**Author's Final Version of:** Doig GS, Simpson F and Delaney AP. A review of the true methodological quality of nutritional support trials conducted in the critically ill: Time for improvement. *Anesthesia & Analgesia* 2005;100(2):527-33.

# A review of the true methodological quality of nutritional support trials conducted in the critically ill: Time for improvement.

Gordon S. Doig [1] Ph.D. (Epidemiology and Biostatistics), Fiona Simpson [2] M.Sc. (Nutrition and Dietetics) and Anthony Delaney [3] MBBS, Fellow of the Australian College of Emergency Medicine and Fellow of the Joint Faculty of Intensive Care Medicine.

[1] Northern Clinical School, University of Sydney, Department of Intensive Care, Royal North Shore Hospital, Sydney, Australia, [2] Department of Nutrition, Royal North Shore Hospital, Sydney, Australia and [3] Foothills Medical Centre, Calgary, Alberta, Canada.

*Corresponding Author:*

Dr. Gordon S. Doig,
Senior Lecturer in Intensive Care
Northern Clinical School
University of Sydney

*Mailing Address:*

Royal North Shore Hospital
Pacific Highway
St. Leonards, NSW
Australia 2065

*phone:* 612 9926 8656
*fax:* 612 9439 8418
*e-mail:* gdoig@med.usyd.edu.au

**Running Head:** Methodological quality of nutritional support RCTs

Reprints will not be available from the authors.

This work was supported in part by grants from the Australian and New Zealand Intensive Care Foundation and the NorthCare Foundation.

**Author's Final Version of:** Doig GS, Simpson F and Delaney AP. A review of the true methodological quality of nutritional support trials conducted in the critically ill: Time for improvement. *Anesthesia & Analgesia* 2005;100(2):527-33.

## Abstract

The purpose of this review is to appraise the true methodological quality of nutritional support studies conducted in critically ill patients and to compare these findings to the methodological quality of sepsis trials. An extensive literature search revealed 111 randomized controlled trials conducted in critically ill patients evaluating the impact of nutritional support interventions on clinically meaningful outcomes. Compared to sepsis trials, nutritional support studies were significantly less likely to use blinding (32/40 vs. 35/111, $p<0.001$) or present an intention-to-treat (ITT) analysis (37/40 vs. 64/111, $p<0.001$). There was a trend towards the less frequent use of randomization methods that are known to maintain allocation concealment (12/40 vs. 19/111, $p=0.10$). Although nutritional support studies demonstrated a significant increase in the use of blinding after the publication of the CONSORT statement in 1996 (9/47 vs. 26/64 post-CONSORT, $p=0.023$), there were no improvements in other key areas. Previous publications have described the overall methodological quality of sepsis trials as "poor". Nutritional support studies were significantly worse that sepsis trials in all aspects of methodological quality and there were few improvements noted over time. In order to detect important differences in clinically meaningful outcomes, the methodological quality of future studies must be improved.

**Key (indexing) words:** parenteral nutrition, enteral nutrition, intensive care, methodological quality, clinical trials, randomized controlled trials

## Introduction

Nutritional support is an important aspect of the management of all hospitalized patients. However, the provision of nutritional support to the critically ill patient is highly variable and often sub-optimal (1). One possible way to reduce the variability and improve the appropriateness of nutritional support practices is to promote the uptake of high-quality evidence (2,3).

A problem that faces clinicians when examining the literature is how to discern which trials are of sufficient methodological quality to merit changes to clinical practice. Previous systematic reviews of nutritional support studies (4,5) have identified high-quality trials using composite methodological quality scales (6). Recent research suggests that composite quality scales may actually mask important differences in true methodological quality (7).

In an evaluation of a 25-item composite quality scale based on the 1996 Consolidated Standards of Reporting Trials (CONSORT) statement (8), Huwiler-Muntener et al. found that "the true quality of a substantial portion of well conducted trials and of trials of low methodological quality will be misjudged." True methodological quality was defined using three key criteria: 1) concealment of random treatment allocation, 2) appropriate use of blinding, and 3) presentation of an intention-to-treat (ITT) analysis. These three key criteria are relevant to all types of clinical trials and when they are not addressed, estimates of benefit (or harm) obtained from any RCT are more likely to be biased (9).

The purpose of this review is to use the three key criteria of true methodological quality to assess nutritional support studies conducted in the critically ill, and to determine if true quality is improving over time. To provide a benchmark reference, the true methodological quality of nutritional support studies will be compared to results obtained from a previously published review of clinical trials of sepsis therapies (10).

## Methods

### Literature Retrieval

A comprehensive literature search was conducted to detect all primary nutritional support studies carried out in critically ill patient populations. A Medline search (from 1966) using the PubMed search engine (www.PubMed.org) was cross referenced with an EMBASE search (from 1980) undertaken using the OVID search engine (www.Ovid.com). Extensive search statements were developed specific to each search engine (PubMed and Ovid) in order to detect all possible primary nutritional support studies, overviews of primary nutritional support studies, and evidence-based guidelines. The reference lists of retrieved overviews and guidelines were hand searched for additional studies. Experts in the field and in industry were contacted in order to contribute any papers that may have been missed. The final close-out date for the search process was April 30, 2003.

### Study Selection

All randomized controlled trials comparing primary nutritional support interventions, conducted in adult critically ill patient populations, and reporting clinically meaningful outcomes (11) were eligible for consideration. Publications based on subgroups of patients from a larger published trial and studies reporting only surrogate outcomes (12) were not eligible. Detailed appraisal was restricted to manuscripts published in the English language (13).

A study was determined to have been conducted in a critically ill patient population if the manuscript reported: 1) the patients were recruited in an intensive care unit (ICU), or 2) the inclusion criteria described were such that the patients would normally be cared for in an ICU (e.g., all patients were receiving invasive mechanical ventilatory support), or 3) the patients were suffering from a condition that usually requires care in an ICU (e.g., severe thermal burns of > 40-50% TBSA, multiple

trauma that required urgent laparotomy), or 4) the patients had an average ICU length of stay (LOS) of greater than two days, or 5) a majority of the patients received a therapy that is delivered in the ICU (e.g., invasive mechanical ventilation), or 6) a severity of illness score was reported that was commensurate with the patient being critically ill.

A study was judged not to involve a critically ill patient population if none of the above criteria were met and: 1) the patients had simple operative procedures that would not normally require admission to an ICU (e.g., simple gastrectomy or hemi-colectomies) or 2) the exclusion criteria were such that patients with complicating medical conditions that might require admission to an ICU, such as cardiac failure, renal failure, diabetes or liver impairment, were not enrolled, or 3) the course of the patients' care was reported as uncomplicated (e.g., routine surgery, oral intake day one and then discharge from hospital day five or six).

We employed a broad definition of trials of nutritional support to include any comparison of the process of providing, or provision of, macronutrients, micronutrients, vitamins, and/or minerals. Parenteral nutrition was defined as an intravenous solution containing protein and a source of nonprotein energy with or without lipids.

## True Methodological Quality Criteria

Manuscripts of all included trials were assessed for the presence of three key criteria: 1) maintenance of allocation concealment in the randomization process, 2) use of blinding at any level (blinding of research personnel, healthcare workers, data collectors, outcome adjudicators, biostatistician, etc.), and 3) presentation of data such that an intention-to-treat (ITT) analysis could be performed.

In addition to the three key criteria listed above, each manuscript was also assessed on the overall adequacy of the statement describing the randomization process, and, if loss to follow-up was

present, whether missing outcomes were reported by treatment arm. Additionally, loss to follow-up was quantified and judged as to whether it was excessive. Excessive loss to follow-up was defined *a priori* as missing outcomes on more than 10% of all randomized patients (10).

The appraisal of nutritional support studies was conducted independently by all three authors (GSD, FS and AD). Any differences of opinion were resolved by discussion.

## Changes over time

The original CONSORT statement was published in 1996 (8). To determine whether methodological quality changed over time, trials published in 1996 or earlier were compared to trials published after 1996 (post-CONSORT).

## Sepsis Trials

The nature of the literature search, study selection, appraisal process, and methodological quality of the included sepsis trials has been reported elsewhere (10). Sepsis trials reporting clinically meaningful outcomes (11) were eligible for comparison. All sepsis trials were originally graded using an extensive 57-item composite scale, which included an assessment of the key domains of true methodological quality listed above. Primary comparisons between nutritional support studies and sepsis trials were restricted to these key domains.

## Statistical Comparisons

All p values reported for dichotomous outcomes were obtained using Fisher's Exact Test. A p-value $\leq 0.05$ was considered to indicate statistical significance. A p value $\leq 0.10$ but greater than 0.05 was considered to indicated a trend towards statistical significance.

# Results

## Literature Retrieval and Study Selection

The initial Medline/EMBASE electronic search retrieved 2,287 abstracts. Hand searching of abstracts and reference lists of all overviews and guidelines (GSD and FS) resulted in the retrieval of 465 papers. Of these 465 papers, 337 appeared to be primary nutritional support studies and were identified for detailed review (GSD, FS and AD). On detailed review 103 studies were found not to report any clinically meaningful outcomes, 42 were not conducted in critically ill patients, 27 were not primary nutritional support studies (i.e., evaluations of recombinant human growth hormone, insulin, etc.), 15 were cross-over studies, 12 evaluated pre-operative interventions, 8 were true observational studies (not controlled trials), 7 were non-English-language studies, 6 were pseudo-randomized, 5 were based on subgroups of patients from a larger published trial, and 1 was a post-operative intervention (oral intake for ten weeks post surgery). The remaining 111 papers were found to be primary nutritional support studies reporting clinically meaningful outcomes (11) conducted in critically ill patient populations. A complete listing of all 111 papers is presented in Appendix A.

## Eligible Studies

Eight thousand, three hundred and one patients were randomized into the 111 eligible nutritional support studies. The median number of patients randomized per study was 51, with a range from 12 to 398. Only 3 of the 111 nutritional support studies adequately addressed all three domains of true quality. Nine of the 111 studies failed to address *any* of the three key criteria. With regard to the maintenance of allocation concealment and presentation of results in an ITT format, 12/111 studies addressed both and 40/111 failed to address either.

Forty sepsis trials reporting clinically meaningful outcomes were eligible for primary comparison. The eligible sepsis trials randomized 13,240 patients, with a median of 121 patients per trial and a range from 22 to 2,199 patients. Further details on the eligible sepsis trials are presented elsewhere (10).

## True Methodological Quality

### *Maintenance of Allocation Concealment and Method of Randomization*

Seventeen percent (19/111) of nutritional support studies used a randomization approach that is known to adequately maintain concealment of the allocation process. Twenty-seven percent (30/111) of studies adequately described the approach used to randomly allocate patients into treatment arms.

Compared to sepsis trials, there was a trend towards reporting the use of randomization methods that are less likely to maintain allocation concealment (12/40 vs. 19/111, p=0.10). Nutritional support studies were also significantly less likely to provide an adequate overall description of the method of random allocation employed (26/40 vs. 30/111, p<0.001).

### *Blinding*

Nutritional support investigators reported the use of blinding, at any level, in 31% of studies (35/111). Sepsis trials used blinding at any level significantly more often than nutritional support studies (32/40 vs. 35/111, p<0.001).

### *Intention-to-Treat Analysis and Loss to Follow-up*

Nutritional support studies presented results in an ITT format (outcomes reported on all randomized patients) 57% of the time (64/111). This was significantly less often than sepsis trials (37/40 vs. 64/111, p<0.001).

Eighteen percent of all nutritional support studies had an excessive degree of loss to follow-up (20/111). Compared to sepsis trials, nutritional support studies were significantly more likely to have excessive loss to follow-up (0/40 vs. 20/111, $p<0.001$).

Seventy percent of nutritional support studies with less than 10% loss to follow-up reported patient loss by allocated treatment arm, whereas only 25% of nutritional support studies with excessive loss to follow-up (greater than 10%) reported losses by treatment arm (19/27 vs. 5/20, $p<0.001$). In nutritional support studies that failed to report patient outcomes, the median loss to follow-up was 9.5%, with a range of less than 1% to 34%.

## Changes over time

Sixty-four of the 111 nutritional support studies were published after the dissemination of the CONSORT statement in 1996 (post-CONSORT). There was no significant improvement over time in the maintenance of allocation concealment (5/47 vs. 14/64 post-CONSORT, $p=0.14$). However, significantly more post-CONSORT trials provided an adequate description of the method of random allocation actually used to assign patients to treatment groups (7/47 vs. 23/64, $p=0.017$).

Although there was a significant increase in the use of blinding in the post-CONSORT period (9/47 vs. 26/64, $p=0.023$), there was no significant improvement in the presentation of ITT results (29/47 vs. 35/64, $p=0.56$). In the 47 trials with documented loss to follow-up, there was no significant improvement in the reporting of losses by study arm (10/18 vs. 14/29 post-CONSORT, $p=0.766$).

## Discussion

Compared to clinical trials of sepsis therapies, nutritional support studies consistently performed worse in all three key domains of true methodological quality: maintenance of allocation concealment, appropriate use of blinding, and presentation of results using an ITT analysis. Although nutritional support studies published after the 1996 CONSORT statement were more likely to use blinding, there were no improvements over time in the other two key measures of true methodological quality.

It is possible that sepsis trials do not represent a “gold standard” for the conduct of clinical trials to which nutritional support studies should be held. Based on the findings of a previously published review (10), the overall methodological quality of sepsis trials has been described as “poor” (14). Indeed, an informal comparison to other multidisciplinary trials suggests that there is much room for improvement in sepsis trials (15). For these reasons, it is possible that sepsis trials provide an unreasonably low standard with which nutritional support studies have been compared.

There are many reasons that could explain the current disparity in methodological quality between sepsis trials and nutritional support studies. It is possible that the more stringent licensing requirements that must be addressed by most sepsis therapies results in better designs. It is also possible that sepsis trials receive better funding, are more often peer reviewed, or are more likely to have a collaborative approach to their design. Any of these explanations is possible; however, none justifies the continued poor conduct of future nutritional support studies. An increase in the use of blinding by nutritional support investigators since the publication of the CONSORT statement suggests that improvements in other domains are possible. The use of sealed, opaque, sequentially numbered envelopes is a cheap and effective way to maintain allocation concealment in any clinical trial.

Similarly, complete follow-up and reporting on all randomized patients, especially if outcomes are readily available, is all that is required to present an ITT analysis.

*Evaluating true methodological quality*

Despite the existence of excellent resources to guide the trialist in the conduct (16,17) and reporting (18) of RCTs, methodological deficiencies have been reported in many areas of medical research (19,20,21,22,23,24). Although there is evidence to suggest that methodological quality is improving over time (15,10), objective research specifically evaluating intensive care-based clinical trials suggests “that more consideration to the methodological quality” (25) is needed.

The majority of methodological appraisals use a composite scale to assess overall quality; however, there has been controversy in the literature over the appropriateness of this approach (14). The goal of a composite scale is to numerically combine information on different features of a trial into an overall score. Many different published scales exist and they “differ from one another in almost every respect: how and why the items were selected for inclusion, the number of items, reliability, approximate time to complete, and scoring range. Little attention has been given to the construct that the scales are assessing. With one exception, the scales are uniformly weak in how they were developed” (26).

Although there may be disagreement over the relative importance of many items included in composite scales, extensive reviews of the methodological literature consistently agree on the importance of three criteria: 1) the maintenance of allocation concealment (9), 2) the presentation of results from all randomized patients in the form of an ITT analysis (27), and 3) the appropriate use of blinding (28). Appropriate application of these three design features consistently results in a more unbiased estimate of treatment effect. Because recent methodological research has demonstrated that

overall scores based on composite scales may not adequately reflect the presence of these three key criteria (7), it has been recommended that an assessment of true methodological quality should be based on an appraisal of these three key criteria separately (9).

*Maintenance of Allocation Concealment and Method of Randomization*

The term *allocation concealment* is used to describe a process of randomizing patients into a clinical trial that protects researchers, clinicians, and patients from predicting upcoming group assignments. Concealing the knowledge of upcoming group assignments "prevents researchers from (unconsciously or otherwise) influencing which participants are assigned to a given intervention group" (29). As early as 1898, medical investigators realized the importance of removing 'subjective judgement' from the process used to assign patients to treatment groups (30). Indeed, although many believe that the 1948 Medical Research Council trial of streptomycin in tuberculosis (31) is a landmark trial because it was the first to use random allocation, it was truly innovative because it was the first trial to use sealed envelops in order to maintain allocation concealment (32).

An extensive review of 148 trials published in the *BMJ, JAMA* and the *Lancet* found that 50% of the trials used a randomization technique that is known to maintain allocation concealment however the reporting of allocation concealment improved significantly, from 39% to 61%, in all three journals after the publication of the original CONSORT statement (15). We found that only 17% of published nutritional support studies used a process for allocating patients to treatment groups that maintained allocation concealment, which was approximately half as often as sepsis trials (30%). Numerous reviews have demonstrated that trials with inadequate or unclear allocation concealment can produce up to 40% greater estimates of treatment effects (33).

In general, it is very difficult to maintain allocation concealment if some form of alternating group allocation process (pseudo-randomization) is used (32). Examples of sentences describing the randomization process obtained from nutritional support studies include: “Patients were randomized into one of two nutritional subgroups (Total Parenteral Nutrition or Naso-Gastric feeding) based on their date of admission” (34) and “Patient selection was carried out in a random fashion, patients being assigned alternately to one or the other group according to the order in which they were admitted to the ICU” (35). Although both of these trials clearly assign patients to treatment based on an alternating group allocation process (pseudo-randomization), both also incorrectly claim to be ‘randomized’. These examples serve to illustrate that in order to determine whether allocation concealment was maintained, it is extremely important for trials to include a sentence adequately describing the randomization process (32).

The simplest randomization process that maintains allocation concealment is the use of ‘sequentially numbered, opaque, sealed envelopes’ that are opened *after* a patient is recruited into a trial (33). Likewise, a centrally controlled randomization process, whereby a randomization center is contacted by phone, fax, e-mail, or via the web *after* a patient is recruited into the trial, is also an excellent way to maintain allocation concealment (33). The use of either approach requires only one sentence to describe adequately, however it should be noted that the minimum description of either approach includes the statement that randomization occurred *after* the patient was screened as truly eligible and recruited into the trial.

*Blinding*

The term ‘double blind’ is frequently used to refer to a process whereby both the patient and the healthcare team are unaware of which study treatment the patient is actually receiving. Because it may

be important to understand whether the researchers, outcome adjudicators, data collectors, and even the analysts were also blinded, use of the term 'double blinded' is discouraged and an explicit list of exactly who was blinded is preferable (36).

Our appraisal found that nutritional support studies employed blinding significantly less often than sepsis trials (31% vs. 80%, $p<0.001$). Although there has been a significant increase in the use of blinding in the nutritional support literature since the publication of the CONSORT statement (19% vs. 40%, $p=0.023$), blinding was still employed half as often as it was in sepsis trials.

We appreciate that it may be easier to achieve blinding in a trial that compares simple IV infusions, such as many sepsis trials. Although some are quick to claim that blinding is not possible when interventions are more complex, the critical care literature is replete with examples where trialists have been able to develop novel and innovative processes for establishing blinding of complex interventions (37,38). Indeed, the appropriate use of blinding may decrease overoptimistic estimates of treatment effect by up to 26% (39). Regardless of the intervention evaluated in a trial, if a subjective outcome such as ventilator-associated pneumonia or suspected infection is important, it is always possible to blind outcome adjudicators. Likewise, in order to ensure that the accuracy and completeness of follow-up and data collection are equivalent in both groups, it is usually possible to blind the primary data collectors.

*Intention-to-Treat Analysis and Loss to Follow-up*

An intention-to-treat analysis compares outcomes obtained from all patients enrolled and randomized into a clinical trial. Inclusion in an ITT analysis does not depend on whether patients actually satisfied the study entry or exclusion criteria, whether they actually received treatment, whether a protocol violation was recorded, or whether treatment was discontinued (27). In order to

conduct an ITT analysis, full and complete follow-up is required on all patients randomized into the trial. As compared to an ITT analysis, an *efficacy subset analysis* may include only patients who received an adequate dose of the study treatment, who were truly eligible for study enrollment (met all eligibility criteria and none of the exclusion criteria), or who did not have any protocol violations.

Although the study protocol may identify the criteria used to define the efficacy subset analysis before the trial begins, because the subset of patients themselves cannot be identified until after randomization, it cannot be claimed that the properties of randomization apply or that the subset provides an unbiased assessment of treatment effects (40). An efficacy subset analysis is more susceptible to bias than many investigators believe.

For example, consider an efficacy subset analysis conducted on 80% of patients selected from a 200-patient study, with 20 patients excluded from each arm. Because the exclusion rate is similar in each arm, it might appear that there is very little chance of bias. Simulation studies conducted under reasonable assumptions have demonstrated that with as little as 10% loss to follow-up in each arm, the chance of obtaining a false positive result can easily double (40). The only way to ensure that bias has not entered an efficacy subset analysis, and resulted in a false positive result, is to report the actual outcomes for each patient not included in the analysis. In essence, this requires complete follow-up on all randomized patients until “the death of the patient…. or the end of the study” (40).

Fifty-seven percent of nutritional support studies presented results in an ITT format. Because patients may withdraw informed consent after randomization, which leads to true loss to follow-up, it may not be reasonable to expect 100% follow-up from every single trial. However, 92% of sepsis trials presented a true ITT analysis.

The majority of nutritional support studies that failed to report outcomes on randomized patients presented an efficacy subset analysis in preference to an ITT analysis. In such a situation, it is always

possible that outcomes *were* available on the missing patients and that investigators simply chose not to report outcomes on patients who did not qualify for the efficacy subset analysis. Failure to record or report an outcome on a patient who is not truly lost to follow-up should be strongly discouraged. The only way to determine whether an efficacy subset analysis is biased is to compare the results to those obtained in the ITT analysis (40). This requires full follow-up on all randomized patients.

*Summary*

Although others have emphasized that many aspects of methodology—such as issues of power and sample size (41) and the use of explicit, repeatable eligibility criteria to define target patient populations—need to be addressed to obtain reliable results from nutritional support studies (42), we evaluated three key methodological quality criteria that are universally accepted to reduce bias. Compared to clinical trials of sepsis, nutritional support studies were significantly less likely to use blinding or present an ITT analysis. There was a statistical trend towards the less frequent use of a randomization process that prevents researchers from predicting the group assignment sequence (maintenance of allocation concealment). Although nutritional support studies published after the 1996 CONSORT statement were more likely to use blinding, there were no improvements over time in the other two key criteria of true quality.

The use of three simple design features can improve the reliability of the results obtained from any clinical trial. Sealed, opaque, sequentially numbered envelopes are a cheap and effective way to maintain allocation concealment. There are many different levels at which blinding can be used to improve the reliability of any trial, and reporting of outcomes of all randomized patients, regardless of whether an efficacy subset analysis is conducted, should be considered mandatory. If appropriate nutritional support of the critically ill truly can improve mortality by 10% to 13% (3), the only way

clinical trials will be able to detect an improvement of this magnitude is by implementing key design criteria that are known to reduce bias.

In the absence of definitive (Level I) evidence (43), systematic reviews and evidence-based guidelines based on high-quality trials can support clinical decision making (44). However, the strength of the conclusions reached by overviews and guidelines is intimately related to the quality of the individual trials included (39). Because trials that assign patients to treatment based on an alternating group allocation process (pseudo-randomization) and those that contain excessive (> 10%) loss to follow-up are prone to severe bias, we strongly recommend that these trials not be included in systematic reviews and evidence-based guidelines. In addition, wherever possible, methodologists should also conduct sensitivity analyses to determine if clinical recommendations differ when they are based on high vs. low quality evidence as determined by the three key measures of true methodological quality.

## Reference List

1. Heyland DK, Schroter-Noppe D, Drover JW, et al. Nutrition support in the critical care setting: current practice in Canadian ICUs—opportunities for improvement? J Parenter Enteral Nutr 2003; 27:74-83.

2. Linton AL, Peachey DK. Guidelines for medical practice: 1. The reasons why. CMAJ 1990; 143:485-90.

3. Martin CM, Doig GS, Heyland DK, et al. Cluster randomized clinical trial of Algorithms for Critical Care Enteral and Parenteral Therapy (ACCEPT). CMAJ 2004;170:197-204.

4. Heyland DK, MacDonald S, Keefe L, Drover JW. Total parenteral nutrition in the critically ill patient: a meta-analysis. JAMA 1998; 280:2013-19.

5. Heyland DK, Novak F, Drover JW, et al. Should immunonutrition become routine in critically ill patients? A systematic review of the evidence. JAMA 2001; 286:944-53.

6. Jadad AR, Moore RA, Carroll D, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? Control Clin Trials 1996; 17:1-12.

7. Huwiler-Muntener K, Juni P, Junker C, Egger M. Quality of reporting of randomized trials as a measure of methodologic quality. JAMA 2002; 287:2801-4.

8. Begg C, Cho M, Eastwood S, et al. Improving the quality of reporting of randomized controlled trials. The CONSORT statement. JAMA 1996; 276:637-9.

9. Egger M, Juni P, Bartlett C, et al. How important are comprehensive literature searches and the assessment of trial quality in systematic reviews? Empirical study. Health Technol Assess 2003; 7:1-76.

10. Graf J, Doig GS, Cook DJ, et al. Randomized, controlled clinical trials in sepsis: has methodological quality improved over time? Crit Care Med 2002; 30:461-72.

11. Prentice RL. Surrogate endpoints in clinical trials: definition and operational criteria. Stat Med 1989; 8:431-40.

12. Novak F, Heyland DK, Avenell A, et al. Glutamine supplementation in serious illness: a systematic review of the evidence. Crit Care Med 2002; 30:2022-9.

13. Juni P, Holenstein F, Sterne J, et al. Direction and impact of language bias in meta-analyses of controlled trials: empirical study. Int J Epidemiol 2002; 31:115-23.

14. Dhingra V, Chittock DR, Ronco JJ. Assessing methodological quality of clinical trials in sepsis: searching for the right tool. Crit Care Med 2002; 30:487-8.

15. Moher D, Jones A, Lepage L. Use of the CONSORT statement and quality of reports of randomized trials: a comparative before-and-after evaluation. JAMA 2001; 285:1992-5.

16. Meinart CL. Clinical trials: Design, conduct and analysis. New York: Oxford University Press, 1986.

17. Pocock SJ. Clinical trials: a practical approach. Chichester: John Wiley & Sons, 1983.

18. Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. Lancet 2001; 357:1191-4.

19. Sanchez-Thorin JC, Cortes MC, Montenegro M, Villate N. The quality of reporting of randomized clinical trials published in *Ophthalmology*. Ophthalmology 2001; 108:410-15.

20. Koes BW, Bouter LM, van der Heijden GJ. Methodological quality of randomized clinical trials on treatment efficacy in low back pain. Spine 1995; 20:228-35.

21. Junker CA. Adherence to published standards of reporting: a comparison of placebo-controlled trials published in English or German. JAMA 1998; 280:247-9.

22. Adetugbo K, Williams H. How well are randomized controlled trials reported in the dermatology literature? Arch Dermatol 2000; 136:381-5.

23. Zhang B, Schmidt B. Do we measure the right end points? A systematic review of primary outcomes in recent neonatal randomized clinical trials. J Pediatr 2001;138:76-80.

24. Linde K, Jonas WB, Melchart D, Willich S. The methodological quality of randomized controlled trials of homeopathy, herbal medicines and acupuncture. Int J Epidemiol 2001;30:526-31.

25. Latronico N, Botteri M, Minelli C, et al. 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; 28:1316-23.

26. Moher D, Jadad AR, Nichol G, et al. Assessing the quality of randomized controlled trials: an annotated bibliography of scales and checklists. Control Clin Trials 1995; 16:62-73.

27. Hollis S, Campbell F. What is meant by intention to treat analysis? Survey of published randomised controlled trials. BMJ 1999; 319:670-74.

28. Kaptchuk TJ. The double-blind, randomized, placebo-controlled trial: gold standard or golden calf? J Clin Epidemiol 2001; 54:541-9.

29. Definition of allocation concealment. CONSORT Statement web site: http://www.consort-statement.org/allocationconcealment.htm. Accessed 13-1-2004.

30. Hrobjartsson A, Gotzsche PC, Gluud C. The controlled clinical trial turns 100 years: Fibiger's trial of serum treatment of diphtheria. BMJ 1998; 317:1243-5.

31. Medical Research Council. Streptomycin treatment of pulmonary tuberculosis: a Medical Research Council investigation. BMJ 1948; ii:769-82.

32. Chalmers I. Comparing like with like: some historical milestones in the evolution of methods to create unbiased comparison groups in therapeutic experiments. Int J Epidemiol 2001; 30:1156-1164.

33. Schulz KF, Grimes DA. Allocation concealment in randomised trials: defending against deciphering. Lancet 2002; 359:614-8.

34. Hadley MN, Grahm TW, Harrington T, et al. Nutritional support and neurotrauma: a critical review of early nutrition in forty-five acute head injury patients. Neurosurgery 1986; 19:367-73.

35. Jimenez Jimenez FJ, Ortiz LC, Morales MS, et al. Prospective study on the efficacy of branched-chain amino acids in septic patients. J Parenter Enteral Nutr 1991; 15:252-61.

36. Schulz KF, Chalmers I, Altman DG. The landscape and lexicon of blinding in randomized trials. Ann Intern Med 2002; 136:254-9.

37. Dellinger RP, Zimmerman JL, Taylor RW, et al. Effects of inhaled nitric oxide in patients with acute respiratory distress syndrome: results of a randomized phase II trial. Inhaled Nitric Oxide in ARDS Study Group. Crit Care Med 1998; 26:15-23.

38. Finfer S, Bellomo R, Myburgh J, Norton R. Efficacy of albumin in critically ill patients. BMJ 2003; 326:559-60.

39. Juni P, Altman DG, Egger M. Systematic reviews in health care: Assessing the quality of controlled clinical trials. BMJ 2001; 323:42-6.

40. Lachin JM. Statistical considerations in the intent-to-treat principle. Control Clin Trials 2000; 21:167-89.

41. Halpern SD, Karlawish JH, Berlin JA. The continuing unethical conduct of underpowered clinical trials. JAMA 2002; 288:358-62.

42. Preiser JC, Chiolero R, Wernerman J. Nutritional papers in ICU patients: what lies between the lines? Intensive Care Med 2003; 29:156-66.

43. LeLorier J, Gregoire G, Benhaddad A, et al. Discrepancies between meta-analyses and subsequent large randomized, controlled trials. N Engl J Med 1997; 337:536-42.

44. Smith R, Chalmers I. Britain's gift: a "Medline" of synthesised evidence. BMJ 2001; 323:1437-8.