). Both the generation of an unpredictable allocation sequence and its concealment until assignment occurs are essential, since inappropriate or unclear randomisation may yield inflated treatment effects [3, 4].

Other aspects which are relevant to the definition of the quality of RCTs include blinding, in order to avoid performance and detection (or ascertainment) bias, and complete reporting of protocol violations and patients lost to follow-up to prevent attrition bias [5]. Among protocol deviations, non-adherence to treatment regimens poses special problems in the analysis. There is agreement that patients should be kept in the group to which they were originally allocated, whether or not they actually received the assigned treatment and even if they received the wrong treatment. This is an essential prerequisite for a study to be valid and is called "intention-to-treat" as opposed to "on treatment" analysis [see 6 for review]. If not clearly stated by the authors, selection bias cannot be excluded, thus nullifying the randomisation process. In recent years, the Consolidated Standards of Reporting Trials (CONSORT) statement has been formulated [7] and implemented [8, 9] to provide a standard checklist and flow diagram for the reporting of RCTs. A revised version has recently been proposed [6].

Attention to achieving the highest standards of transparent reporting of RCTs is important not only for researchers but also for editors, as it is commonly thought that journals publishing high quality RCTs gain in prestige and consideration. The aim of this work was to assess the number and quality of RCTs published in *Intensive Care Medicine* (formerly *European Journal of Intensive Care Medicine*) over 26 years from the first issue (1975) until December 2000.

Methods

Randomised controlled trials were analysed only if they evaluated the efficacy of treatments. RCTs evaluating different diagnostic strategies were excluded. Studies were identified by MEDLINE (limits: "human" and "randomised controlled trial") and our own research. To evaluate the publication trend, the proportion of RCTs to the total number of original articles was calculated for each year [10].

Assessment of methodological quality by means of a summary scale

All RCTs were assessed for quality reporting by a scale developed with an appropriately rigorous standard [11]. This scale contains two questions for randomisation and double-blinding and one question evaluating the reporting of withdrawals and dropouts. Each question entails a yes or no response option. Total scores range from 0 to 5 points (2 points each for randomisation and double-blinding, 1 point for withdrawals), with higher scores indicating superior quality. A summary score greater than 2 defined adequate reporting [12]. Two clinicians assessed the studies using a standardised form and discrepancies were resolved by discussion with a senior clinician until consensus was reached. To evaluate the reliability of the clinicians' assessment, a clinical epidemiologist was asked to assess independently a random sample of 25 studies (about 15% of the total number), using the same evaluation form. The degree of the agreement was evaluated for each item by the overall proportion of agreement (p_o) and the kappa statistic (k). The latter is the correction of the former obtained by subtracting the proportion of agreement attributable to chance [13]. When dealing with more than two ordered categories (e.g., randomisation – 0, 1, 2), a weighted k was also calculated [14]. The following classes of agreement for k and weighted k were adopted: k 0.20 or less = poor; k 0.21–0.40= fair; k 0.41–0.60= moderate; k 0.61–0.80= good; k more than 0.80= very good.

Assessment of methodological quality by means of individual components

Three methodological components were evaluated: the randomisation process, blinding and reporting and handling of loss to follow-up (both dropouts and withdrawals). In general, a "guilty until proved innocent" approach was adopted, and omission of information was equated to inadequate quality [5]. In addition, a clear description of the method used was required to define the adequacy of the assessed component (e.g., random number table or a computerised random number generation of the sequence; sealed, opaque, sequentially numbered or coded envelopes for concealment).

Other information extracted

In addition to evaluating the quality of reporting of RCTs, we also extracted information about design characteristics and analytical approaches [12]. To assess the design characteristics we extracted the sex and age (adult versus others) of participants; type of outcome considered (survival or specific outcomes); presence of parallel groups (yes/no); number of intervention groups (two or more); characteristics of the control group (placebo or active treatment); number of patients included; blinded subjects (patient, clinician, researcher, outcome assessor, statistician). In order to evaluate the analytical approaches we checked if the primary outcome was specified; if an "a priori" calculation of the sample size was performed and the rationale for sample size estimation stated; if simple or restricted randomisation was used; if withdrawals were indicated; if the trial reported statistically significant results or not. Data were analysed using simple descriptive statistics.

Results

We found 181 RCTs published in *Intensive Care Medicine*, 153 (84.5%) identified by electronic search in MEDLINE and 28 (15.5%) by our own research. Volumes of the journal published from 1979 to 2000 were manually checked and selected on the basis of the title. RCTs were identified through reading abstracts. For volumes published from 1975 to 1978, only the titles from MEDLINE were available. Numbers 1 and 3 of Volume

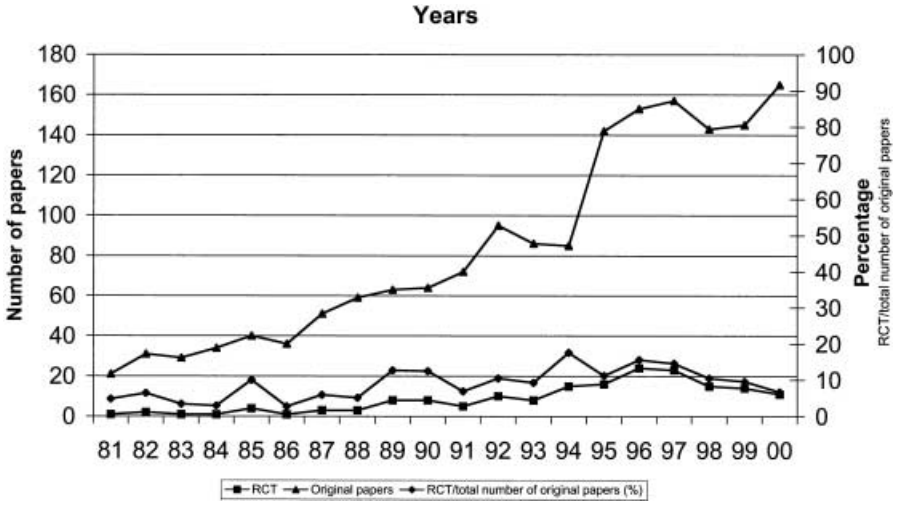

Fig. 1 Publication trend of original papers, randomised controlled trials (RCT) and their proportions

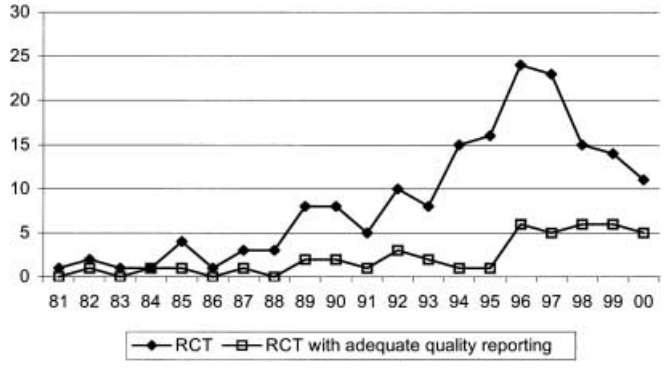

Fig. 2 Total number of randomised controlled trials (RCT) and RCT with adequate quality reporting according to the Jadad scale score. Note that despite the decline of published RCT after 1997, the number of good quality RCTs remained stable

Table 1 Details of quality reporting according to the Jadad scale (the higher the score, the better the quality)

	Score		
Item	0	1	2
Randomisation		127 (73.4%)	46 (26.6%)
Double blinding	131 (75.7%)	24 (13.9%)	18 (10.4%)
Withdrawals	88 (50.9%)	85 (49.1%)	–

1 were impossible to find and no data were available. Eight papers (4.4%), classified as RCTs in MEDLINE, were excluded because they actually were reviews (4), editorials (1), letters to the Editor commenting on other papers (1) or papers dealing with diagnosis (2). This left 173 (95.6%) papers available for the analysis.

The number of original papers and RCTs increased slowly from 1981 (the data of the first RCT published) to 1989 and then increased at a faster rate. Starting from 1995, when the journal changed from 8 to 12 issues per year, the total number of original papers increased steeply, while RCTs increased less until 1997 and then decreased. As a result, the proportion of RCTs compared to the total number of original papers decreased (Fig. 1).

The large majority of RCTs published (112, 64.7%) came from Germany (33, 19.1%), France (31, 17.9%), United Kingdom (17, 9.8%), Spain (11, 6.0%), Switzerland (11, 6.0%) and Italy (9, 5.2%). Researchers were from 27 nations, including other European countries, United States, Oceania, Middle East, Africa, Japan and other Asian countries. One-hundred and fifty-three RCTs included adult patients (88.4%). Both genders were usually studied (145 RCT, 83.9%), but we also found studies including only males (5, 2.9%) or females (2, 1.1%). In five papers (2.9%) age was not indicated and in 21 (12.1%) gender was omitted.

There were 44 (25.4%) adequately reported RCTs according to a Jadad scale score higher than 2. Their number per year remained unchanged until 1996, when quality of reporting increased (chi-square 3.5775, p=0.059) and remained stable despite the declining number of RCTs published (Fig. 2). Details of the Jadad scale score are summarised in Table 1. The degree of agreement between clinicians and epidemiologists in the sample of 25 articles was good for all items, the overall proportion of agreement ranging from 0.84 to 0.92, and the kappa statistic (which was a weighted kappa for all items with more than two ordered categories) ranging from 0.63 to 0.89.

Quality analysis with both the summary Jadad scale score (Fig. 2) and its individual components (Fig. 3) showed improvement over time starting from 1995. Analysis of other components not included in the Jadad scale showed a similar trend (Fig. 3), however frequencies of adequate reporting were low. Allocation concealment and restricted randomisation were explicitly cited only in, respectively, 6 (3.5%) and 4 (2.3%) studies pub-

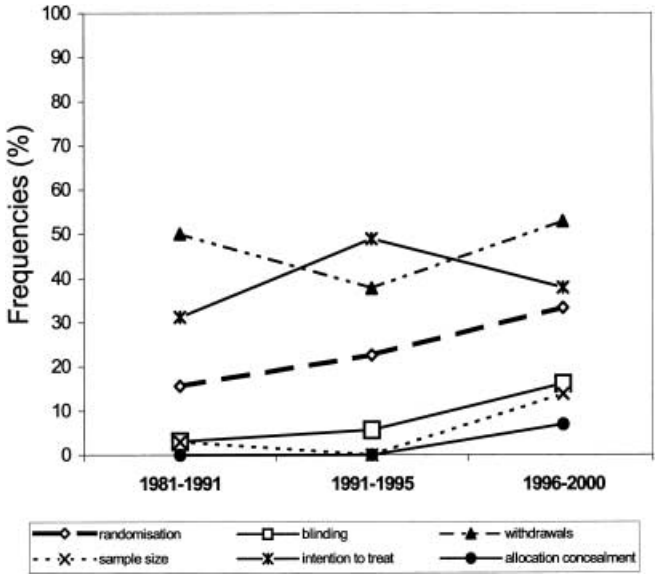

Fig. 3 Trend over time of key methodological components of published randomised controlled trials (RCT). Randomisation, double blinding and withdrawals are the components of the Jadad score. Frequencies of papers with maximum score are shown. Sample size, intention to treat and allocation concealment are not part of the Jadad score. Frequencies of papers with adequate reporting are shown

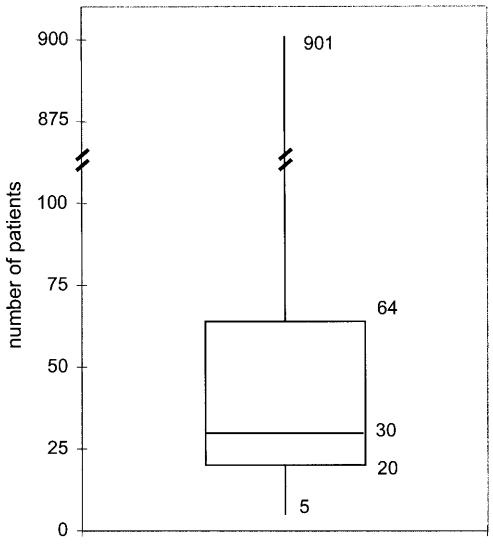

Fig. 4 Sample sizes of randomised controlled trials published in *Intensive Care Medicine* from its birth (1975) to the December 2000 issue. Minimum (5 patients) and maximum (901), and 25th (20 patients), 50th (30 patients, the median) and 75th (64 patients) percentiles are shown. Fifty percent of studies had a sample size between 20 and 64 (*the box*)

Table 2 Design characteristics of the randomised controlled trials

		Number of articles	Percentage
Assessed change in specific outcome	Survival	4	2.3
	Specific outcome	169	97.7
Parallel groups	Yes	135	78
	No	38	22
Number of intervention groups	2 Groups	116	67
	> 2 Groups	57	33
Comparison group receiving active treatment	Yes	117	67.6
	No	56	32.4
Blinding	"Double blinding" no further specified	23	54.8
	Patient and clinician blinded	13	30.9
	Patient, clinician and researcher blinded	4	9.5
	Patient, clinician and statistician blinded	1	2.4
	Patient, clinician and lab technician blinded	1	2.4

lished as from 1993; sample size estimation was stated in 13 (7.5%, all but one published as from 1998). Double blinding was cited in 42 studies (24.3%) and in 19 (11.0%) the blinding status was further specified (Table 2). In 18 (10.4%, all but one published as from 1994) the procedure for blinding was described. Adequate reporting of withdrawals and of intention-to-treat analysis was substantially higher, 85 (49.1%) and 70 studies (40.5%), respectively. Overall, the quality assessed by individual methodological components of studies was worse than that estimated by the Jadad scale.

The design characteristics of trials analysed are summarised in Table 2. The most common design was that of two parallel groups. 38 studies (21.9%) were classified as non-parallel groups, of which 26 (15.0%) were crossover studies. Information regarding the analytical approaches are summarised in Table 3. The median sample size was 30 patients; 89 (51%) papers had 30 patients or less, 48 (28%) 20 patients or less and 16 (9%) 10 patients or less (Fig. 4).

Table 3 Analytical approaches of randomised controlled trials

		Number of articles	Percentage
Primary outcome specified	Yes	173	100
	No	0	0
Rationale for sample size estimation stated	Yes	12	6.9
	No	161	93.1
Withdrawals	Yes	68	39.3
	No	17	9.8
	Not specified	88	50.9
Participants randomised and included in analysis (intention-to-treat)	Yes	69	39.9
	No	29	16.8
	Not specified	75	43.3
Statistically positive result	Yes	141	81.5
	No	29	16.8
	Yes\no	3	1.7

Discussion

Randomised controlled trials provide the best evidence of the efficacy of medical intervention. Although the first example of a study with random allocation was reported in 1884 [15], only in the last 30 years has the number of RCTs increased substantially, particularly in general medical journals [10]. The results of the present study show that a similar trend has taken place in *Intensive Care Medicine*. After the first RCT was published in 1981 the number of RCTs published grew slowly until 1989 and then increased more quickly. However, when the proportion of RCTs to the total number of published original papers was considered, the proportion increased until 1994 and then decreased when the journal changed from 8 to 12 issues yearly. This suggests that other types of study design (that is, observational, rather than experimental, design) are more commonly used, possibly because of the difficulty, either perceived or objective [16, 17], in performing RCTs in this area. However, the evidence that observational studies may give comparable results to RCTs is not convincing as yet [18, 19, 20, 21]. It is also possible that, due to the complexity of performing RCTs and their relevance, they are published more widely in general medical journals. Indeed, in the period 2000–2001, 22 important randomised trials were published, the majority (11, 50%) by European researchers, in the *Annals of Internal Medicine*, *JAMA*, *Lancet* and *New England Journal of Medicine* [22, 23, 24, 25, 26, 27, 2829, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43], indicating that intensive care physicians do produce influential clinical research.

Quality of randomised controlled trials

Analysis of the quality of reporting led to discordant results when this was assessed by means of a summary score, the Jadad scale, and by individual methodological components. In the first case, 25.4% of papers were adequate (score >2), while in the second percentages varied between 3.5% and 49.1% depending on the methodological component examined. Whether one method or the other should be preferred is debatable. The Jadad scale is the only known scale developed with standard scale development techniques and has been used extensively in several clinical areas as it is efficient to use [12]. It also provides investigators with a numeric indicator of quality of reporting that is useful when describing trends over time. The agreement between clinicians' and epidemiologists' evaluation in this study was high and good for all the items investigated, indicating a lower inter-observer variability. On the other hand, the use of composite scales is problematic for several reasons. In particular, there are several scales available and it is unclear which would produce the best assessment. Jüni et al. have recently demonstrated that different scales lead to different interpretation and grading of the quality of studies [44].

As a matter of fact, assessment of key methodological components – proper randomisation, blinding and reporting and handling of protocol violation, dropouts and withdrawals – should always be done when dealing with quality evaluation of RCTs [5, 44]. Our results indicate that the most critical aspects of quality were the reporting of randomisation (particularly the allocation concealment) and blinding, rather than the handling of patients' attrition. This is further confirmed by the fact that only four papers indicated a procedure for restricted randomisation, despite the fact that the majority of studies were dealing with small samples. Unrestricted, simple randomisation in such cases may cause imbalance in baseline prognostic variables, thus altering the trial's results [6].

The validity of the terms "single", "double" and "triple" blind has recently been questioned in that they have become almost a matter of convention [4, 45]. Both physicians and textbooks vary greatly in their interpretations and definitions of these terms. Thus, explicit statements about the blinding status of specific groups involved in

RCTs is suggested to be preferable to the current ambiguous terminology [45]. This is confirmed by our result which shows that, in 23 (55%) of 64 studies, the term "double-blind" was used without any statement about who was actually blinded. When a detailed description was provided, six studies were actually triple-blind: researchers, statisticians or laboratory technicians were blinded in addition to patients and clinicians, important information which is missed if conventional terminology is used.

Interestingly, the number of adequately reported RCTs increased steadily starting from 1996, after the publication of CONSORT [7]. This temporal relationship suggests that guidelines were included in the editorial policy of the journal and/or that they gained acceptance among authors and were partly responsible for the improved quality of reporting. In fact, Moher et al. recently demonstrated a similar improvement in RCTs published in general medical journals since the CONSORT statement [9].

Despite the fact that the quality of reporting of RCTs published in *Intensive Care Medicine* was not particularly encouraging, data from other medical journals or areas are not substantially better. Randomisation is usually clearly stated in articles, because "it is something to be proud of" [1], however the method for generation of the allocation sequence and its concealment rarely is. Actually, allocation concealment was stated in 23% of trials in head trauma [46], 11% of trials in rheumatoid arthritis [47], 52% of trials in obstetrics and gynaecology journals [48] and 56% of trials in general medical journals [49]. Only 5 of 73 (7%) RCTs published in a dermatology journal between 1976 and 1997 reported the method used to allocate treatments [50].

Sample size

Sample size estimation is important when planning RCTs and should be based on a balance of clinical, statistical and economic considerations. The larger the sample size, the higher the probability of detecting a clinically significant difference, but also the higher the costs. Estimation is based on the primary outcome, which was specified in all 173 RCTs evaluated. However, the rationale for sample size estimation was reported in only 12 (6.9%). In addition, sample sizes were small, with some studies including only 5–10 patients. The high number of studies reporting statistically different results (141, 81.5%) suggests that either treatment effects were strong, a publication bias (that is, the selected publication of papers with positive results) existed or the uncertainty principle (according to which a patient should be entered only if the clinician is substantially uncertain which of the trial treatments would be most appropriate [51]) was not fulfilled.

Studies with insufficient power to detect a clinically significant difference often occur in medicine [51]. In a recent systematic review of RCTs in head trauma patients, the average number of randomised participants was 82. None of the trials would have been large enough to detect reliably the difference between a 20% and a 15% risk of death or disability [46]. The median sample size was 54 patients in 196 trials in arthritis [47], 46 patients in 73 trials in dermatology [50] and 65 patients in 2000 trials in schizophrenia [52].

In conclusion, the vast majority of RCTs published in *Intensive Care Medicine* were from European countries, confirming that the journal is an important reference for European intensive care research [53]. Their number has increased and their quality has improved over the 26year period we examined. However, our study suggests that more consideration to the methodological quality, particularly to randomisation, blinding and sample size estimation, could be helpful for the journal's pre-eminence to be maintained in the challenging world of intensive care medicine. Adherence to recently published guidelines (CONSORT II) [6] may help to refine the peer review process for publication of future RCTs.

References

1. Sackett DL, Richardson SW, Rosenberg W, Haynes BR (eds) (1998) Evidence-based medicine. How to practice and teach EBM. Churchill Livingstone, Edinburgh New York Toronto
2. Harbour R, Miller J for the Scottish Intercollegiate Guidelines Network Grading Review Group (2001) A new system for grading recommendations in evidence based guidelines. BMJ 323:334–336
3. Schulz KF, Chalmers I, Hayes RJ, Altman DG (1995) Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. JAMA 273:408–412
4. Moher D, Pham B, Jones A, Cook DJ, Jadad AR, Moher M et al. (1998) Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? Lancet 352:609–613
5. Altman DG (2001) Systematic reviews in health care. Systematic reviews of evaluations of prognostic variables. BMJ 323:224–228
6. Altman DG, Schulz KF, Moher D, Egger M (2001) The revised CONSORT statement for reporting randomized trials: explanation and elaboration. Ann Intern Med 134:663–694
7. Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I et al. (1996) Improving the quality of reporting of randomized controlled trials. The CONSORT statement. JAMA 276:637–639

8. Egger M, Juni P, Bartlett C (2001) Value of flow diagrams in reports of randomized controlled trials. The CONSORT Group. JAMA 285:1996–1999
9. Moher D, Jones A, Lepage L for the CONSORT group (2001) Use of the CONSORT statement and quality of reports of randomized trials. A comparative before-and-after evaluation. JAMA 285:1992–1995
10. Simini B (1998) Randomised controlled trials in *The Lancet*. Lancet 351:682
11. Jadad AR, Moore RA, Carroll D et al. (1996) Assessing the quality of report of randomized clinical trial: is blinding necessary? Control Clin Trials 17:1–12
12. Moher D, Fortin P, Jadad AR, Juni P, Klassen T, Le Lorier J et al. (1996) Completeness of reporting of trials published in languages other than English: implications for conduct and reporting of systematic reviews. Lancet 347:363–366
13. Cohen J (1960) A coefficient of agreement for nominal scales. Educ Psychol Meas 20:37–46
14. Altman D G (1990) Some common problems in medical research. In: Practical statistics for medical research. Chapman & Hall, Boca Raton London New York Washington, pp 403–409
15. Peirce CS, Jastrow J (1884) On small differences of sensation. Memoirs Nat Acad Sci 3:73–83
16. Sutcliffe AJ, Walia S (2001) Normalisation of blood pressure in bleeding patients. Lancet 357:1976
17. Roberts I, Evans P (2001) Normalisation of blood pressure in bleeding patients. Lancet 357:1977
18. Concato J, Shah N, Horwitz RI (2000) Randomized, controlled trials, observational studies and the hierarchy of research designs. N Engl J Med 342:1887–1892
19. Benson K, Hartz AJ (2000) A comparison of observational studies and randomized, controlled trials. N Engl J Med 342:1878–1886
20. Pocock SJ, Elbourne DR (2000) Randomized trials or observational tribulations? N Engl J Med 342:1907–1909
21. Ioannidis JPA, Haidich AB, Pappa M et al. (2001) Comparison of evidence of treatment effects in randomized and nonrandomized studies. JAMA 286:821–830
22. Fagon JY, Chastre J, Wolff M, Gervais C, Parer-Aubas S, Stephan F, Similowski T, Mercat A, Diehl JL, Sollet JP, Tenaillon A (2000) Invasive and noninvasive strategies for management of suspected ventilator-associated pneumonia. A randomized trial. Ann Intern Med 132:621–630
23. Angus DC, Birmingham MC, Balk RA, Scannon PJ, Collins D, Kruse JA, Graham DR, Dedhia HV, Homann S, MacIntyre N (2000) E5 murine monoclonal anti-endotoxin antibody in gram-negative sepsis: a randomized controlled trial. E5 Study Investigators. JAMA 283:1723–1730
24. The ARDS Network Authors for the ARDS Network. (2000) Ketoconazole for early treatment of acute lung injury and acute respiratory distress syndrome A randomized controlled trial. JAMA 283:1995–2002
25. Delclaux C, L'Her E, Alberti C, Mancebo J, Abroug F, Conti G, Guerin C, Schortgen F, Lefort Y, Antonelli M, Lepage E, Lemaire F, Brochard L (2000) Treatment of acute hypoxemic nonhypercapnic respiratory insufficiency with continuous positive airway pressure delivered by a face mask A randomized controlled trial. JAMA 284:2352–2360
26. Antonelli M, Conti G, Bufi M, Costa MG, Lappa A, Rocco M, Gasparetto A, Meduri GU (2000) Noninvasive ventilation for treatment of acute respiratory failure in patients undergoing solid organ transplantation. A randomized trial. JAMA 283:235–241
27. Warren BL, Eid A, Singer P, Pillay SS, Carl P, Novak I, Chalupa P, Atherstone A, Penzes I, Kubler A, Knaub S, Keinecke HO, Heinrichs H, Schindel F, Juers M, Bone RC, Opal SM, Kyber-Sept Trial Study Group (2001) Caring for the critically ill patient. High-dose antithrombin III in severe sepsis: a randomized controlled trial. JAMA 286:1869–1878
28. Merrer J, De Jonghe B, Golliot F, Lefrant JY, Raffy B, Barre E, Rigaud JP, Casciani D, Misset B, Bosquet C, Outin H, Brun-Buisson C, Nitenberg G, French Catheter Study Group in Intensive Care (2001) Complications of femoral and subclavian venous catheterization in critically ill patients: a randomized controlled trial. JAMA 286:700–707
29. Bellomo R, Chapman M, Finfer S, Hickling K, Myburgh J (2000) Low-dose dopamine in patients with early renal dysfunction: a placebo-controlled randomised trial. Australian and New Zealand Intensive Care Society (ANZICS) Clinical Trials Group. Lancet 356:2139–2143
30. Ronco C, Bellomo R, Homel P, Brendolan A, Dan M, Piccinni P, La Greca G (2000) Effects of different doses in continuous veno-venous haemofiltration on outcomes of acute renal failure: a prospective randomised trial. Lancet 355:26–30
31. Masip J, Betbese AJ, Paez J, Vecilla F, Canizares R, Padro J, Paz MA, De Otero J, Ballus J (2000) Non-invasive pressure support ventilation versus conventional oxygen therapy in acute cardiogenic pulmonary oedema: a randomised trial. Lancet 356:2126–2132
32. Schortgen F, Lacherade JC, Bruneel F, Cattaneo I, Hemery F, Lemaire F, Brochard L (2001) Effects of hydroxyethylstarch and gelatin on renal function in severe sepsis: a multicentre randomised study. Lancet 357:911–916
33. Nouira S, Marghli S, Belghith M, Besbes L, Elatrous S, Abroug F (2001) Once daily oral ofloxacin in chronic obstructive pulmonary disease exacerbation requiring mechanical ventilation: a randomised placebo-controlled trial. Lancet 358:2020–2025
34. Bennett CC, Johnson A, Field DJ, Elbourne D, UK Collaborative ECMO Trial Group (2001) UK collaborative randomised trial of neonatal extracorporeal membrane oxygenation: follow-up to age 4 years. Lancet 357:1094–1096
35. Gattinoni L, Tognoni G, Pesenti A, Taccone P, Mascheroni D, Labarta V, Malacrida R, Di Giulio P, Fumagalli R, Pelosi P, Brazzi L, Latini R for the Prone-Supine Study Group (2001) Effect of prone positioning on the survival of patients with acute respiratory failure. New Engl J Med 345:568–573
36. The Acute Respiratory Distress Syndrome Network (2000) Ventilation with lower tidal volumes as compared with traditional tidal volumes for acute lung injury and the acute respiratory distress syndromes. N Engl J Med 342:1301–1308
37. Kress JP, Pohlman AS, O'Connor MF, Hall JB (2000) Daily interruption of sedative infusions in critically ill patients undergoing mechanical ventilation. N Engl J Med 342:1471–1477
38. Hilbert G, Gruson D, Vargas F, Valentino R, Gbikpi-Benissan G, Dupon M, Reiffers J, Cardinaud JP (2001) Noninvasive ventilation in immunosuppressed patients with pulmonary infiltrates, fever, and acute respiratory failure. N Engl J Med 344:481–487
39. Bernard GR, Vincent JL, Laterre PF, LaRosa SP, Dhainaut JF, Lopez-Rodriguez A, Steingrub JS, Garber GE, Helterbrand JD, Ely EW, Fisher CJ Jr (2001) Efficacy and safety of recombinant human activated protein C for severe sepsis. N Engl J Med 344:699–709

40. Van den Berghe G, Wouters P, Weekers F, Verwaest C, Bruyninckx F, Schetz M, Vlasselaers D, Ferdinande P, Lauwers P, Bouillon R (2001) Intensive insulin therapy in the surgical intensive care unit. N Engl J Med 345:1359–1367
41. Rivers E, Nguyen B, Havstad S, Ressler J, Muzzin A, Knoblich B, Peterson E, Tomlanovich M, Early Goal-Directed Therapy Collaborative Group (2001) Early goal-directed therapy in the treatment of severe sepsis and septic shock. N Engl J Med 345:1368–1377
42. Herndon DN, Hart DW, Wolf SE, Chinkes DL, Wolfe RR (2001) Reversal of catabolism by beta-blockade after severe burns. N Engl J Med 345:1223–1229
43. Clifton GL, Miller ER, Choi SC, Levin HS, McCauley S, Smith KR Jr, Muizelaar JP, Wagner FC Jr, Marion DW, Luerssen TG, Chesnut RM, Schwartz M (2001) Lack of effect of induction of hypothermia after acute brain injury. N Engl J Med 344:556–563
44. Jüni P, Witschi A, Bloch R, Egger M (1999) The hazards of scoring the quality of clinical trials for meta-analysis. JAMA 282:1054–1060
45. Devereaux PJ, Manns BJ, Ghali WA, Quan H, Lacchetti C, Montori VM, Bhandari M, Guyatt GH (2001) Physician interpretations and textbook definitions of blinding terminology in randomized controlled trials. JAMA 285:2000–2003
46. Dickinson K, Bunn F, Wentz R, Edwards P, Roberts I (2000) Size and quality of randomised controlled trials in head injury: review of published studies. BMJ 320:1308–1311
47. Gøtzsche PC (1989) Methodology and overt and hidden bias in reports of 196 double-blind trials of nonsteroidal antiinflammatory drugs in rheumatoid arthritis. Control Clin Trials 10:31–56
48. Schulz KF, Chalmers I, Grimes DA, Altman DG (1994) Assessing the quality of randomization from reports of controlled trials published in obstetrics and gynaecology journals. JAMA 272:125–128
49. Altman DG, Dore CJ (1990) Randomisation and baseline comparisons in clinical trials. Lancet 335:149–153
50. Adetugbo K, Williams H (2000) How well are randomized controlled trials reported in the dermatology literature? Arch Dermatol 136:381–385
51. Peto R, Baigent C (1998) Trials: the next 50 years. Large scale randomised evidence of moderate benefits. BMJ 317:1170–1171
52. Thornley B, Adams C (1998) Content and quality of 2000 controlled trials in schizophrenia over 50 years. BMJ 317:1181–1184
53. Lemarie F (1995) Intensive Care Medicine. A European journal. Intensive Care Med 21:1–2