Improvement of quality of reporting in randomised controlled trials to prevent hypotension after spinal anaesthesia for caesarean section

A. Herdan · R. Roth · D. Grass · M. Klimek · S. Will · B. Schauf · R. Rossaint · M. Heesen

Received: 18 November 2010 /Accepted: 3 December 2010 /Published online: 18 December 2010 © The Author(s) 2010. This article is published with open access at Springerlink.com

Abstract Hypotension is a frequent complication of spinal anaesthesia for caesarean section and can threaten the well-being of the unborn child. Numerous randomised controlled trials (RCTs) dealt with measures to prevent hypotension. The aim of this study was to determine the reporting quality of RCTs using the Consolidated Standards of Reporting Trials (CONSORT) statement since low quality can lend false credibility to a study and overestimate the effect of an intervention. We performed a systematic literature search in PubMed to identify relevant RCTs in a pre-CONSORT period (1990–1994) and a post-CONSORT period (2004–2008). A comparative evaluation was done between the two periods, and the trials were assessed for compliance with each of the 22 CONSORT items. A total of 37 RCTs was identified. The CONSORT score increased significantly ($p<0.05$) from 66.7% ($±12.5%$) in the pre-CONSORT period to 87.4% ($±6.9%$) in the post-CONSORT period. A statistically significant improvement was found for eight items, including randomization, blinding and intention-to-treat analysis. The CONSORT score in the post-CONSORT era was fairly good, also in comparison to other medical fields. In the post-CONSORT era, reporting of important items improved, in particular in the domains that are crucial to avoid bias and to improve internal validity. Use of CONSORT should be encouraged in order to keep or even improve the reporting quality.

Keywords Caesarean section · Spinal anaesthesia · Hypotension · Randomised controlled · Trials · CONSORT statement

Background

Hypotension is considered a major side effect of spinal anaesthesia for caesarean section that can compromise foetal circulation, causing hypoxia and foetal acidosis in the unborn baby [1]. Many randomised controlled trials (RCTs), the gold standard of clinical research, were undertaken to study measures aiming to prevent hypotension. In a previous report, we demonstrated that research in this field is limited by a wide variety of definitions of hypotension making it difficult if not impossible to compare the results of the RCTs [2].

The present study sought to evaluate the quality of RCTs in this field. Incomplete data reporting and methodological flaws can limit the value of a RCT. In order to avoid that poor studies receive false credibility, readers need to
evaluate the quality of the methods. In response to growing concerns about the quality of clinical trials, a group of 30 experts including medical journal editors, clinical trialists, epidemiologists and methodologists identified 22 items for which there was evidence that inadequate reporting can introduce bias. A checklist containing all 22 items and a flow diagram were designed, both instruments examining transparent reporting of chronological enrolment, intervention allocation, follow-up and data analysis. So the Consolidated Standards of Reporting Trials (CONSORT) statement was developed and first published in 1996; a revised version was issued in 2001 [3].

There is evidence that studies with a low-quality reporting trend to overestimate the effect of the evaluated intervention by 30–50% [4]. For the assessment of quality of reporting of clinical studies, the CONSORT tool offers some advantages. Proper performance of randomisation, blinding of randomisation and blinding of the subjects and investigators to study interventions will help to reduce bias in clinical studies [5]. Eligibility criteria relate to the external validity of the study, assisting the clinician to decide whether the results of the trial can be applied to his own patients. Scientific validity is reflected by aspects involving statistical analysis. In addition, standards for abstract and title facilitate the process of finding relevant articles.

Since quality of reporting RCTs of interventions to prevent hypotension due to spinal anaesthesia in parturients undergoing caesarean section has not been systematically assessed the present study applied the CONSORT checklist to RCTs. We compared a period before CONSORT (1990–1994) with a post-CONSORT period (2004–2008).

Methods

To identify relevant literature, we performed a PubMed with the search terms “caesarean section” and “hypotension” and “randomised controlled trial” as well as a hand-search of anaesthesiologic journals, the journals Obstetrics and Gynecology as well as the American Journal of Obstetrics and Gynecology. The search was restricted to articles in English language. Two periods were searched: A first period comprised the time from January 1, 1990 until December 31, 1994. The second period ranged from January 1, 2004 until December 31, 2008.

The retrieved articles were then independently screened for eligibility by two authors. Studies were included into this review when healthy parturients scheduled for caesarean section under spinal anaesthesia were randomly assigned to at least one intervention arm that aimed to prevent or treat anaesthesia-induced hypotension. In agreement with others [6], blinding of the graders was not performed since research on the impact of blinding on the evaluation of quality of RCTs has yielded ambiguous results [7, 8].

The CONSORT statement was used to evaluate the reporting quality of the included trials. Each of the 22 CONSORT items was graded as either “yes” or “no”. In case that blinding was not possible, a “not applicable” was also possible. The CONSORT score of each RCT was calculated by adding the correctly reported domains of the CONSORT checklist and expressed as percentage of the maximally available number (in general out of 22, out of 21 in cases that blinding was not possible). All domains had the same weight.

For the purpose of this study, the reviewers underwent systematic training. Initially, they received the German version of the “Revised CONSORT statement for reporting randomised trials: Explanation and elaboration document” which provides the meaning of each item and examples of good practise. Consecutively, they were trained by scoring two articles that were not part of this review.

All reports were independently graded by two investigators. According to our protocol, a third reviewer had to be involved if consensus was not reached between the two first investigators.

To assess chance-adjusted intrarater reliability, Cohen’s kappa statistic was determined. This method involves the degree of reviewers’ agreement on whether a domain was reported or not. It was calculated for three articles that were not included into this study.

The articles were grouped by publication period (1990–1994: pre-CONSORT; 2004–2008: post-CONSORT). The pre- and post-CONSORT periods were compared by calculating the odds ratio and the 95% confidence interval for each domain. CONSORT scores are given as mean ± standard deviation. Student’s t test served for the comparison of CONSORT scores in the two periods. The level of statistical significance was set at the two-sided 0.05 value.

Findings

We identified 48 articles with our search strategy of which 37 [9–45] met the inclusion criteria (Fig. 1). All articles found by the hand-search were also retrieved by the PubMed search. Of these 37 articles, 13 were published in the pre-CONSORT period (1990–1994) and 24 in the post-CONSORT period (2004–2008). Table 1 shows the sources of the retrieved articles, 33 articles were published in anaesthesiologic journals, two in obstetric journals and two others in general medical journals.

The CONSORT scores increased from 66.7±12.5% in the pre-CONSORT period to 87.4±6.9% in the post-CONSORT era (p<0.01). Agreement between the evaluators was good with a k=0.94 (0.92–0.96).
Figure 2 portrays the percentage of correctly described CONSORT items in the pre- and post-CONSORT era. More than a third of all articles from both periods reported 90–100% correctly, and only 5% had a correct reporting of 40–50% of the CONSORT items. When comparing the two time periods, a significant improvement was observed in eight items including sample size calculation, method of randomization, implementation of randomization, blinding, statistical methods, participant flow, intention-to-treat analysis (ITT) and generalizability. A non-significant improvement was found for endpoints, allocation concealment, baseline data, outcomes and estimation of effect, ancillary analysis, interpretation of results and overall evidence (Table 2). For two items, a decrease of correctly reporting articles was found, i.e. recruitment and follow-up as well as adverse effects. In the pre-CONSORT period, five domains were correctly reported in all 13 articles.

**Discussion**

To our knowledge, this is the first study evaluating the compliance with CONSORT of RCTs dealing with hypotension due to spinal anaesthesia for caesarean section. Our approach of studying a period before publication of CONSORT with a time period thereafter has several precedents [6, 46, 47].

As a major result, we found a significant increase in the CONSORT scores from 66.7% to 87.4%. Also, a statistically significant improvement was observed for eight items after CONSORT was published and an improvement in seven items that did not reach statistical significance. Of the total of 22 items, five items were already correctly reported in the pre-CONSORT era.

A remarkable improvement in the items relating to randomization was found. This is critical for the detection of selection bias, and deficiencies were associated with an exaggeration of the treatment effect [4, 48]. Another area of improvement was the participant flow diagram which is intended to explicitly report the number of subjects undergoing randomization, receiving treatment, the number of dropouts and the number finally being analysed [49].

ITT was used by all reports of the post-CONSORT era versus ten out of 13 pre-CONSORT articles. ITT is strongly recommended since it preserves the randomization process and allows for non-compliance and deviations from policy.
ITT together with methods of randomization and blinding is crucial for internal validity and helps to avoid selection, performance, detection and attrition bias [5]. Sample size calculation was an item that showed the sharpest increase over time, i.e. the percentage of trials correctly reporting on this item increased from 23% (three of 13 reports) to 71% (17 of 24 reports) in the post-CONSORT period. Sample size calculation is required to quantitatively estimate the power of a trial to answer the studied question [46]. Underpowered trials are prone to bias and can negatively affect the quality of meta-analyses [46]. However, it has to be emphasized that this CONSORT item only evaluates whether a sample size calculation has been performed but it does not give information of whether this calculation is correct or not.

Interestingly, our analysis found a lower percentage of correct reporting for the items recruitment and follow-up as well as adverse effects in the post- versus the pre-CONSORT period, although statistically not significant. It could be speculated that the reporting of adverse effects was omitted since authors feared that the occurrence of side effects and complications could question their positive results and lead to a rejection of the manuscript. The fact that inadequate reporting is found in the more recent papers could be due to the increasing competition in the academic field leading to a growing pressure to produce publications, since the number of authored papers is one of the parameters for evaluating scientific careers [51, 52]. A report that evaluated surgical papers which were published in 2005 confirmed our finding of a high rate of inadequate reporting of adverse effects [53]. Editors should probably draw the conclusion that authors have to be encouraged to report on undesired effects. Proper definition and reporting of adverse events is crucial for critical appraisal of study results and, in addition, facilitates systematic reviews and meta-analysis. Previous studies showed that acute surveillance, i.e. actively asking study subjects whether undesired events occurred by use of structured questionnaires, inter-

| Data items                | Combined | Pre-CONSORT | Post-CONSORT | Odds ratio (95% CI) | p value |
|---------------------------|----------|-------------|--------------|---------------------|---------|
| Title and abstract        | 1.0 (37) | 1.0 (13)    | 1.0 (24)     | 1.0                 |         |
| Introduction              | 1.0 (37) | 1.0 (13)    | 1.0 (24)     | 1.0                 |         |
| Scientific background     | 1.0 (37) | 1.0 (13)    | 1.0 (24)     | 1.0                 |         |
| Eligibility criteria for participants | 1.0 (37) | 1.0 (13)    | 1.0 (24)     | 1.0                 |         |
| Details of Intervention   | 1.0 (37) | 1.0 (13)    | 1.0 (24)     | 1.0                 |         |
| Objectives                | 1.0 (37) | 1.0 (13)    | 1.0 (24)     | 1.0                 |         |
| Endpoints                 | 0.89 (33)| 0.77 (10)   | 0.96 (23)    | 6.9 (0.69; 74.51)   | 0.115   |
| Sample size               | 0.54 (20)| 0.23 (3)    | 0.71 (17)    | 8.1 (1.70; 38.58)   | 0.007   |
| Methods of randomization  | 0.70 (26)| 0.38 (5)    | 0.88 (21)    | 11.2 (2.11; 58.37)  | 0.003   |
| Allocation concealment    | 0.57 (21)| 0.38 (5)    | 0.67 (16)    | 3.2 (1.17; 8.74)    | 0.096   |
| Implementation of randomization | 0.41 (15)| 0.08 (1)    | 0.58 (14)    | 16.8 (1.87; 150.73) | 0.003   |
| Blinding                  | 0.65 (24)| 0.31(4)     | 0.8 (20)     | 11.25 (2.29; 55.35) | 0.0023  |
| Statistical methods       | 0.89 (33)| 0.69 (9)    | 1.0 (24)     | 0.011               |         |
| Results                   |          |             |              |                     |         |
| Participant flow          | 0.91 (34)| 0.77 (10)   | 1.0 (24)     | 0.037               |         |
| Recruitment + follow-up   | 0.51 (19)| 0.61 (8)    | 0.46 (11)    | 0.53 (0.13; 2.10)   | 0.895   |
| Baseline data             | 0.95 (35)| 0.85 (11)   | 1.0 (24)     | 0.12                |         |
| Intention-to-treat analysis| 0.92 (34)| 0.76 (10)   | 1.0 (24)     | 0.037               |         |
| Outcome and estimation    | 0.95 (35)| 0.85 (11)   | 1.0 (24)     | 0.12                |         |
| Ancillary analysis        | 0.29 (11)| 0.15 (2)    | 0.38 (9)     | 3.3 (0.59; 18.40)   | 0.15    |
| Adverse effects           | 0.81 (30)| 0.92 (12)   | 0.75 (18)    | 0.25 (0.03; 2.34)   | 0.9     |
| Discussion                |          |             |              |                     |         |
| Interpretation of results | 0.95 (35)| 0.84 (11)   | 1.0 (24)     | 0.12                |         |
| Generalizability          | 0.86 (32)| 0.69 (9)    | 0.96 (23)    | 10.22 (1.00; 104.0) | 0.042   |
| Overall evidence          | 0.95 (35)| 0.85 (11)   | 1.0 (24)     | 0.12                |         |

CI confidence interval
views or diagnostic tests at predefined time intervals, is more effective than passive disclosure [44, 45]. Reporting of adverse events should already be considered during study design since data on adverse events are less susceptible to bias and confounders when they are collected prospectively rather than retrospectively [54, 55]. Our finding is confirmed by a study about the reporting quality of surgical trials [53].

There are many other tools to evaluate the quality of RCTs. We chose the CONSORT checklist for several reasons. Firstly, CONSORT is officially supported by the World Association of Medical Authors and the International Committee of Medical Journal Editors. Secondly, several reports demonstrated that poor compliance with CONSORT criteria is associated with an exaggeration of the effect size. Moher and colleagues found that RCTs that were not double-blind overestimated the effects by 17% [4]. Inadequate or unclear allocation concealment exaggerated odds ratios by 41% and 30%, respectively. A similar finding was obtained by another group [56]. An exaggeration of the treatment effect in single RCTs by inappropriate reporting has also consequences for subsequent meta-analyses, considered to be the highest level of evidence-based medicine and often guiding our clinical practice. In their meta-analysis on anticoagulants, Lensing et al. [57] clearly favoured low molecular heparins. However, when taking quality of reporting of the underlying RCTs into consideration, superiority of low molecular weight heparin on mortality due to venous thrombosis was no longer given [4]. Exclusion of low-quality trials may thus directly impact our clinical practice.

It is probably a consequence of its widespread use that CONSORT has already been proven effective: A meta-analysis including 248 articles clearly demonstrated an improvement of the reporting of RCTs by the adoption of this tool [49]. It is noteworthy that these authors compared a time period before CONSORT publication with a period thereafter; this approach was also chosen in our report.

Thirdly, CONSORT has in the meantime been applied in many other medical specialities [53, 58]. We used the CONSORT checklist in its 2001 version. An update has been published very recently [59], after our study was completed. However, use of the 2001 version allows us to compare our results with those obtained in other medical specialities and therapeutic areas.

In only 12% of trials of analgesics given for pain relief after trauma or orthopaedic surgery sample size calculation was reported [60]. This percentage was 54% in our survey, combining the pre- and post-CONSORT period. The reason for the discrepancy could be due to the fact that the other group included trials published from 1966 on, thus covering a time period when probably little or no attention was paid to this issue, whereas our early period began much later, in 1990. In a study of nutritional support trials [58], use of blinding was found to increase significantly from 19% to 41% from a pre- (before 1996) to the post-CONSORT (after 1996) era. We observed an increase from 31% to 80%. In our study, intention-to-treat analysis was used in 76% of the pre-CONSORT and in 100% of the post-CONSORT articles, an improvement reaching statistical significance. No improvement was found for this parameter in the study by Doig and colleagues [58].

The CONSORT score of urological and non-urological surgical trials published between 2000 and 2003 was 11.1 and 11.2 (corresponding to 50.45% and 50.90%) [61], respectively, which is much lower than the CONSORT scores in either of our study periods.

Comparing our results with a study of obstetric anaesthesia trials could be of particular interest. This report by Halpern and colleagues [61] focussed on articles published from June 2000 to June 2002, providing no information on whether reporting quality might have changed over time. When comparing the reports included into this study [61], which were published between June 2000 and June 2002, with articles from our post-CONSORT era (2004–2009), we found a higher percentage of correctly reported CONSORT items in our study. More than 80% correctly reported domains were observed in 62% of our articles whereas less than 5% of the articles from the study by Halpern et al. [61] reach this value. The majority of the obstetric anaesthesia articles only ranged between 50% and 70% correctly reported items, compared to less than 5% in our sample of RCTs of interventions to prevent hypotension after spinal anaesthesia. We think it is the time difference between the two studies (2004 to 2009) in our post-CONSORT period versus June 2000 to June 2002 in the study by Halpern et al. [61] which accounts for the higher percentage of compliance with some CONSORT items in our study, since awareness of the CONSORT statement published in 1996 (respective in 2001 for the revised version) had less time to spread in the study period of Halpern et al. [61]. This notion receives further support from studies by others: Only few improvements could be observed when comparing trials from before 1996 with those appearing immediately after 1996 [60]. Moher et al. [46] also considered it a flaw of their study that their post-CONSORT period began only 12 to 18 months after publication of CONSORT. They argued that “effective dissemination is a slow process and that to estimate the true influence of CONSORT requires more time.”

It is a limitation of our study that we do not provide data on the association of quality of reporting and CONSORT adoption. We did not test for the association between the CONSORT score in adopters and non- or late adopters since every additional analysis increases the likelihood of a type I error, i.e. incorrectly inferring that there is a significant difference when there is no such difference.
In summary, the CONSORT scores increased over time from 66.7% to 87.4%, reflecting a remarkable improvement. We observed a significantly better reporting in eight of 22 items in the post-CONSORT compared to the pre-CONSORT period. However, we saw a decrease, although statistically not significant, in the reporting of adverse events which deserves further attention.

Conclusion

We conclude from our study that the reporting quality has improved significantly in the period after dissemination of the CONSORT statement. Therefore, journal editors, reviewers and authors should be encouraged to adhere to the CONSORT checklist in order to ensure high-quality trials. Consequently, clinicians can spend more time considering the findings, rather than scrutinizing quality of reporting of the trial.

Declaration of interest  The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

Open Access  This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

1. Corke BC, Datta S, Ostheimer GW, Weiss JB, Alper MH (1982) Spinal anaesthesia for caesarean section. The influence of hypotension on neonatal outcome. Anaesthesia 37:658–662
2. Kloth S, Roth R, Hofmann T, Rossaint R, Heesen M (2010) Definitions of hypotension after spinal anaesthesia for caesarean section: systematic literature search and application to parturients. Acta Anaesthesiol Scand 1:1–13
3. Moher D, SchulzKF, Altman DG (2001) The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. Lancet 357:1191–1194
4. Moher D, Pham B, Jones A, Cook DJ, Jadad AR, Moher M, Tugwell P, Klassen TP (1998) Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? Lancet 352:609–613
5. Jüni P, Altman DG, Egger M (2001) Systematic reviews in health care: assessing the quality of controlled clinical trials. BMJ 323:42–46
6. Moher D, Jones A, Lepage L, CONSORT group (2001) Use of the CONSORT statement and quality of reports of randomized trials: a comparative before- and after-evaluation. JAMA 285:1992–1995
7. Berlin JA (1997) Does blinding of readers affect the results of meta-analysis? University of Pennsylvania Meta-Analysis Blinding Study Group. Lancet 350:185–186
8. Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynold DJ, Gavaghan DJ, McQuay HJ (1996) Assessing the quality of reports of randomized clinical trials: is blinding necessary? Control Clin Trials 17:1–12
9. Bhagwanjee S, Rocke DA, Rout CC, Koovarjee RV, Brijball R (1990) Prevention of hypotension following spinal anaesthesia for elective caesarean section by wrapping of the legs. Br J Anaesth 70:819–822
10. Moran DH, Perillo M, La Porta RF, Bader AM, Datta S (1991) Phenylephrine in the prevention of hypotension following spinal anesthesia for cesarean delivery. J Cin Anesth 3:301–305
11. Wright PMC, Littkohar M, Fitzpatrick KT, Moore J, Thompson W (1992) Vasopressor therapy for hypotension during epidural anesthesia for cesarean section: effects on maternal and fetal flow velocity ratios. Anesth Analg 75:56–63
12. Rout CC, Rocke DA, Brijball R, Koovarjee RV (1992) Prophylactic intramuscular ephedrine prior to caesarean section. Anesth Intensive Care 20:448–452
13. Rout CC, Akoojee SS, Rocke DA, Gouws E (1992) Rapid administration of crystalloid preload does not decrease the incidence of hypotension after spinal anaesthesia for elective caesarean section. Br J Anaesth 68:394–397
14. Rout CC, Rocke DA, Levin J, Gouws E, Reddy D (1993) A reevaluation of the role of crystalloid preload in the prevention of hypotension associated with spinal anaesthesia for elective caesarean section. Anesthesiology 79:262–269
15. Rout CC, Rocke DA, Gouws E (1993) Leg elevation and wrapping in the prevention of hypotension following spinal anaesthesia for elective caesarean section. Anæsthesia 48:304–308
16. Kafle SK (1993) Intrathecal meperidine for elective caesarean section: a comparison with lidocaine. Can J Anaesth 40:718–721
17. Campbell DC, Douglas MJ, Pavy TJG, Merrick P, Flanagan ML, McMorland GH (1993) Comparison of the 25-gauge Whitacre with the 24-gauge Sprotte spinal needle for elective caesarean section: cost implications. Can J Anaesth 40:1131–1135
18. Thoren T, Holmström B, Rawal N, Schollin J, Lindeberg S, Skeppner G (1994) Sequential combined spinal epidural block versus spinal block for caesarean section: effects on maternal hypotension and neurobehavioral function of the newborn. Anesth Analg 78:1087–1092
19. Ramin SM, Ramin KD, Cox K, Magness RR, Shearer VE, Gant NF (1994) Comparison of prophylactic angiotensin II versus ephedrine infusion for prevention of maternal hypotension during spinal anaesthesia. Am J Obstet Gynecol 171:734–739
20. Hall PA, Bennett A, Wilkes MP, Lewis M (1994) Spinal anaesthesia for caesarean section: comparison of infusion of phenylephrine and ephedrine. Br J Anaesth 73:471–474
21. Filos KS, Goudas LC, Patroni O, Polyzou V (1994) Hemodynamic and analgesic profile after intrathecal clonidine in humans. Anesthesiology 81:591–601
22. Loughrey JPR, Yao S, Datta S, Segal S, Pian-Smith M, Tsen LC (2005) Colloid vs. crystalloid preloading to prevent maternal hypotension during spinal anaesthesia for elective caesarean section. Anaesthesia 60:348–352
23. Hallworth SP, Fernando R, Columb MO, Stocks GM (2005) The effect of posture and baricity on the spread of intrathecal bupivacaine for elective cesarean delivery. Anesth Analg 100:1159–1165
24. Koo WDN, Khaw KS, Ng FF (2005) Prevention of hypotension during spinal anaesthesia for cesarean delivery. Anesthesiology 103:744–750
25. Dahlgren G, Granath F, Pregner K, Rööblad PG, Wessel H, Irested L (2005) Colloid vs. crystalloid preloading to prevent maternal hypotension during spinal anaesthesia for elective cesarean section. Acta Anaesthesiol Scand 49:1200–1206
