

A systematic review of the quality of publications reporting coronary artery bypass grafting trials

Forough Farrokhyar, PhD;*† Rong Chu, BSc;‡ Richard Whitlock, MD;* Lehana Thabane, PhD†§

Objective: Several studies have shown that the quality of reports of randomized controlled trials (RCTs) in medicine is variable and often poor, whereas the quality of those in surgery is unknown. We aimed to assess the quality of reports of RCTs in coronary artery bypass grafting (CABG) surgery when comparing off- and on-pump techniques. **Methods:** From electronic searches of MEDLINE, the Cochrane Library, CINAHL, HealthSTAR and EMBASE, we identified RCTs published between 2000 and 2005 comparing off- and on-pump CABG. We assessed the report quality, using 35 items from the Consolidated Standards for Reporting Trials (CONSORT) statement and 54 additional indicators relevant to CABG surgery. Some of the indicators comprised several small parts, making the maximum possible total score 105. Two authors independently reviewed and assessed the reporting quality of each RCT. The level of agreement was assessed with kappa statistics, and disagreements were resolved by consensus. We expressed descriptive analyses as median and interquartile range; we used a generalized estimating equation (GEE) for data analysis. **Results:** We included 50 trials, for a total of 5134 patients. The kappa value was greater than 0.6 for 73 of 105 (70%) indicators. The overall report quality score varied from 35 to 93 of 105. The CONSORT score reporting quality varied from 16 to 39 of 42. The quality of reporting was poor and insufficient for the methods (particularly, the sample size, allocation and blinding subsections), results and discussion sections. With GEE modelling, the reporting quality had a strong association with trial size, publication year, trial location and funding source, but not with the results and type of primary outcome. **Conclusion:** The quality of the publications' reporting methods, results and discussion sections was suboptimal. It is critical that, in reporting surgical trials, authors follow the CONSORT guidelines as well as consider the surgical factors.

Objectif : Plusieurs études ont démontré que la qualité des rapports d'études contrôlées randomisées (ECR) en médecine est variable et souvent médiocre, tandis qu'on ne connaît pas celle des rapports d'ECR en chirurgie. Nous voulions évaluer la qualité des rapports d'ECR sur le pontage aortocoronarien (PAC) lorsque l'on compare des techniques sans pompe et avec pompe. **Méthodes :** Des recherches effectuées dans les banques de données MEDLINE, Cochrane Library, CINAHL, HealthSTAR et EMBASE nous ont permis de repérer des rapports publiés entre 2000 et 2005 portant sur des ECR au cours desquelles on a comparé le PAC sans pompe et avec pompe. Nous avons évalué la qualité du rapport en nous fondant sur 35 éléments de l'énoncé sur le regroupement des normes relatives aux rapports d'études (CONSORT) et sur 54 indicateurs supplémentaires pertinents au PAC. Certains des indicateurs comportaient plusieurs sous-éléments, ce qui a porté à 105 le score total maximum possible. Deux auteurs ont critiqué et évalué indépendamment la qualité du rapport de chaque ECR. On a évalué le niveau de convergence au moyen de statistiques kappa et résolu les divergences de vues par consensus. Nous avons exprimé les analyses descriptives sous forme d'intervalle médian et interquartile. Nous avons utilisé une équation d'estimation généralisée (EEG) pour analyser les données. **Résultats :** Nous avons inclus 50 études portant sur 5134 patients au total. La valeur kappa a dépassé 0,6 pour 73 indicateurs sur 105 (70 %). Le score global de la qualité des rapports a varié de 35 à 93 sur 105. Le score CONSORT de qualité des rapports a varié de 16 à 39 sur 42. La qualité des rapports était médiocre et insuffisante dans les sections portant sur les méthodes (en particulier les sous-sections sur la taille de l'échantillon, la répartition et le masquage), les résultats et la discussion. La modélisation EEG a révélé qu'il y

From the Departments of *Surgery, †Clinical Epidemiology and Biostatistics and ‡Computing and Software, McMaster University, and the §Centre for Evaluation of Medicines, St. Joseph's Healthcare, Hamilton, Ont.

Accepted for publication Sept. 25, 2006

Correspondence to: Dr. Lehana Thabane, Centre for Evaluation of Medicines, Father Sean O'Sullivan Research Centre, 105 Main St. E., Level P1, Hamilton ON L8N 1G6; fax 905 528-7386; thabanl@mcmaster.ca

avait un lien solide entre la qualité du rapport et l'ampleur de l'étude, l'année de publication, le lieu où se déroulait l'étude et la source de financement, mais non avec les résultats et le type de résultat principal. **Conclusion** : La qualité des sections des publications sur les méthodes, les résultats et la discussion n'est pas optimale. Il est crucial que dans leur rapport sur des études chirurgicales, les auteurs suivent les lignes directrices CONSORT et tiennent compte des facteurs chirurgicaux.

Coronary artery bypass grafting (CABG) plays a central role in the management of coronary artery disease. For more than 3 decades, surgeons have used cardiopulmonary bypass (CPB) to provide a still and bloodless field in which to accomplish optimal revascularization. The recent resurgence of operating on a beating heart, or off-pump, as an alternative procedure for coronary artery revascularization is intended to decrease the adverse events typically associated with CPB (on-pump CABG). Despite the growing body of literature, the results of off-pump CABG are controversial regarding its benefits or adverse effects, compared with on-pump CABG. Most of the available information on off-pump surgery is from observational studies. Currently, randomized controlled trials (RCTs) are applied to assess the superiority of off-pump CABG as a new technique against conventional on-pump CABG, but the results of these RCTs can only be understood if they are reported with high quality. The results of meta-analyses¹⁻³ performed on the available RCTs show no reduction in mortality but a significant reduction in several indicators of morbidities. Studies in other branches of medicine suggest that the quality of publications describing RCTs vary considerably in terms of their reporting quality,⁴ but no study has explicitly reviewed the quality of reports on RCTs in the surgical field.⁵⁻¹³ The poor quality of published information could be misleading, and it might introduce difficulties in decision making and during peer review, systematic review or meta-analysis. Hence, we aimed to systematically review the quality of published reports of RCTs comparing off- and on-pump CABG surgery. We discuss the results of this review, highlighting areas where improvements are needed and providing some direction on additional resources that can be used.

Methods

This systematic review was performed in accordance with a protocol that prescribed eligibility criteria, search strategy, outcomes and statistical analyses.

We performed an Ovid literature search to select potential reports according to the strategy for MEDLINE, in the Cochrane Reviewers' Handbook.¹⁴ We used similar strategies to search other databases, such as CINAHL, HealthSTAR and EMBASE. We also performed manual searches to identify articles missed in the computer-assisted searches. We limited the data searches to randomized controlled trials published in English between Jan. 1, 2000 and Dec. 31, 2005; subjects were aged 18 years and over. To refine the articles, we used the keywords "coronary artery bypass," "CABG," "off-pump," "beating heart," "OPCAB," "on-pump," "cardiopulmonary bypass," "conventional CABG," "ONCAB," "valve," "repair," "balloon," "stent," "PCI," "PTCA," "angioplasty" and "drug."

Eligibility criteria for study selection

Two authors independently identified trials for inclusion and extracted information on interventions and outcomes. A study was considered for inclusion if it was an RCT designed to compare off-pump CABG and on-pump CABG in adult populations, on the basis of at least 1 relevant clinical outcome. All studies were published in English and used human subjects. We included all RCTs with the primary objective of assessing a categorical or continuous variable as an outcome. For this reason, the number of trials in our study is different from the recently published meta-analysis.³ We excluded duplicate publications of previously

published trials on the same patients. We included substudies of larger trials with a predefined protocol, eligibility criteria and different sample size. The time period for the inclusion of published studies was Jan. 1, 2000 to Dec. 31, 2005. We excluded 5 RCTs before the year 2000 (1 from 1995, 1 from 1998 and 3 from 1999). We chose this time period to better assess the reporting quality of CABG publications, because the medical journals adopted the Consolidated Standards for Reporting Trials (CONSORT) statement by 1996, which was developed for clinical trials. The CONSORT statement would have had a very marginal effect in CABG surgical trials until the emergence of the revised CONSORT statement.^{4,15} We excluded studies that involved a combination of CABG surgery with another procedure, drug or device trials and subgroup analyses of an original study and those that only explained the rationale and design of the study.

Data extraction and trial assessment

The revised CONSORT statement¹⁵ has been used by many impact journals to criticize the quality of articles. We used a modified CONSORT statement⁴ and added the factors relevant to surgical trials and CABG surgery (Appendix 1). We created a data extraction form before data collection, and we used 89 indicators to assess the selected trials (35 indicators identified by the modified CONSORT statement and 54 indicators relevant to surgical trials or CABG surgery). Using a standardized data abstraction form, 2 authors independently reviewed the articles to assess the reporting quality of each RCT. The form was previously pilot tested and revised. Disagreements were re-

solved by consensus. One credit was awarded to each appropriately reported indicator, and no credit was assigned otherwise. Some of the indicators comprised several small parts. Therefore, the maximum possible score was greater than the number of indicators and varied from 0 to 42 for the CONSORT score and from 0 to 105 for the total score. We collected each reviewer's data set separately, using Microsoft Access (Microsoft Corporation, Redmond, Wash., 2000).

Definition of variables

When we found significant differences comparing off- and on-pump CABG, we defined the result of the primary outcomes as positive; otherwise, results were defined as negative. The trial size was decided based on the number of patients enrolled in each trial. The journal of publication refers to the journal in which the trial was published; the trial location refers to the location where the trial was conducted or, in the case of multicentre trials, to the location of the corresponding author. Type of primary outcome was defined as a categorical or continuous variable.

Data analysis

We performed Cohen's kappa analysis¹⁶ to measure the level of agreement between the reviewers on all items on the data abstraction form. A kappa of > 0.8 was considered good, 0.6–0.8 was substantial, 0.4–0.6 was moderate, 0.2–0.4 was fair and < 0.2 was poor. We performed 2 sets of analysis on extracted data, one on all 105 indicators and the other on the 42 CONSORT statement indicators as a sensitivity analysis. We reported categorical data as frequencies, percentages and 95% confidence intervals (CIs), using the Wilson Score method. Continuous data were reported as mean and standard deviation (SD), median, interquartile range (IQR) and minimum (min) and maximum (max) score. Tests of

associations between total reporting score and sectional reporting scores were performed using partial correlation coefficient when the trial size, publication year, primary outcome results, journal of publication, type of primary outcome, funding source and trial location were held constant. We performed univariate analysis and multivariable analysis, using a generalized estimating equation (GEE)¹⁷ to model the total quality score as a function of final sample size, publication year, funding source, trial location, and type and result of primary outcome. GEE was chosen to incorporate the possible correlation between quality scores of publications in the same journal. We expressed the GEE results as a coefficient, corresponding standard error (SE), 95% CI and associated *p* value. We included variables, in the multivariable analysis, with a *p* value of < 0.1 on univariate analysis. We assessed goodness of fit, using the likelihood ratio

statistic.¹⁸ All analyses were conducted with SPSS 15.0 and Stata 8.0.

Results

The MEDLINE database search identified 129 RCT articles published between Jan. 1, 2000 and Dec. 31, 2005. The combined search for additional databases yielded 114 trials for the same time period. After reviewing both databases, we identified 133 articles for systematic review. Another 4 articles were found by reviewing the references of the articles or meta-analyses. Of 137 articles, 50 trials fulfilled the eligibility criteria for this study (Fig. 1).^{19–68} Seven of 9 duplicates or nonoriginal publications were published in 2005. A total of 5134 patients were studied, and the trial enrolment varied from 20 to 401 patients. Table 1 summarizes the characteristics of these trials. Thirty-two (64%) of 50 trials were published in the surgical journals, 31 (62%) were completely or

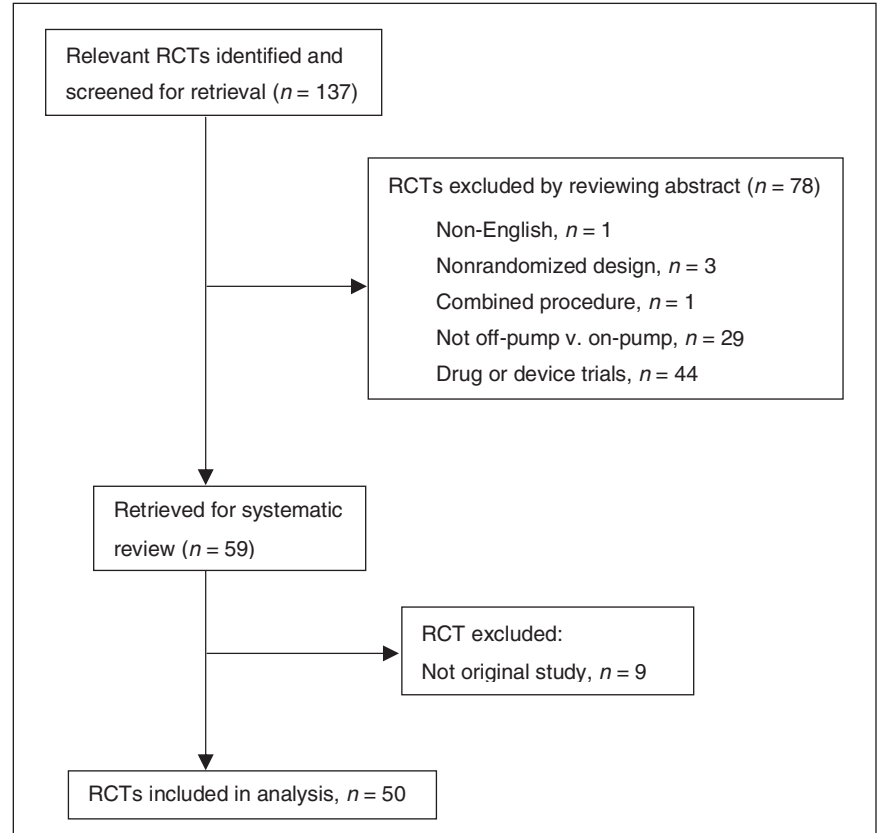


FIG. 1. Identification of eligible randomized controlled trials between Jan. 1, 2000 and Dec. 31, 2005.^{19–68}

partially funded by governmental agencies or foundations and 36 (72%) were conducted in Europe. The inter-reviewer level of agreement was good ($\kappa > 0.8$) for 47 (45%) and substantial ($\kappa = 0.6-0.8$) for 26 (25%) of 105 indicators; it was moderate ($\kappa = 0.4-0.6$)

for 17 (16%) and fair ($\kappa = 0.2-0.4$) for 15 (14%) indicators.

Table 2 summarizes the total scores, the CONSORT scores and the breakdown of scores for each section. The total report quality varied from 35 to 93 out of a possible score of 105. The report quality assessed by CONSORT criteria varied from 16 to 39 out of a possible score of 42. The report quality varied from 7 to 39 out of a possible score of 47 for methods sections and from 13 to 35 of a possible 39 for results sections. Within method subsections, sample size (min = 0, max = 11 of a possible score of 17), allocation or blinding (min = 0, max = 7 of a possible 7) and protocol (min = 1, max = 10) of a possible 10) had the least report quality score, respectively. There was a significant positive correlation between the total report quality score and report quality scores of methods (correlation coefficient [r] = 0.85, $p < 0.001$), results ($r = 0.38$, $p = 0.011$), and discussion ($r = 0.36$, $p = 0.018$) sections of the trials while controlling for trial size, publication year, primary outcome results, trial location, funding source and type of primary outcome. The report quality scores of components of methods sections were also significantly and positively correlated with the total score. The correlation coefficient was 0.49 ($p = 0.001$) for protocol, 0.49 ($p = 0.001$) for allocation and blinding, 0.54 for sample size calculation ($p < 0.001$) and 0.76 ($p < 0.001$) for statistical methods, respectively.

Table 3 presents the percentages of the trials that reported the CONSORT and surgical indicators adequately. The abstract or title described the study as an RCT, the introduction, description of study participants, surgical procedure, statistical methods of analysis for primary outcome, baseline data, results of primary outcomes and ancillary analyses were reported adequately in more than 90% of the trials. Indicators that were reported less frequently included the number of surgeons

(46%), the learning curve for off-pump technique (20%), sample size (28%), method of generating allocation sequence (30%), person or group that generated allocation sequence (10%), blinding (34%), intention-to-treat analysis (30%), the participants' flow chart for recruitment and follow-up (16%), the study period (66%) and the strengths and weaknesses of the findings (51%).

The detailed report quality for each item is presented in Appendix 1. An overall explanation of the study design was described adequately in only 24 trials (48%). In the protocol sections, 30% reported whether patients were consecutive, 60% reported the number of centres and 34% defined secondary outcomes. Although a sample size calculation was provided, there was little information on the desired significance level (28%), power (26%), the statement of smallest difference between groups (28%), the type of test used for primary outcomes (8%) or the method of analysis for the sample size calculation (4%). Other poorly presented indicators included handling missing information (40%), stating the method used in adjusted (40%) or subgroup (8%) analyses, and stating the software used for analysis (58%). Research Ethics Board approval and patient's informed consent were reported in more than 80% of the trials.

GEE modelling results

Univariate GEE analysis showed that the trial size, publication year, location of trial, funding source and type of primary outcome were significantly associated with the total quality of reporting score when we adjusted for the journal of publication (Table 4). The results of primary outcome (negative or positive) were not a significant factor in relation to the quality of reporting CABG trials. When we adjusted for the journal of publication, multivariable analysis of GEE revealed that the sample size of the trial was significantly and positively associated with

Table 1

Characteristics of included CABG RCTs ¹⁹⁻⁶⁸	
Trial features	Trials: no (and %) (n = 50)
Publication year	
2000	7 (14)
2001	8 (16)
2002	5 (10)
2003	11 (22)
2004	12 (24)
2005	7 (14)
Trial location	
North America	7 (14)
United Kingdom	13 (26)
Scandinavia	11 (22)
Other parts of Europe	12 (24)
Other countries	7 (14)
Publication journal	
<i>Annals of Thoracic Surgery</i>	16 (32)
<i>Journal of Thoracic and Cardiovascular Surgery</i>	6 (12)
<i>European Journal of Cardiothoracic Surgery</i>	6 (12)
<i>Heart Forum Surgery</i>	4 (8)
<i>Circulation</i>	5 (10)
<i>Journal of the American Medical Association</i>	2 (4)
<i>New England Journal of Medicine</i>	2 (4)
<i>Lancet</i>	1 (2)
10 other journals	8 (16)
Source of funding	
Government, foundation or both	26 (52)
Partial government, foundation or both	5 (10)
Companies	1 (2)
Not reported	18 (36)
Results of primary outcome	
Positive	22 (44)
Negative	28 (56)
Type of primary outcome	
Categorical	17 (34)
Continuous	33 (66)
CABG = coronary artery bypass grafting; RCT = randomized controlled trial.	

the quality of reporting (Table 5). With respect to the publication year, the report quality of publications for CABG surgery was significantly and positively improved in 2002, 2004 and 2005, compared with 2000, and this improvement was greatest in 2005. The report quality in 2001 and 2003 was similar to that in 2000. With respect to the trial location, the quality of reporting was similar between North America and the United Kingdom, but that of CABG trials was significantly lower for Scandinavia and other countries. The funding source and type of primary outcome (continuous v. categorical) were not significantly associated with the quality of reporting. Most trials with a continuous variable as the primary outcome had positive results (66%), compared with those with categorical data as the primary outcome (34%).

Discussion

We comprehensively searched the databases to identify RCT articles comparing any type of outcome between off- and on-pump CABG surgery. Only 50 original RCT articles published between Jan. 1, 2000 and Dec. 31, 2005 met our inclusion

criteria. The total reporting quality of trials comparing off- and on-pump CABG surgery in this systematic review varied substantially between publications (35–93 of a possible score of 105). The results of this systematic review showed that there is room for improvement. The reporting quality of the methods was poor, with insufficient and inadequate information. Information on the methods and study design, allocation sequence (30%), blinding mechanism (34%), sample size calculation (28%), intention-to-treat approach (30%), number of surgeons in the study (46%), learning curve and surgeon's experience with the off-pump procedure (20%) was minimal in most of the published trials. Although all of the trials reported the number of patients (sample size) in the final reports, the details on sample size calculation were lacking from most of the studies. Because the present study only reviewed trials published in peer-reviewed journals, we expected that the quality of reports on CABG trials (when including non-peer-reviewed reports) would be even lower. The publication year and the size and location of the trials were significantly associated with the

reporting quality of the CABG trial findings. The reasons for improving the reporting quality of CABG trials in recent years are as follows: 1) authors of surgical trials are gaining more experience and skill at conducting RCTs and writing manuscripts and 2) referees and journals are demanding better quality in reports of RCTs, since medical journals adopted the CONSORT statement. However, these improvements are not as evident in reporting indicators relevant in surgical trials.

In the hierarchy of evidence, RCTs represent the highest quality design for evaluating clinical practice. They are most valuable when conducted with precision and when the results are adequately reported. The main reason for undertaking RCTs is to inform and alter medical practice. Without adequate information on allocation, randomization or blinding, the validity of a trial becomes questionable. Until recently, trials of surgical methods have been uncommon, and the decision making on surgical methods has been modelled on the experience gained from observational studies on a small number of patients. With the increasing popularity of RCTs in surgery, several

Table 2

Report quality scores by report section and CONSORT statement

Section	Maximum possible	Maximum obtained	Median (Q1, Q3)	Minimum obtained	Mean	SD
Title	1	1	1 (0, 1)	0	0.5	0.5
Abstract	6	5	3 (3, 3)	1	2.9	0.7
Introduction	3	3	3 (3, 3)	2	2.9	0.2
Methods	47	39*	16 (13, 24)	7	17.2	8.1
Protocol	10	10	6 (5, 8)	1	6.3	1.9
Sample size	17	11	0 (0, 6)	0	2.5	4.2
Allocation/blinding	7	7	3 (2, 3)	0	2.8	1.7
Statistical methods	13	13	6 (5, 7)	4	6.6	2.0
Results	39	35*	23 (19, 26)	13	23.0	5.0
Baseline data	21	17	10 (7, 12)	3	9.3	3.5
Outcomes	18	18	14 (12, 15)	7	14.0	2.5
Discussion	9	9	5 (4, 6)	2	4.5	1.4
CONSORT	42	39	26 (23, 30)	16	26.4	4.9
Total	105	93	53 (44, 61)	35	51.7	11.5

CONSORT = Consolidated Standards for Reporting Trials; Q1 = first quartile; Q3 = third quartile; SD = standard deviation.

*The maximum and minimum obtained scores do not necessarily add up to the total maximum and minimum obtained scores.

pitfalls related to their conduct may introduce significant biases.

There are different types of RCTs that are applicable in surgery patients.⁶⁹ Type I studies the superiority of a new surgical procedure versus the conventional one. Type II compares a surgical technique to a laparoscopic method or to medical treatment. Type III does not study surgical methods but, instead, relates to such medical treatments as pain relief and prophylaxis (i.e., infection, thrombosis, bleeding) in surgical pa-

tients.⁷⁰ Although most of the difficulties apply to all 3 types, the first 2 require a more challenging design process. This systematic review focuses on type I studies.

A surgical intervention can be distinguished from a drug intervention by the skill required to administer the treatment. Drug trials do not risk any differential skills in administering an active medication versus placebo to patients; however, surgery is a skilled, complex and multistep process, which makes the design of RCTs

problematic in surgical studies for several reasons.⁷⁰⁻⁷² Training experience is required to develop expertise in a surgical technique. There is a learning process in every new surgical technique, and the quality of performance improves with frequent repetition over time.⁷⁰ This learning process most likely varies from surgeon to surgeon. To avoid serious bias, the learning curve for the new procedure needs to be acknowledged and controlled at the design stage or evaluated at the analysis stage.⁷⁰ In a surgical trial, if a participating surgeon is performing the new technique as well as the conventional technique and has restricted expertise with the former, the results will be biased in favour of the conventional technique. Devereaux and colleagues⁷³ refer to this as differential bias. Surgeon's skill variation is another issue that poses a problem in conducting surgical trials. There is an inherent variation in different surgeons performing the same procedure.⁷⁰ The effect of trainees, fellows and other surgical teams in the operating room (OR) makes the problem even more profound. The recruitment time is a challenging issue in surgical trials, especially in rare conditions. It is difficult to recruit when patient accrual time is long or when the numbers needed to achieve an effect is large.⁶⁹ Another concern is that patients who have been recruited might become ineligible by the time they have surgery, because of long waiting lists. The classic method preferred to avoid outcome bias in clinical trials is double blinding, wherein neither the patients nor the investigators know which treatment the patient receives. In surgical RCTs, surgeons cannot be blinded to the type of surgery they are performing; patients usually know the type of surgery they are having because it is unethical and difficult for the surgeon not to tell them.⁶⁹ The blinded approach would be impossible when a surgical technique is being compared with a nonsurgical one. The best approach is to conduct a single-blind

Table 3

Overall quality of reporting using CONSORT and surgery-based items

Criteria	Description	% reported (and 95% CI)
Title and abstract	Title or abstract identified the study as a randomized trial	98 (89 to 99)
Introduction	Scientific background and explanation of rationale	98 (89 to 99)
Methods		
Participants	Study population and eligibility criteria described	94 (83 to 98)
Surgeons	Number provided	46 (33 to 60)
Learning curve	Learning curve for new intervention described	20 (21 to 46)
Intervention	Surgical technique for each group described	88 (76 to 94)
Outcomes	Primary outcome measures described	74 (60 to 84)
Sample size	Calculated	28 (17 to 42)
Randomization		
Allocation	Clear description of method	30 (18 to 45)
Implementation	Group or person who generated the random sequence provided	10 (4 to 21)
Blinding	Mechanism described	34 (22 to 48)
Intention-to-treat	Statement of analysis provided	30 (18 to 45)
Statistical method	Detailed method of analysis for primary outcomes, or subgroup and adjusted analysis	96 (86 to 99)
Results		
Participant flow chart	Flow chart of participants' recruitment and follow-up process, i.e., no. who were randomized, received treatment, completed protocol and were analyzed	16 (8 to 28)
Recruitment	Dates define the period of recruitment and follow-up	66 (58 to 82)
Baseline data	Outline of baseline demographic and clinical characteristics of each group	98 (89 to 99)
Outcomes	Clear summary results with measures of precision for the primary outcome	100 (93 to 100)
Ancillary analyses	Clear statement of subgroup and/or adjusted analyses were prespecified or explanatory	91 (81 to 97)
Discussion	Strengths and weaknesses of the findings	51 (37 to 63)

CONSORT = Consolidated Standards for Reporting Trials; CI = confidence interval.

study in which the people who assess the outcome and the data analysts are unaware of the patient's treatment. This is also difficult to guarantee because the assessor is usually the operating surgeon. It becomes difficult to get reliable results if a long time was required to observe the end points. It is often difficult to track a group of patients for a long time or to keep them interested in clinical visits if no treatment or minimal treatment is given after surgical intervention.

Within the aspects of study design, the lack of reporting adequate information on the timing of the trial, the definition of study population, the learning curve for the new procedure, the technical skill and ability of the surgeon(s), a trainee's assistance in the OR, variation in surgeons' experience with the surgical technique, allocation sequence and

concealment, blinding mechanism, intention-to-treat analysis and recruitment time period all encounter certain barriers that undermine the validity, applicability and ethical integrity of RCTs in surgery. Surgical factors that are not incorporated in the CONSORT statement are very important and should be regarded with significance in surgical trials. Researchers and surgeons are encouraged to consider these factors when designing their trials and when reporting them in their publications. In the case of impracticality or lack of feasibility, the limitations inherent in their approach should be discussed in detail to allow readers to interpret the use of the findings. While our review did not investigate the accuracy of reporting statistical aspects, there were some instances where common errors, such as lack of reporting

goodness of fit and assessment of model assumptions, were apparent. As part of the user's guide to the surgical literature series, Thoma and colleagues⁷⁴ critically reviewed a surgical RCT in general surgery and reached a similar conclusion.

Strengths and limitations

Due to the rigor of this systematic review, we are confident that it is a complete summary of the available evidence. There are, however, several limitations to our study. First, our study evaluated only the reporting quality of the trials and not the quality of the trials themselves. Soares and colleagues⁸ evaluated the methodology of the RCTs by the Radiation Therapy Oncology Group and concluded that "poor reporting of RCTs may not indicate poor quality of the

Table 4

Univariate GEE estimates for report quality in 50 trials adjusted for the journal of publication

Characteristics	Median score (Q1, Q3)	Coefficient (and 95% CI)	Corresponding SE	p value
Trial size	60 (37, 140)	0.07 (0.03 to 0.11)	0.02	0.001
Year of publication				
2000 (reference) *	41 (40, 50)	—	—	—
2001	45 (40, 58)	1.80 (-0.43 to 4.02)	1.13	0.114
2002	57 (54, 77)	17.50 (4.96 to 30.04)	6.40	0.006
2003	51 (46, 56)	8.58 (4.58 to 12.57)	2.03	< 0.001
2004	53 (47, 57)	6.96 (5.53 to 8.38)	0.73	< 0.001
2005	61 (59, 67)	19.47 (13.43 to 25.51)	3.08	< 0.001
Trial location				
North America (reference) *	56 (49, 70)	—	—	—
United Kingdom	56 (48, 57)	-2.01 (-9.55 to 5.52)	3.84	0.600
Scandinavia	55 (46, 66)	-2.86 (-4.93 to -0.78)	1.06	0.007
Others	47 (42, 57)	-7.16 (-14.86 to 0.54)	2.71	0.068
Source of funding				
Government or foundations (reference) * †	57 (47, 66)	—	—	—
Companies/not reported‡	47 (41, 52)	-8.12 (-12.21 to -4.03)	2.08	< 0.001
Results of primary outcome				
Negative (reference) *	56 (47, 69)	—	—	—
Positive	50 (43, 57)	4.83 (-2.44 to 12.11)	3.71	0.193
Type of primary outcome				
Continuous (reference) *	49 (41, 57)	—	—	—
Categorical	57 (51, 69)	5.30 (3.11 to 7.50)	1.12	< 0.001

GEE = generalized estimating equation; Q1 = first quartile; Q3 = third quartile; CI = confidence interval; SE = standard error.
 *Other categories were compared to this category.
 †Completely or partially supported by government or foundations.
 ‡Only 1 was funded by a company.

trials themselves.” Because only published reports are available to readers, researchers should be encouraged to publish manuscripts that contain more details of experimental methods and to discuss the strengths and weaknesses of their protocol in greater detail. Several indicators of our instrument are subjective, for example, “Is the study design well explained?” Because indicators are recommended by the CONSORT statement and are necessary to report in clinical trials, we kept them and examined their accuracy by interreviewer analysis. Criteria for both the CONSORT statement and extended variables relevant to surgical RCTs or CABG surgery were not weighted for scoring purposes, based on the assumption that all of the indicators have equal importance. We avoided weighting because it would have been arbitrary and subjective and thus subject to criticism.⁴ This means that the CONSORT and

surgical indicators, such as the learning curve for the off-pump technique, number of participating surgeons, sample size calculation, allocation and blinding, and title identifying the study as an RCT, are treated equally. To the best of our knowledge, we included all known baseline risk factors in the protocol; these depend on the patient population under study, which may vary between trials. To avoid underscoring bias, we included the risk factors in the inclusion and exclusion criteria. It is, however, possible that the baseline data section in some trials was underscoring. These trials might have included a healthier patient population and did not need to report some of the risk factors. Finally, the modified CONSORT statement introduced in this manuscript and its scoring system are not validated, nor is there any validated tool to evaluate the reporting quality of the RCTs reported in the field of surgery.

Conclusion

Our findings are consistent with other studies on the quality of RCT publications in clinical trials,^{4,75-77} although no studies have focused on surgical trials. The published CABG RCTs lack good reporting quality, especially with respect to the methods and discussion sections. The results of this review should strongly encourage journal editors to change the instructions to authors to ensure that the issues that affect the understanding of a manuscript, whether by referees or readers, and how the study was undertaken are adequately described. The implementation of CONSORT statements and surgical indicators by referees in reviewing surgical trials would facilitate peer review and enhance the scientific quality of the data retrieved from the publications. The CONSORT statements provide trialists, referees and editors with guidelines for improving the

Table 5

Multivariable GEE estimates for report quality in 50 trials adjusted for the journal of publication

Characteristics	Median score (Q1, Q3)	Coefficient (and 95% CI)	Corresponding SE	p value
Trial size	60 (37, 140)	0.08 (0.03 to 0.12)	0.02	0.001
Year of publication				
2000 (reference) *	41 (40, 50)	—	—	—
2001	45 (40, 58)	-0.71 (-6.91 to 5.50)	3.16	0.822
2002	57 (54, 77)	8.30 (6.82 to 9.78)	0.75	< 0.001
2003	51 (46, 56)	3.29 (-1.79 to 8.34)	2.59	0.204
2004	53 (47, 57)	4.33 (2.71 to 5.94)	0.82	< 0.001
2005	61 (59, 67)	20.94 (14.57 to 27.30)	3.24	< 0.001
Trial location				
North America (reference) *	56 (49, 70)	—	—	—
United Kingdom	56 (48, 57)	-0.33 (-5.52 to 4.86)	2.64	0.900
Scandinavia	55 (46, 66)	-5.28 (-8.66 to -1.90)	1.72	0.002
Others	47 (42, 57)	-9.74 (-16.18 to -3.30)	3.29	0.003
Source of funding				
Government/foundations (reference)*†	57 (47, 66)	—	—	—
Companies/not reported‡	47 (41, 52)	-0.85 (-5.97 to 4.26)	2.61	0.743
Type of primary outcome				
Continuous (reference) *	49 (41, 57)	—	—	—
Categorical	57 (51, 69)	-0.23 (-5.55 to 5.09)	2.71	0.932

GEE = generalized estimating equation; Q1 = first quartile; Q3 = third quartile; CI = confidence interval; SE = standard error; — = there is no coefficient for this category because it is the reference category.

*Other categories are compared to this category.

†Completely or partially supported by government or foundations.

‡Only 1 was funded by a company.

quality of RCT reports, but they do not cover the issues related to the surgical trials. Therefore, authors, editors and referees are encouraged to consider the CONSORT statements and factors relevant to the surgical trials described in this review. These criteria are likely generalizable to other fields of surgery, when an investigator is conducting an RCT while adjusting for baseline risk factors. To embrace this tool and its scoring system as a practical tool in the evaluation of surgical RCTs, it needs to be established and validated in the different fields of surgery.

Acknowledgements: Thanks to Susan Tomlinson for proofreading the manuscript. The study was supported in part by a grant from Human Resources Development Canada.

Competing interests: None declared.

Contributors: Drs. Farrokhyar and Thabane designed the study. Drs. Farrokhyar and Whitlock and Ms. Chu acquired the data, which Drs. Farrokhyar and Thabane and Ms. Chu analyzed. Dr. Farrokhyar wrote the article, and all authors reviewed it. All authors gave final approval for the article to be published.

References

- Parolari A, Alamanni F, Cannata A, et al. Off-pump versus on-pump coronary artery bypass: meta-analysis of currently available randomized trials. *Ann Thorac Surg* 2003; 76:37-40.
- van der Heijden GJ, Nathoe HM, Jansen EW, et al. Meta-analysis on the effect of off-pump coronary bypass surgery. *Eur J Cardiothorac Surg* 2004;26:81-4.
- Cheng DC, Bainbridge D, Martin JE, et al. Does off-pump coronary artery bypass reduce mortality, morbidity, and resource utilization when compared with conventional coronary artery bypass? A meta-analysis of randomized trials. *Anesthesiology* 2005;102:188-203.
- Bath FJ, Owen VE, Bath PM. Quality of full and final publications reporting acute stroke trials: a systematic review. *Stroke* 1998;29:2203-10.
- Chan KB, Man-Son-Hing M, Molnar FJ, et al. How well is the clinical importance of study results reported? An assessment of randomized controlled trials. *CMAJ* 2001;165:1197-202.
- Ko CY, Sack J, Chang JT, et al. Reporting randomized, controlled trials: where quality of reporting may be improved. *Dis Colon Rectum* 2002;45:443-7.
- Thakur A, Wang EC, Chiu TT, et al. Methodology standards associated with quality reporting in clinical studies in pediatric surgery journals. *J Pediatr Surg* 2001;36:1160-4.
- Soares HP, Daniels S, Kumar A, et al. Bad reporting does not mean bad methods for randomised trials: observational study of randomised controlled trials performed by the Radiation Therapy Oncology Group. *BMJ* 2004;328:22-4.
- Hebert RS, Wright SM, Dittus RS, et al. Prominent medical journals often provide insufficient information to assess the validity of studies with negative results. *J Negat Results Biomed* 2002;1:1.
- Altman DG, Moher D, Schulz KF. Peer review of statistics in medical research. Reporting power calculations is important. *BMJ* 2002;325:491.
- Moher D, Dulberg CS, Wells GA. Statistical power, sample size, and their reporting in randomized controlled trials. *JAMA* 1994;272:122-4.
- 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.
- Mills E, Loke YK, Wu P, et al. Determining the reporting quality of RCTs in clinical pharmacology. *Br J Clin Pharmacol* 2004;58:61-5.
- Alderson P, Green S, Higgins J. Cochrane Reviewers' Handbook 4.2.2. In: Alderson P, Green S, Higgins J, editors. *The Cochrane Library, Issue 1*. Chichester (UK): John Wiley & Sons Ltd; 2004.
- Altman DG, Schulz KF, Moher D, et al. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med* 2001;134: 663-94.
- Landis JR, Koch GG. An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics* 1977;33:363-74.
- Zeger SL, Liang KY. Longitudinal data analysis for discrete and continuous outcomes. *Biometrics* 1986;42:121-30.
- Horton NJ, Bebchuk JD, Jones CL, et al. Goodness-of-fit for GEE: an example with mental health service utilization. *Stat Med* 1999;18:213-22.
- Wan IY, Arifi AA, Wan S, et al. Beating heart revascularization with or without cardiopulmonary bypass: evaluation of inflammatory response in a prospective randomized study. *J Thorac Cardiovasc Surg* 2004;127:1624-31.
- Legare JF, Buth KJ, King S, et al. Coronary bypass surgery performed off pump does not result in lower in-hospital morbidity than coronary artery bypass grafting performed on pump. *Circulation* 2004; 109:887-92.
- Ahonen J, Sahlman A, Yli-Hankala A, et al. No effect of cardiopulmonary bypass on hypnosis in patients anaesthetized with propofol and alfentanil. *Br J Anaesth* 2004;92:137-9.
- Gerola LR, Buffolo E, Esbik W, et al. Off-pump versus on-pump myocardial revascularization in low-risk patients with one or two vessel disease: perioperative results in a multicenter randomized controlled trial. *Ann Thorac Surg* 2004;77:569-73.
- Straka Z, Widimsky P, Jirasek K, et al. Off-pump versus on-pump coronary surgery: final results from a prospective randomized study PRAGUE-4. *Ann Thorac Surg* 2004;77:789-93.
- Khan NE, De Souza A, Mister R, et al. A randomized comparison of off-pump and on-pump multivessel coronary-artery bypass surgery. *N Engl J Med* 2004;350:21-8.
- Dorman BH, Kratz JM, Multani MM, et al. A prospective, randomized study of endothelin and postoperative recovery in off-pump versus conventional coronary artery bypass surgery. *J Cardiothorac Vasc Anesth* 2004;18:25-9.
- Alwan K, Falcoz PE, Alwan J, et al. Beating versus arrested heart coronary revascularization: evaluation by cardiac troponin I release. *Ann Thorac Surg* 2004;77:2051-5.
- Puskas JD, Williams WH, Mahoney EM, et al. Off-pump vs conventional coronary artery bypass grafting: early and 1-year graft patency, cost, and quality-of-life outcomes: a randomized trial. *JAMA* 2004; 291:1841-9.
- Gulielmos V, Menschikowski M, Dill H, et al. Interleukin-1, interleukin-6 and myocardial enzyme response after coronary artery bypass grafting—a prospective randomized comparison of the conventional and three minimally invasive surgical techniques. *Eur J Cardiothorac Surg* 2000; 18:594-601.
- Baker RA, Andrew MJ, Ross IK, et al. The Octopus II stabilizing system: biochemical and neuropsychological outcomes in coronary artery bypass surgery. *Heart Surg Forum* 2001;4(Suppl 1):S19-S23.
- Muneretto C, Bisleri G, Negri A, et al. Off-pump coronary artery bypass surgery technique for total arterial myocardial revascularization: a prospective randomized study. *Ann Thorac Surg* 2003;76:778-82.
- Lund C, Hol PK, Lundblad R, et al. Comparison of cerebral embolization during off-pump and on-pump coronary artery bypass surgery. *Ann Thorac Surg* 2003;76:765-70.
- Parolari A, Alamanni F, Juliano G, et al. Oxygen metabolism during and after cardiac surgery: role of CPB. *Ann Thorac Surg* 2003;76:737-43.
- Lee JD, Lee SJ, Tsushima WT, et al. Ben-

- efits of off-pump bypass on neurologic and clinical morbidity: a prospective randomized trial. *Ann Thorac Surg* 2003; 76:18-25.
34. Penttila HJ, Lepojarvi MV, Kiviluoma KT, et al. Myocardial preservation during coronary surgery with and without cardiopulmonary bypass. *Ann Thorac Surg* 2001; 71:565-71.
 35. Guler M, Kirali K, Toker ME, et al. Different CABG methods in patients with chronic obstructive pulmonary disease. *Ann Thorac Surg* 2001;71:152-7.
 36. Wandschneider W, Thalmann M, Trampitsch E, et al. Off-pump coronary bypass operations significantly reduce S100 release: an indicator for less cerebral damage? *Ann Thorac Surg* 2000;70:1577-9.
 37. Lingaas PS. Clinical and angiographic outcome of coronary surgery with and without cardiopulmonary bypass: a prospective randomized trial. *Heart Surg Forum* 2004;7:37-41.
 38. Ascione R, Lloyd CT, Underwood MJ, et al. Inflammatory response after coronary revascularization with or without cardiopulmonary bypass. *Ann Thorac Surg* 2000;69:1198-204.
 39. Sahlman A, Ahonen J, Nemlander A, et al. Myocardial metabolism on off-pump surgery; a randomized study of 50 cases. *Scand Cardiovasc J* 2003;37:211-5.
 40. Matata BM, Sosnowski AW, Galinanes M. Off-pump bypass graft operation significantly reduces oxidative stress and inflammation. *Ann Thorac Surg* 2000;69:785-91.
 41. Wehlin L, Vedin J, Vaage J, et al. Activation of complement and leukocyte receptors during on- and off pump coronary artery bypass surgery. *Eur J Cardiothorac Surg* 2004;25:35-42.
 42. Wildhirt SM, Schulze C, Schulz C, et al. Reduction of systemic and cardiac adhesion molecule expression after off-pump versus conventional coronary artery bypass grafting. *Shock* 2001;16(Suppl 1):55-9.
 43. Motallebzadeh R, Kanagasabay R, Bland M, et al. S100 protein and its relation to cerebral microemboli in on-pump and off-pump coronary artery bypass surgery. *Eur J Cardiothorac Surg* 2004;25:409-14.
 44. Selvanayagam JB, Petersen SE, Francis JM, et al. Effects of off-pump versus on-pump coronary surgery on reversible and irreversible myocardial injury: a randomized trial using cardiovascular magnetic resonance imaging and biochemical markers. *Circulation* 2004;109:345-50.
 45. Puskas JD, Williams WH, Duke PG, et al. Off-pump coronary artery bypass grafting provides complete revascularization with reduced myocardial injury, transfusion requirements, and length of stay: a prospective randomized comparison of two hundred unselected patients undergoing off-pump versus conventional coronary artery bypass grafting. *J Thorac Cardiovasc Surg* 2003;125:797-808.
 46. Nathoe HM, van Dijk D, Jansen EW, et al. A comparison of on-pump and off-pump coronary bypass surgery in low-risk patients. *N Engl J Med* 2003;348:394-402.
 47. Zamvar V, Williams D, Hall J, et al. Assessment of neurocognitive impairment after off-pump and on-pump techniques for coronary artery bypass graft surgery: prospective randomised controlled trial. *BMJ* 2002;325:1268.
 48. Al Ruzzeq S, Hoare G, Marczin N, et al. Off-pump coronary artery bypass surgery is associated with reduced neutrophil activation as measured by the expression of CD11b: a prospective randomized study. *Heart Surg Forum* 2003;6:89-93.
 49. Tang AT, Knott J, Nanson J, et al. A prospective randomized study to evaluate the renoprotective action of beating heart coronary surgery in low risk patients. *Eur J Cardiothorac Surg* 2002;22:118-23.
 50. Carrier M, Perrault LP, Jeanmart H, et al. Randomized trial comparing off-pump to on-pump coronary artery bypass grafting in high-risk patients. *Heart Surg Forum* 2003;6:E89-92.
 51. van Dijk D, Jansen EW, Hijman R, et al. Cognitive outcome after off-pump and on-pump coronary artery bypass graft surgery: a randomized trial. *JAMA* 2002;287:1405-12.
 52. Lloyd CT, Ascione R, Underwood MJ, et al. Serum S-100 protein release and neuropsychologic outcome during coronary revascularization on the beating heart: a prospective randomized study. *J Thorac Cardiovasc Surg* 2000;119:148-54.
 53. van Dijk D, Nierich AP, Jansen EW, et al. Early outcome after off-pump versus on-pump coronary bypass surgery: results from a randomized study. *Circulation* 2001;104:1761-6.
 54. Czerny M, Baumer H, Kilo J, et al. Complete revascularization in coronary artery bypass grafting with and without cardiopulmonary bypass. *Ann Thorac Surg* 2001;71:165-9.
 55. Ascione R, Caputo M, Calori G, et al. Predictors of atrial fibrillation after conventional and beating heart coronary surgery: A prospective, randomized study. *Circulation* 2000;102:1530-5.
 56. Diegeler A, Hirsch R, Schneider F, et al. Neuromonitoring and neurocognitive outcome in off-pump versus conventional coronary bypass operation. *Ann Thorac Surg* 2000;69:1162-6.
 57. Velissaris T, Tang AT, Murray M, et al. A prospective randomized study to evaluate stress response during beating-heart and conventional coronary revascularization. *Ann Thorac Surg* 2004;78:506-12.
 58. Ascione R, Williams S, Lloyd CT, et al. Reduced postoperative blood loss and transfusion requirement after beating-heart coronary operations: a prospective randomized study. *J Thorac Cardiovasc Surg* 2001;121:689-96.
 59. Caputo M, Yeatman M, Narayan P, et al. Effect of off-pump coronary surgery with right ventricular assist device on organ function and inflammatory response: a randomized controlled trial. *Ann Thorac Surg* 2002;74:2088-95.
 60. Rankin KP, Kochamba GS, Boone KB, et al. Presurgical cognitive deficits in patients receiving coronary artery bypass graft surgery. *J Int Neuropsychol Soc* 2003;9: 913-24.
 61. Celik JB, Gormus N, Topal A, et al. Effect of off-pump and on-pump coronary artery bypass grafting on renal function. *Ren Fail* 2005;27:183-8.
 62. Lund C, Sundet K, Tennoe B, et al. Cerebral ischemic injury and cognitive impairment after off-pump and on-pump coronary artery bypass grafting surgery. *Ann Thorac Surg* 2005;80:2126-31.
 63. Wehlin L, Vedin J, Vaage J, et al. Peripheral blood monocyte activation during coronary artery bypass grafting with or without cardiopulmonary bypass. *Scand Cardiovasc J* 2005;39:78-86.
 64. Parolari A, Mussoni L, Frigerio M, et al. Increased prothrombotic state lasting as long as one month after on-pump and off-pump coronary surgery. *J Thorac Cardiovasc Surg* 2005;130:303-8.
 65. Blacher C, Neumann J, Jung LA, et al. Off-pump coronary artery bypass grafting does not reduce lymphocyte activation. *Int J Cardiol* 2005;101:473-9.
 66. Rastan AJ, Bittner HB, Gummert JF, et al. On-pump beating heart versus off-pump coronary artery bypass surgery-evidence of pump-induced myocardial injury. *Eur J Cardiothorac Surg* 2005;27:1057-64.
 67. Angelini GD, Taylor FC, Reeves BC, et al. Early and midterm outcome after off-pump and on-pump surgery in Beating Heart Against Cardioplegic Arrest Studies (BHACAS 1 and 2): a pooled analysis of two randomised controlled trials. *Lancet* 2002;359:1194-9.
 68. Kobayashi J, Tashiro T, Ochi M, et al. Early outcome of a randomized comparison of off-pump and on-pump multiple arterial coronary revascularization. *Circulation* 2005;112(9 Suppl):I338-I343.
 69. Carey TS. Randomized controlled trials in surgery: an essential component of scientific progress. *Spine* 1999;24:2553-5.
 70. McCulloch P, Taylor I, Sasako M, et al. Randomised trials in surgery: problems and possible solutions. *BMJ* 2002;324: 1448-51.
 71. McLeod RS, Wright JG, Solomon MJ, et al. Randomized controlled trials in surgery: issues and problems. *Surgery* 1996;119:483-6.

72. Solomon MJ, McLeod RS. Surgery and the randomised controlled trial: past, present and future. *Med J Aust* 1998;169:380-3.
73. Devereaux PJ, Bhandari M, Clarke M, et al. Need for expertise based randomised controlled trials. *BMJ* 2005;330:88.
74. Thoma A, Farrokhyar F, Bhandari M, et al. Users' guide to the surgical literature. How to assess a randomized controlled trial in surgery. *Can J Surg* 2004;47:200-8.
75. Nicolucci A, Grilli R, Alexanian AA, et al. Quality, evolution, and clinical implications of randomized, controlled trials on the treatment of lung cancer. A lost opportunity for meta-analysis. *JAMA* 1989;262:2101-7.
76. Liberati A, Himel HN, Chalmers TC. A quality assessment of randomized control trials of primary treatment of breast cancer. *J Clin Oncol* 1986;4:942-51.
77. Gotzsche PC. Methodology and overt and hidden bias in reports of 196 double-blind trials of nonsteroidal antiinflammatory drugs in rheumatoid arthritis. *Control Clin Trials* 1989;10:31-56.

Appendix 1: Criteria of report quality of 54 indicators relevant to CABG and surgical trials and 35 from the CONSORT statement (modified from Bath and colleagues⁴) n = 50

Trials	Criteria provided	% reported
	Title	
1	Identifies study as an RCT*	50
	Abstract	
2	Structured format*	94
3	No. of patients	96
4	States whether analysis was by intention-to-treat	4
5	Hypothesis (or rationale)	32
6	Method of randomization	2
7	Follow-up period	64
	Introduction	
8	Study background*	100
9	Study rationale*	96
10	Clinical objectives*	98
	Methods	
	Protocol	
11	Intended study population*	96
12	Inclusion criteria for participants*	92
13	Exclusion criteria for participants*	94
14	Identifies whether patients were consecutive	30
15	States no. of centres	60
16	States no. of surgeons	46
17	Defines surgical intervention*	88
18	Learning curve for off-pump technique justified	20
19	Primary outcome measure(s)*	74
20	Secondary outcome measure(s)*	34
	Sample size	
21	Calculation	28
22	Calculation for multiple primary outcomes*	26
23	States and justifies the desired significance level	28
24	States and justifies the desired power	26
25	States and justifies the desired smallest difference	28
26	States and justifies the type(s) of test(s) for primary outcomes	8
27	States and justifies the estimates of measures of dispersion	28
28	States and justifies the type of statistic/method of analysis	4
29	States the covariates for which to control in sample size calculation	10
30	Gives software/reference used in sample size calculation	0
	Statistical methods	
31	Study design is well-explained	48
32	Describes study as an RCT	100
33	States whether trial has a standard care/control group	100
34	States intention-to-treat or completers' analysis*	30
35	Defines/justifies methods used to compare primary outcomes*	96
36	Gives method used in interim analysis if applicable*	4
37	Gives method used in subgroup analysis if applicable*	8
38	Gives method used in adjusted analysis if applicable*	40
39	Gives methods used to enhance the quality of measurement*	90
40	States how missing data were handled	40
41	Gives software used in analysis	58
	Allocation and blinding	
42	Gives generation of random allocation sequence*	30
43	Details any restriction of the randomization*	30
44	Defines separation of generator from executor*	10
45	States Declaration of Helsinki/ good clinical practice guidelines	8

Continued...

Appendix 1 continued

Trials	Criteria provided	% reported
46	Gives approval from research ethics committee/institutional review board	84
47	Gives written informed consent from patients	80
48	Gives blinding mechanism*	34
Results		
Baseline data provided		
49	Ages/ age range	98
50	Sex of patients	92
51	History of smoking	36
52	History of hypertension	60
53	History of hyperlipidemia/ hypercholesterolemia	38
54	History of myocardial infarction	74
55	History of renal insufficiency or renal dialysis	62
56	History of diabetes mellitus	74
57	History of carotid stenosis	22
58	History of transient ischemic attack	26
59	History of stroke	56
60	History of chronic obstructive pulmonary disease	34
61	History of congestive heart failure	14
62	History of pulmonary hypertension	2
63	History of peripheral vascular disease	20
64	History of left ventricular aneurysm	4
65	Canadian Cardiovascular Society angina class	42
66	New York Heart Association functional class	24
67	Left ventricular ejection fraction	92
68	Priority of surgery	48
69	Body mass index or obesity	24
Outcomes reported		
70	No. of patients randomly allocated	98
71	No. of patients who completed treatment	70
72	No. of patients in each study stage	62
73	Recruitment period*	48
74	Follow-up period*	82
75	Results of primary and secondary outcome measures*	100
76	Outcome measures compared by treatment group	100
77	Baseline characteristics by treatment group	98
78	Results as absolute numbers*	90
79	Appropriately labelled participant flow chart*	16
80	Results in interim analysis if applicable	88
81	Justified adjusted analyses if applicable*	90
82	Justified subgroup analysis if applicable*	92
83	No. of subjects with at least 1 protocol deviation*	62
84	Adverse events/side-effects in each intervention group*	78
Discussion items		
85	Study objectives*	100
86	Clinical importance of the primary outcome(s)	82
87	Weaknesses or strengths of the trial findings*	51
88	External validity of the trial findings*	58
89	Generalized interpretation of results in context of current evidence*	98

CABG = coronary artery bypass grafting; CONSORT = Consolidated Standards for Reporting Trials; RCT = randomized controlled trial.
*Criteria from the CONSORT statement.