

Systematic Evaluation of the Quality of Randomized Controlled Trials in Diabetes

VICTOR M. MONTORI, MD, MSc^{1,4}
YAQIAN GRACE WANG, BHSC²

PABLO ALONSO-COELLO, MD³
SUMIT BHAGRA, MBBS⁴

OBJECTIVE — We sought to systematically ascertain the quality of randomized controlled trials (RCTs) in diabetes.

RESEARCH DESIGN AND METHODS — We identified the 10 most recently published trials as of 31 October 2003 in each of six general medical, five diabetes, and five metabolism and nutrition journals and further enriched our sample with 10 additional RCTs from each of five journals that published the most eligible RCTs in a year. We explored the association between trial characteristics and reporting quality using univariate analyses and a preplanned multivariate regression model.

RESULTS — After excluding redundant reports of included trials and one trial that measured outcomes on the health system and not on patients, we included 199 RCTs: 119 assessed physiological and other laboratory outcomes, 42 assessed patient-important outcomes (e.g., morbidity and mortality, quality of life), and 38 assessed surrogate outcomes (e.g., disease progression or regression, HbA_{1c}, cholesterol). Fifty-three percent were of low methodological quality, as were one-third (36–40%) of trials reporting patient-important or surrogate outcomes and two-thirds (64%) of laboratory investigations. Independent predictors of low quality were nonprofit funding source (odds ratio 3.1 [95% CI 1.5–6.2]), measure of physiological and laboratory outcomes (2.3 [1.2–4.4]), and cross-over design (2.3 [1.1–4.8]), all characteristics of laboratory clinical investigations.

CONCLUSIONS — There is ample room for improving the quality of diabetes trials. To enhance the practice of evidence-based diabetes care, trialists need to pay closer attention to the rigorous implementation and reporting of important methodological safeguards against bias in randomized trials.

Diabetes Care 29:1833–1838, 2006

A key principle of evidence-based practice is that one should seek to apply the best available evidence from clinical research (1,2). The expression “best available” suggests a hierarchy of evidence; one ought to draw stronger inferences from evidence that comes from high-quality studies with optimal safeguards to prevent random and systematic error (bias). Most hierarchies of evidence

about interventions place high-quality randomized controlled trials (RCTs) at the top of the hierarchy (3). Following this principle, diabetes practitioners should pay particular attention to RCTs to guide their practice.

Not all RCTs share the same quality; that is, not all RCTs yield unbiased results. In the laboratory clinical investigation tradition, “quality” has referred

almost exclusively to the rigorousness and reproducibility of the experimental procedures performed on the volunteers as well as the precise and accurate nature of the laboratory determinations. In conducting systematic reviews of RCTs in diabetes (4–6), we have noticed that investigators seem to pay little attention (as judged by the extent to which they report these methods) to methodological safeguards that limit the introduction of bias into RCTs. As a result, these potentially biased RCTs could mislead clinicians. Readers only have access to the methods as reported. When reports leave out critical information about methodological safeguards against bias, readers cannot ascertain whether those safeguards were present or not during trial conduct.

To our knowledge, there is no contemporary systematic assessment of the quality of RCTs in diabetes. Ten years ago, McIver and Dinneen (7) conducted the first evaluation of the quality of RCTs in diabetes and found it lacking. In updating their work, we sought to systematically evaluate the methodological quality of RCTs in diabetes (as reported in major medical journals), taking into account the advances made in the last 10 years in the conduct and reporting of RCTs.

RESEARCH DESIGN AND METHODS

Eligibility criteria

Eligible articles were reports of studies describing random allocation of human participants to at least two interventions, one of which was a control intervention; participants had to be patients with any form of diabetes or people at risk for developing any form of diabetes. When several reports referred to the same trial, we retained the original report.

Search strategy

To describe the best RCTs in diabetes, we sought to identify RCTs published in the top journals judged by 2003 impact-factor rankings. Furthermore, we purposefully selected both general journals ($n = 6$, *New England Journal of Medicine*, *JAMA*, *Lancet*, *Annals of Internal Medicine*, *BMJ*, and *Archives of Internal Medicine*) as

From the ¹Knowledge and Encounter Research Unit, Division of Diabetes and Endocrinology, Mayo Clinic College of Medicine, Rochester, Minnesota; the ²Bachelor of Health Sciences Program, Faculty of Health Sciences, McMaster University, Hamilton, Ontario, Canada; the ³Iberoamerican Cochrane Center, Department of Clinical Epidemiology and Public Health, Hospital de Sant Pau, Sant Antoni Maria Claret, Barcelona, Spain; and the ⁴Department of Medicine, Mayo Clinic College of Medicine, Rochester, Minnesota.

Address correspondence and reprint requests to Victor M. Montori, MD, MSc, Mayo W18A, Mayo Clinic, 200 First St. SW, Rochester, MN 55905. E-mail: montori.victor@mayo.edu.

Received for publication 11 January 2006 and accepted in revised form 28 April 2006.

Additional information on this article can be found in an online appendix available at <http://care.diabetesjournals.org>.

Abbreviations: CONSORT, Consolidated Standards of Reporting Trials; RCT, randomized controlled trial.

A table elsewhere in this issue shows conventional and Système International (SI) units and conversion factors for many substances.

DOI: 10.2337/dc06-0077

© 2006 by the American Diabetes Association.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

well as pertinent specialty journals ($n = 5$ in diabetes: *Diabetes*, *Diabetes Care*, *Diabetic Medicine*, *Diabetologia*, and *Diabetes and Metabolism Research*; and $n = 5$ in metabolism and nutrition: *Journal of Clinical Endocrinology and Metabolism*, *Metabolism: clinical and experimental*, *American Journal of Clinical Nutrition*, *Journal of the American College of Nutrition*, and *European Journal of Clinical Nutrition*).

To identify RCTs published in each of these journals, Y.G.W. (under the supervision of V.M.M.) conducted an online search using the PubMed interface (www.pubmed.gov) and the search terms “journal name” [Journal] and (diabet* [title] or niddm [title] or iddm [title]), limited to human studies in English indexed with abstracts. We then retrieved each apparently eligible abstract and read the article in full to determine its eligibility. We identified the 10 most recently published and eligible RCTs from each journal as of 31 October 2003. We further enriched our sample with 10 additional RCTs from each of the five journals that published the most eligible RCTs in a year.

Data extraction

We developed, pilot tested, and used a standardized form to abstract data from each of the eligible RCTs about research methods (allocation concealment, blinding, intention to treat, and loss to follow-up), statistical reporting (between-arm versus within-arm comparisons and use of CIs versus P values), and funding source. Also, we noted the patients enrolled (number and description) and the intervention types (e.g., drugs, procedures, diet).

Outcome and trial classification

To classify the reported primary outcomes and adverse events, we noted outcomes expected to directly impact on patients' quality of life (which we refer to as “patient-important outcomes” [8]), those that assess the response to physiological and other laboratory maneuvers (“physiological and laboratory outcomes”), and those that lie intermediate to these two classes, such as measures that may indicate an increased risk for patient-important outcomes (“surrogate outcomes”).

We considered trials to have low methodological quality when they met three or more of these criteria (online appendix [available at <http://care.diabetesjournals.org>]): inadequate (or not reported) allocation concealment, inadequate (or not reported) blinding of pa-

tients and of caregivers, failure to adhere to the intention-to-treat principle, or a reporting of $>10\%$ (or did not report information to calculate) loss to follow-up. Allocation concealment refers to the extent to which researchers assessed eligibility and to which enrolled patients were kept unaware of the randomization sequence such that they could not predict the arm of the trial to which the next patient would be allocated. Examples of adequate allocation concealment include central (online or phone-in) randomization and medication dispensing in coded containers. The explicit reporting of blinding of patients or caregivers or the statement that the trial was “double blind” and tested one intervention against a placebo qualified as adequate blinding. Given that most modern trials estimate small treatment effects, loss to follow-up $>10\%$ was considered inadequate. Adherence to the intention-to-treat principle requires minimal loss to follow-up and minimal cross over. We limited the assessment of adherence to this principle to the explicit reporting of the conduct of analyses by the intention-to-treat principle (i.e., statement or evidence that patients were analyzed in the arm to which they were randomized). In the case of trials with several reports (e.g., the index report referenced papers published earlier and describing the methods in detail), we considered all available reports to ascertain methodological quality.

Reproducibility

Since these classifications required judgments, we sought to standardize the use of the form through iterative use, item reduction or clarification, and re-review by the four abstractors (Y.G.W., P.A., S.B., and V.M.M.). We extracted data in duplicate until we achieved adequate reproducibility (chance-adjusted interrater reliability [κ] $>90\%$) and thereafter continued with individual data extraction.

Statistical analyses

This study uses descriptive statistics to characterize the quality of RCTs in diabetes. All univariate analyses explored associations to generate hypotheses related to the quality of RCTs and its relation to RCT characteristics.

From the literature and our previous work, we proposed the following predictors of quality of RCTs: journal of publication (general versus specialist [9]), publication of Consolidated Standards of Reporting Trials (CONSORT) flow chart

(as a surrogate marker for adherence to CONSORT statement [10,11]), parallel versus cross-over design (there are no recently published and widely endorsed standards for reporting of cross-over trials), and funding source (12). We also put forth that trials seeking to impact clinical practice by measure of patient-important outcomes and surrogate outcomes were of better quality than laboratory investigations.

To test these predictors, we constructed a multivariable model with all these predictors entered at once and tested this model, using logistic regression, with low quality (yes/no) as the dependent variable. Associations were described using odds ratios (ORs) and their associated 95% CIs.

RESULTS— After screening 357 potentially eligible abstracts, we selected 209 eligible RCTs, of which 9 were additional reports of RCTs already represented in the sample (e.g., subanalyses of the Diabetes Control and Complications Trial and U.K. Prospective Diabetes Study). One RCT focused its intervention on the delivery of health care and reported only process outcomes (i.e., degree of adherence to guidelines) and was excluded from further description. Table 1 describes the characteristics of the included 199 RCTs classified by type of outcome.

Types of trials and funding

Typical RCTs measuring patient-important and surrogate outcomes were single-center parallel-design drug trials enrolling patients with type 2 diabetes and funded by for-profit agencies (e.g., pharmaceutical companies). Typical RCTs measuring physiological and other laboratory measures were single-center parallel or cross-over design drug trials enrolling patients with type 2 diabetes and funded through mixed or not-for-profit funding (e.g., National Institutes of Health).

Outcomes

Most trials reported physiological and other laboratory outcomes ($n = 118$). Trials describing patient-important outcomes ($n = 42$) reported on the effect of interventions on one or more of the following: mortality ($n = 12$), major morbidity (such as myocardial infarction or stroke, $n = 10$), minor morbidity (such as transient ischemic attack and severe hypoglycemia, $n = 5$), physical or mental disability ($n = 4$), discomfort that hinders

Table 1—Description of included RCTs

Characteristics	Patient-important outcomes	Surrogate outcomes	Physiological and other laboratory measures
<i>n</i>	42	38	119
Year of publication	2002 (1998–2003)	2002 (1995–2003)	2002 (1987–2003)
Journal of publication*			
General medical	26 (62)	20 (53)	21 (18)
Diabetes	16 (38)	6 (16)	47 (39)
Metabolism and nutrition	0 (0)	12 (32)	51 (43)
Region of work			
North America	12 (29)	11 (29)	37 (31)
Europe	19 (45)	17 (45)	52 (44)
Asia	0 (0)	4 (11)	10 (8)
Other	11 (26)	6 (16)	20 (17)
Number of centers			
Single center	24 (57)	24 (63)	90 (76)
Few centers (<10)	3 (7)	6 (16)	20 (17)
Many centers (>10)	15 (36)	8 (21)	9 (8)
Type of patients			
Type 2 diabetes	15 (36)	19 (50)	82 (69)
Type 1 diabetes	8 (19)	5 (13)	20 (17)
Type 1 or 2 diabetes	9 (21)	4 (11)	3 (2.5)
At risk of diabetes	4 (10)	3 (8)	3 (2.5)
Other	6 (14)	7 (18)	11 (9)
Exclusion of women	3 (7)	1 (3)	13 (11)
Exclusion of elderly (were elderly could be eligible)	20/31 (64)	14/22 (64)	33/44 (75)
Exclusion of ethnic group	2 (5)	1 (3)	2 (2)
Type of interventions			
Drug	30 (71)	29 (76)	94 (79)
vs. placebo	14	15	62
vs. active drug	13	12	30
vs. usual care or no intervention	3	2	2
Nutrition	0 (0)	2 (5)	17 (14)
Education, counseling, or psychological interventions	8 (19)	4 (10)	4 (3)
Other interventions	4 (10)	3 (8)	4 (3)
Funding			
For-profit sources only	25 (60)	18 (47)	53 (45)
For-profit and not-for-profit sources	9 (21)	7 (18)	29 (24)
Not-for-profit sources only	8 (19)	13 (34)	37 (31)

Data are median (range) or *n* (%). *General medical journals: *New England Journal of Medicine* (*n* = 10), *Lancet* (*n* = 11), *JAMA* (*n* = 10), *BMJ* (*n* = 8), *Annals of Internal Medicine* (*n* = 10), and *Archives of Internal Medicine* (*n* = 10). Diabetes journals: *Diabetes Care* (*n* = 19), *Diabetes* (*n* = 17), *Diabetic Medicine* (*n* = 18), *Diabetologia* (*n* = 19), and *Diabetes/Metabolism Research and Reviews* (*n* = 10). Metabolism and nutrition journals: *Journal of Clinical Endocrinology and Metabolism* (*n* = 18), *Metabolism: clinical and experimental* (*n* = 11), *American Journal of Clinical Nutrition* (*n* = 10), *European Journal of Clinical Nutrition* (*n* = 8), and *Journal of the American College of Nutrition* (*n* = 10).

daily living (such as minor hypoglycemia, *n* = 8), and specific measure of quality of life (*n* = 7). Trials reporting surrogate outcomes (*n* = 38) reported on the effect of interventions on disease progression or regression (*n* = 12) or on laboratory measures such as HbA_{1c} (A1C) or cholesterol measures (*n* = 31).

Reported methodological quality of RCTs and its predictors

Table 2 describes the methodological characteristics of the included RCTs; 106 trials (53%) were considered of low methodological quality. While 36 and 40% of

trials reporting patient-important or surrogate outcomes were of low quality, respectively, 64% of the laboratory investigations were of low reported methodological quality.

Apart from lack of publication of a patient flow chart (OR 1.0 [95% CI 0.5–2.5]) and publication in a specialist journal (1.5 [0.7–3.2]), all variables included in the predefined model were independent significant predictors. Independent predictors of low quality were nonprofit funding source (3.1 [1.5–6.2]), report of physiological and laboratory outcomes (2.3 [1.2–4.4]), and cross-over design

(2.3 [1.1–4.8]). Much of the variability in methodological quality, however, remains unexplained; the complete model significantly predicted 14% of the variability in quality in the included RCTs (*P* < 0.0001).

Report elements to support evidence-based clinical decision making

Laboratory investigations were as likely as RCTs reporting patient-important and surrogate outcomes to report adverse events (*n* = 64 [54%] vs. *n* = 37 [46%]). These trials also followed fewer patients (median of 54 vs. 160 or 474, respec-

Table 2—Methodological quality of RCTs in diabetes

	Patient-important outcomes	Surrogate outcomes	Physiological and other laboratory outcomes	Overall
<i>n</i>	42	38	119	199
Study characteristics				
Parallel/cross-over trials	36/6	31/7	78/41	145/54
Study duration (weeks)				
Parallel trials	136 (26–235)	33 (24–129)	16 (10–48)	26 (12–104)
Cross-over trials	24 (5–42)	16 (1.3–51)	12 (2.6–24)	13 (3–27)
Randomized				
Parallel trials	474 (162–1,513)	160 (46–404)	54 (37–105)	97 (46–327)
Cross-over trials	23 (20–27)	35 (22–63)	16 (9–27)	20 (11–32)
Reporting				
Patient flow chart	14 (33)	8 (21)	18 (15)	40 (20)
IRB approval, informed consent	40 (95)	37 (97)	115 (97)	192 (97)
Estimation of sample size	25 (60)	19 (50)	34 (29)	78 (39)
Adverse events reported	16 (38)	21 (55)	64 (54)	101 (51)
Methods to minimize bias				
Adequate allocation concealment	9 (21)	4 (11)	8 (7)	21 (11)
Blinding				
Participants	20 (48)	17 (45)	62 (52)	99 (50)
Health care providers	19 (45)	20 (53)	52 (44)	91 (46)
Data collectors	9 (21)	3 (8)	11 (9)	23 (12)
Data analysts	3 (7)	1 (3)	7 (6)	11 (6)
Judicial assessors of outcomes	8 (19)	1 (3)	8 (7)	17 (9)
Only described as double blind	16 (38)	16 (42)	57 (48)	89 (45)
Intention-to-treat analysis	27 (64)	22 (58)	22 (18)	71 (36)
Loss to follow-up				
Parallel trials (%)	4 (0–24)	9 (2–20)	5 (0–14)	5 (0–17)
Cross-over trials (%)	3 (0–19)	0 (0–1)	0 (0–0)	0 (0–2)
Not reported	10 (24)	10 (26)	37 (31)	57 (29)
Low methodological quality	15 (36)	15 (40)	76 (64)	106 (53)

Data are median (interquartile range) or *n* (%) unless otherwise indicated. IRB, institutional review board.

tively) for a shorter period (median of 16 vs. 33 or 136 weeks, respectively). Laboratory investigations were less likely than RCTs measuring patient-important or surrogate outcomes to use estimation (CIs) to describe the precision of the results ($n = 24$ [20%] vs. $n = 40$ [50%], respectively). Also, we found that 13 of 78 (17%) parallel-design laboratory investigations failed to present the results of comparisons between intervention and control arms, presenting before-after comparisons within each arm.

CONCLUSIONS— RCTs in diabetes published in pertinent top journals, both general and specialized, have important deficiencies in their report of key methodological features. These deficiencies are most remarkable in laboratory investigations with trials that measured patient-important outcomes showing better reporting. Many RCTs measured patient-important outcomes, but very few of these assessed nonpharmacological interven-

tions. Despite the worldwide explosion of diabetes as a major public health problem, most trials came from researchers working in the northern hemisphere.

Limitations and strengths

By the nature of our selection process (i.e., sampling exclusively from top journals), our work likely overestimates the quality of diabetes RCTs in general as well as the proportion of these trials that measured patient-important outcomes. Thus, the deficiencies in reporting or methodological quality documented here may very well represent a “best case” scenario, therefore strengthening any calls for improvement in the conduct and report of RCTs in diabetes. Further, our reproducible methods using multiple judges and our focused analyses strengthen the inferences drawn from these data.

Comparison with previous research

McIver and Dinneen (7) evaluated 79 RCTs related to type 2 diabetes published

nearly 10 years ago (1994–1995) and indexed in Medline. Almost half (42%) of the trials they assessed were published in the journals *Diabetes Care* or *Diabetic Medicine*. We sampled 28% of our RCTs from these two journals. Compared with our cohort of RCTs, RCTs in the McIver and Dinneen cohort enrolled fewer patients (median patients randomized 40 [range 5–2,769]), followed them for shorter periods (median duration 22 weeks [0.2–260]), and were less likely to measure patient-important outcomes (9%). However, the proportion of RCTs adequately reporting allocation methods (15%) and blinding (58%, limited to “double blind”) seems similar 10 years later.

Of note, recent empirical evidence supports the explicit reporting of the allocation concealment process (13–15), of which groups (participants, clinicians, data collectors, data analysts, and judicial assessors of outcomes) were blinded (not just the term “double blind” [16,17]) and

the extent to which the trial was conducted under the intention-to-treat principle (18,19). While the revised CONSORT statement offers guidelines for the reporting of parallel RCTs (10), including the reporting of harms (20), we and others have documented gaps in the reporting of RCTs, even in RCTs published in journals that endorse, but fail to enforce, CONSORT (9,21,22).

Also, it is important to separate the reported methodological quality of an RCT with the actual rigor of its conduct. Generally critical readers have assumed that if something important (e.g., a methodological safeguard against bias) was not mentioned, it likely did not happen. We undertook an evaluation of this assumption (in which we interviewed the authors and compared their answers with what they reported) and found it excessively pessimistic (23). For example, while only 58% of trials reported adequate allocation concealment, >90% of the trials actually had implemented an adequate strategy to conceal the allocation sequence from personnel enrolling patients and assessing their eligibility. Extrapolating those results to this cohort of RCTs, it is likely that RCTs in diabetes have better methods than may be apparent from the careful review we conducted of their published reports. To mitigate this uncertainty, it is fair to suggest to trialists to work toward improving the report of their trials by adhering more closely to standards of reporting such as the CONSORT statement. Journals also can do their part in enforcing adherence to the CONSORT standards they endorse (9). Furthermore, key reporting elements that are thought to use too much printed space (even though CONSORT sets a minimal reporting requirement) by editors and authors could appear in electronic appendixes in journal websites linked to the original publication instead of being edited out.

Despite our sampling from top journals, only 20% of the included diabetes trials reported patient-important outcomes. While there is ample room for improving the reporting of key methodological features (e.g., only one in five reports allocation concealment) and of adverse events in these trials, they follow more patients for a longer time and half of them lost <4% of patients to follow-up. This situation is encouraging because these trials play a major role in informing evidence-based clinical decision making and clinical policy.

There are, however, some other areas

of concern. First, it is worrisome that 17% of laboratory investigations report before-after comparisons within experimental and control groups but do not report the relevant comparison (i.e., between outcomes in the experimental and control groups). In these trials, the inferences drawn ignore the inferential strength that the randomization offers, thus foregoing the advantages associated with having a concomitant control group with similar prognosis. Second, diabetes trialists seem to favor hypotheses testing and reporting of *P* values rather than the more informative CIs. For clinicians, only knowing the answer to the test of a hypothesis (i.e., whether treatment changed the outcome) is often insufficient and could be misleading (24). It is often more informative to know the extent to which the treatment works (e.g., the size of the reduction in risk) and the precision around this estimate (i.e., the CI). For example, CIs can help readers determine whether a trial was not large enough to rule out important treatment effects when the association between intervention and outcomes is “not significant” (25). We believe these concerns should be corrected as a service to readers seeking to use RCT evidence in their clinical practice.

In conclusion, RCTs in diabetes published in pertinent top journals, both general and specialized, have important deficiencies in their report of key methodological features. These deficiencies are most remarkable in laboratory investigations with trials that measured patient-important outcomes showing better reporting with some room for improvement. To enhance the practice of evidence-based diabetes care, trialists need to pay closer attention to the rigorous implementation and reporting of important methodological safeguards against bias. Furthermore, clinicians and their patients with diabetes need more RCTs assessing the effect of promising interventions, including nonpharmacological interventions, on patient-important outcomes.

Acknowledgments—While conducting this work, V.M.M. was a Mayo Foundation Scholar and P.A.-C. held a postgraduate research fellowship at the Instituto Carlos III, Spanish Ministry of Health. P.A.-C.’s activities are supported in part by the Red Temática MBE (FIS G03/090).

References

1. Montori VM: Evidence-based endocrinology: how far have we come? (Review). *Treat Endocrinol* 3:1–10, 2004
2. Montori VM, Guyatt GH: What is evidence-based medicine? (Review). *Endocrinol Metab Clin North Am* 31:521–526, 2002
3. Montori VM: Evidence-based endocrine practice. *Endocr Pract* 9:321–323, 2003
4. Montori VM, Basu A, Erwin PJ, Velosa JA, Gabriel SE, Kudva YC: Posttransplantation diabetes: a systematic review of the literature (Review). *Diabetes Care* 25:583–592, 2002
5. Montori VM, Farmer A, Wollan PC, Dinneen SF: Fish oil supplementation in type 2 diabetes: a quantitative systematic review (Review). *Diabetes Care* 23:1407–1415, 2000
6. Montori VM, Helgemoe PK, Guyatt GH, Dean DS, Leung TW, Smith SA, Kudva YC: Telecare for patients with type 1 diabetes and inadequate glycemic control: a randomized controlled trial and meta-analysis. *Diabetes Care* 27:1088–1094, 2004
7. McIver B, Dinneen SF: An overview of randomized controlled trials in non-insulin dependent diabetes mellitus (Abstract). *Diabetologia* 39:A193, 1996
8. Guyatt G, Montori V, Devereaux PJ, Schunemann H, Bhandari M: Patients at the center: in our practice, and in our use of language (Editorial). *ACP J Club* 140:A11–A12, 2004
9. Mills E, Wu P, Gagnier J, Heels-Ansdell D, Montori VM: An analysis of general medical and specialist journals that endorse CONSORT found that reporting was not enforced consistently. *J Clin Epidemiol* 58:662–667, 2005
10. Moher D, Schulz KF, Altman D: The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *JAMA* 285:1987–1991, 2001
11. Egger M, Juni P, Bartlett C: Value of flow diagrams in reports of randomized controlled trials. *JAMA* 285:1996–1999, 2001
12. Bhandari M, Busse JW, Jackowski D, Montori VM, Schunemann H, Sprague S, Mears D, Schemitsch EH, Heels-Ansdell D, Devereaux PJ: Association between industry funding and statistically significant pro-industry findings in medical and surgical randomized trials. *CMAJ* 170:477–480, 2004
13. Pildal J, Chan AW, Hrobjartsson A, Forfang E, Altman DG, Gotzsche PC: Comparison of descriptions of allocation concealment in trial protocols and the published reports: cohort study (Review). *BMJ* 330:1049, 2005
14. Schulz KF, Altman DG, Moher D: Allocation concealment in clinical trials (Letter).

- JAMA 288:2406–2407, 2002 [author reply JAMA 288:2408–2409, 2002]
15. Schulz KF, Grimes DA: Allocation concealment in randomised trials: defending against deciphering (Review). *Lancet* 359: 614–618, 2002
 16. Devereaux PJ, Manns BJ, Ghali WA, Quan H, Lacchetti C, Montori VM, Bhandari M, Guyatt GH: Physician interpretations and textbook definitions of blinding terminology in randomized controlled trials. *JAMA* 285:2000–2003, 2001
 17. Montori VM, Bhandari M, Devereaux PJ, Manns BJ, Ghali WA, Guyatt GH: In the dark: the reporting of blinding status in randomized controlled trials. *J Clin Epidemiol* 55:787–790, 2002
 18. Hollis S, Campbell F: What is meant by intention to treat analysis? Survey of published randomised controlled trials (Review). *BMJ* 319:670–674, 1999
 19. Montori VM, Guyatt GH: Intention-to-treat principle (Review). *CMAJ* 165:1339–1341, 2001
 20. Ioannidis JP, Evans SJ, Gotzsche PC, O'Neill RT, Altman DG, Schulz K, Moher D: Better reporting of harms in randomized trials: an extension of the CONSORT statement. *Ann Intern Med* 141:781–788, 2004
 21. Devereaux PJ, Manns BJ, Ghali WA, Quan H, Guyatt GH: The reporting of methodological factors in randomized controlled trials and the association with a journal policy to promote adherence to the Consolidated Standards of Reporting Trials (CONSORT) checklist. *Control Clin Trials* 23:380–388, 2002
 22. Mills EJ, Wu P, Gagnier J, Devereaux PJ: The quality of randomized trial reporting in leading medical journals since the revised CONSORT statement. *Contemp Clin Trials* 26:480–487, 2005
 23. Devereaux PJ, Choi PT, El-Dika S, Bhandari M, Montori VM, Schunemann HJ, Garg AX, Busse JW, Heels-Ansdell D, Ghali WA, Manns BJ, Guyatt GH: An observational study found that authors of randomized controlled trials frequently use concealment of randomization and blinding, despite the failure to report these methods. *J Clin Epidemiol* 57:1232–1236, 2004
 24. Bhandari M, Montori VM, Schemitsch EH: The undue influence of significant p-values on the perceived importance of study results. *Acta Orthop* 76:291–295, 2005
 25. Montori VM, Kleinbart J, Newman TB, Keitz S, Wyer PC, Moyer V, Guyatt G: Tips for learners of evidence-based medicine. 2. Measures of precision (confidence intervals). *CMAJ* 171:611–615, 2004 [erratum in *CMAJ* 172:162, 2005]