Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals (Review)

Turner L, Shamseer L, Altman DG, Weeks L, Peters J, Kober T, Dias S, Schulz KF, Plint AC, Moher D

Turner L, Shamseer L, Altman DG, Weeks L, Peters J, Kober T, Dias S, Schulz KF, Plint AC, Moher D.
Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals.
Cochrane Database of Systematic Reviews 2012, Issue 11. Art. No.: MR000030.
DOI: 10.1002/14651858.MR000030.pub2.

TABLE OF CONTENTS

[Methodology Review]

Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Lucy Turner[1], Larissa Shamseer[1], Douglas G Altman[2], Laura Weeks[3], Jodi Peters[1], Thilo Kober[4], Sofia Dias[5], Kenneth F Schulz[6], Amy C Plint[7], David Moher[1]

[1]Ottawa Hospital Research Institute, Ottawa, Canada. [2]Centre for Statistics in Medicine, Wolfson College Annexe, Oxford, UK. [3]Ottawa Integrative Cancer Centre, Ottawa, Canada. [4]Smiling Press International, Canberra, Australia. [5]School of Social and Community Medicine, University of Bristol, Bristol, UK. [6]Quantitative Sciences, FHI 360 and UNC School of Medicine, Research Triangle Park, North Carolina, USA. [7]Departments of Pediatrics and Emergency Medicine, University of Ottawa, Ottawa, Canada

Contact address: Lucy Turner, Ottawa Hospital Research Institute, 501 Smyth Rd, Box 201B, Ottawa, ON, K1H 8L6, Canada. lturner@ohri.ca.

Editorial group: Cochrane Methodology Review Group.
Publication status and date: Edited (no change to conclusions), published in Issue 1, 2013.

Citation: Turner L, Shamseer L, Altman DG, Weeks L, Peters J, Kober T, Dias S, Schulz KF, Plint AC, Moher D. Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals. *Cochrane Database of Systematic Reviews* 2012, Issue 11. Art. No.: MR000030. DOI: 10.1002/14651858.MR000030.pub2.

ABSTRACT

Background

An overwhelming body of evidence stating that the completeness of reporting of randomised controlled trials (RCTs) is not optimal has accrued over time. In the mid-1990s, in response to these concerns, an international group of clinical trialists, statisticians, epidemiologists, and biomedical journal editors developed the CONsolidated Standards Of Reporting Trials (CONSORT) Statement. The CONSORT Statement, most recently updated in March 2010, is an evidence-based minimum set of recommendations including a checklist and flow diagram for reporting RCTs and is intended to facilitate the complete and transparent reporting of trials and aid their critical appraisal and interpretation. In 2006, a systematic review of eight studies evaluating the "effectiveness of CONSORT in improving reporting quality in journals" was published.

Objectives

To update the earlier systematic review assessing whether journal endorsement of the 1996 and 2001 CONSORT checklists influences the completeness of reporting of RCTs published in medical journals.

Search methods

We conducted electronic searches, known item searching, and reference list scans to identify reports of evaluations assessing the completeness of reporting of RCTs. The electronic search strategy was developed in MEDLINE and tailored to EMBASE. We searched the Cochrane Methodology Register and the Cochrane Database of Systematic Reviews using the Wiley interface. We searched the Science Citation Index, Social Science Citation Index, and Arts and Humanities Citation Index through the ISI Web of Knowledge interface. We conducted all searches to identify reports published between January 2005 and March 2010, inclusive.

Selection criteria

In addition to studies identified in the original systematic review on this topic, comparative studies evaluating the completeness of reporting of RCTs in any of the following comparison groups were eligible for inclusion in this review: 1) Completeness of reporting of RCTs published in journals that have and have not endorsed the CONSORT Statement; 2) Completeness of reporting of RCTs published in CONSORT-endorsing journals before and after endorsement; or 3) Completeness of reporting of RCTs before and after the publication of the CONSORT Statement (1996 or 2001). We used a broad definition of CONSORT endorsement that includes any of the following: (a) requirement or recommendation in journal's 'Instructions to Authors' to follow CONSORT guidelines; (b) journal editorial statement endorsing the CONSORT Statement; or (c) editorial requirement for authors to submit a CONSORT checklist and/or flow diagram with their manuscript. We contacted authors of evaluations reporting data that could be included in any comparison group(s), but not presented as such in the published report and asked them to provide additional data in order to determine eligibility of their evaluation. Evaluations were not excluded due to language of publication or validity assessment.

Data collection and analysis

We completed screening and data extraction using standardised electronic forms, where conflicts, reasons for exclusion, and level of agreement were all automatically and centrally managed in web-based management software, DistillerSR®. One of two authors extracted general characteristics of included evaluations and all data were verified by a second author. Data describing completeness of reporting were extracted by one author using a pre-specified form; a 10% random sample of evaluations was verified by a second author. Any discrepancies were discussed by both authors; we made no modifications to the extracted data. Validity assessments of included evaluations were conducted by one author and independently verified by one of three authors. We resolved all conflicts by consensus.

For each comparison we collected data on 27 outcomes: 22 items of the CONSORT 2001 checklist, plus four items relating to the reporting of blinding, and one item of aggregate CONSORT scores. Where reported, we extracted and qualitatively synthesised data on the methodological quality of RCTs, by scale or score.

Main results

Fifty-three publications reporting 50 evaluations were included. The total number of RCTs assessed within evaluations was 16,604 (median per evaluation 123 (interquartile range (IQR) 77 to 226) published in a median of six (IQR 3 to 26) journals. Characteristics of the included RCT populations were variable, resulting in heterogeneity between included evaluations. Validity assessments of included studies resulted in largely unclear judgements. The included evaluations are not RCTs and less than 8% (4/53) of the evaluations reported adjusting for potential confounding factors.

Twenty-five of 27 outcomes assessing completeness of reporting in RCTs appeared to favour CONSORT-endorsing journals over non-endorsers, of which five were statistically significant. 'Allocation concealment' resulted in the largest effect, with risk ratio (RR) 1.81 (99% confidence interval (CI) 1.25 to 2.61), suggesting that 81% more RCTs published in CONSORT-endorsing journals adequately describe allocation concealment compared to those published in non-endorsing journals. Allocation concealment was reported adequately in 45% (393/876) of RCTs in CONSORT-endorsing journals and in 22% (329/1520) of RCTs in non-endorsing journals. Other outcomes with results that were significant include: scientific rationale and background in the 'Introduction' (RR 1.07, 99% CI 1.01 to 1.14); 'sample size' (RR 1.61, 99% CI 1.13 to 2.29); method used for 'sequence generation' (RR 1.59, 99% CI 1.38 to 1.84); and an aggregate score over reported CONSORT items, 'total sum score' (standardised mean difference (SMD) 0.68 (99% CI 0.38 to 0.98)).

Authors' conclusions

Evidence has accumulated to suggest that the reporting of RCTs remains sub-optimal. This review updates a previous systematic review of eight evaluations. The findings of this review are similar to those from the original review and demonstrate that, despite the general inadequacies of reporting of RCTs, journal endorsement of the CONSORT Statement may beneficially influence the completeness of reporting of trials published in medical journals. Future prospective studies are needed to explore the influence of the CONSORT Statement dependent on the extent of editorial policies to ensure adherence to CONSORT guidance.

PLAIN LANGUAGE SUMMARY

CONsolidated Standards Of Reporting Trials (CONSORT) and the completeness of reporting of randomised controlled trials published in medical journals

A group of experts has developed a checklist and flow diagram called the CONSORT Statement. The checklist is designed to help authors in the reporting of randomised controlled trials (RCTs). This systematic review aims to determine whether the CONSORT Statement has made a difference to the completeness of reporting of RCTs. Reporting of RCTs published in journals that encourage authors to use the CONSORT Statement with those that do not is compared. We found that some items in the CONSORT Statement were fully reported more often when journals encouraged the use of CONSORT. While the majority of items are reported more often when journals endorse CONSORT, the data only showed a statistically significant improvement in reporting for five of 27 items. No items suggest that CONSORT decreases the completeness of reporting of RCTs published in medical journals.

None of the evaluations included in this review used experimental designs, and their methodological approaches were mostly poorly described and variable when they were described. Furthermore, evaluations assessed the completeness of reporting of RCTs within a wide range of medical fields and in journals with a wide variation in the enforcement of CONSORT endorsement. Our results do have some limitations, but given the number of included evaluations and the number of assessed RCTs, we conclude that while most RCTs are incompletely reported, the CONSORT Statement beneficially influences their reporting quality.

BACKGROUND

An overwhelming body of evidence demonstrating that the completeness of reporting of randomised controlled trials (RCTs) is sub-optimal has accrued over time (Chan 2005; Glasziou 2008; Hopewell 2008; Moher 2010). In the mid-1990s, in response to concerns about this issue, an international group of clinical trialists, statisticians, epidemiologists, and biomedical editors developed the CONsolidated Standards Of Reporting Trials (CONSORT) Statement (Begg 1996), which has been twice revised and updated (Moher 2001a; Schulz 2010). The CONSORT Statement is an evidence-based set of recommendations for reporting two-arm, parallel-group RCTs, including a minimum set of items to be reported pertaining to the rationale, design, analysis, and interpretation of the trial (i.e. CONSORT checklist) and a diagram describing flow of participants through a trial (i.e. flow diagram). It is intended to facilitate the complete and transparent reporting of RCTs and in turn aid in their critical appraisal and interpretation.

The CONSORT Statement was first published in 1996 (Begg 1996). It included 21 checklist items pertaining to the rationale, design, analysis, and interpretation of a trial (i.e. CONSORT checklist) and a flow diagram outlining the progress of participants through a trial. In 2001, the CONSORT checklist, updated to 22 items, and flow diagram were revised to reflect emerging evidence indicating that lack of, or poor reporting of particular elements of RCTs is associated with biased estimates of treatment effect (Moher 2001a). Some new items were also added because reporting them was found to increase the ability to judge the validity or relevance of trial findings (Moher 2001a). Evidence and examples for each checklist item are found in an accompanying Explanation and Elaboration (E&E) document (Altman 2001). The second revision, and current version, of the CONSORT Statement (CONSORT 2010) was published in March 2010 (Schulz 2010). It contains an updated 25-item checklist and flow diagram, also accompanied by an E&E document (Moher 2010). All CONSORT materials are available on the CONSORT website (www.consort-statement.org; CONSORT Group 2009). For ease, henceforth, 'CONSORT' will refer to this collective body of literature, unless otherwise stated.

To date, the CONSORT Statement has received positive attention, in part, by way of endorsement by biomedical journals. To date, over 600 journals have endorsed the CONSORT Statement. Such endorsement is typically evidenced by a statement in a journal's 'Instructions to Authors' regarding the use (suggested or required) of CONSORT while preparing trial reports for publication. Some journals publish editorials indicating their support, while others institute mandatory submission of a guideline checklist and/or flow diagram along with manuscript submission. As such, while the CONSORT Statement is widely endorsed, there is huge variation in terms of how CONSORT policies are implemented.

Description of the problem or issue

Concurrent with the publication of the 2001 CONSORT Statement, Moher and colleagues reported the first evaluation of endorsement of the CONSORT checklist. The authors reported that the completeness of reports of RCTs in CONSORT-endorsing journals was higher than one non-endorsing journal (Moher 2001). Since then, other evaluations have been published which assess the influence of CONSORT either directly or indirectly on the completeness of reporting of RCTs. In 2006, Plint and colleagues (Plint 2006) published a systematic review synthesising data from all such evaluations to gauge their combined findings

about the influence of CONSORT endorsement on the completeness of reporting of RCTs. Despite methodological weaknesses of the eight included evaluations, the review found that endorsement of CONSORT may influence the completeness of reporting in some checklist items. For example, reporting of the method of sequence generation, allocation concealment, and overall number of CONSORT items (i.e. 'total sum score') was more common in RCTs published in CONSORT-endorsing compared to non-endorsing journals, but CONSORT endorsement seemed to have less effect on the reporting of participant flow and blinding (Plint 2006).

In the six years since this systematic review was published, a number of additional evaluations of the effects of CONSORT on the completeness of reporting have been published. Some of these evaluations directly assess the effect of CONSORT on complete reporting (e.g. Hopewell 2010), others assess complete reporting based on CONSORT criteria in a specific medical field or research area, for example RCTs investigating weight loss (e.g. Thabane 2007), glaucoma (Llorca 2005), and surgery (Agha 2007). For these latter evaluations, the effect of CONSORT can be assessed through a post hoc comparison of completeness of reporting of RCTs published in CONSORT-endorsing versus non-endorsing journals.

This systematic review updates Plint et al's review to include and synthesise results that have been published in the time since the first review was conducted.

Why it is important to do this review

The Plint et al systematic review included evaluations published between January 1996 and July 2005 (Plint 2006). Over six years have passed since the search for literature in that review was carried out and a considerable number of additional evaluations have been published that are relevant to include in this update. For readers looking to know whether CONSORT endorsement influences the completeness of reporting, it is necessary to update Plint et al's review and to incorporate the most comprehensive corpus of literature on this topic. This updated review provides a more complete perspective regarding the possible influence of CONSORT on the completeness of reporting of RCTs and, subsequently, will allow journal editors, methodologists, and trialists to understand the potential benefits of using CONSORT when reporting the design, analysis, and interpretation of RCTs.

OBJECTIVES

To assess whether journal endorsement of CONSORT is associated with more complete reporting of RCTs, by examining the following comparisons:

- comparison 1: completeness of reporting of RCTs published in journals that have and have not endorsed the CONSORT Statement; and/or
- comparison 2: completeness of reporting of RCTs published in CONSORT-endorsing journals before and after endorsement; or
- comparison 3: completeness of reporting of RCTs before and after the publication of CONSORT (i.e. 1996 and 2001).

During the review process, two additional comparisons were identified and reported in already included evaluations, namely completeness of reporting of RCTs published before endorsement in endorsing and non-endorsing journals and completeness of reporting of RCTs published in non-endorsing journals before and after endorsement (where after endorsement was determined by their endorsing counterparts). These comparisons were formed in evaluations to assess, by proxy, potential confounding. We collected data for these comparisons as encountered as they provided information on potential confounders (i.e. the effect of non-endorsement over time and the effect of potential pre-existing differences in completeness of reporting between endorsing and non-endorsing journals). Data for these comparisons were sparse and we carried out no meta-analyses; these data are available upon request.

METHODS

Criteria for considering studies for this review

Types of studies

Any report evaluating the completeness of reporting of RCTs, potentially eligible for any of the three main comparisons, was included; such studies are termed 'evaluations' for the remainder of this report.

We identified evaluations for potential inclusion using the following pre-specified screening questions:

- Does the evaluation involve a relevant comparison (e.g. pre CONSORT publication versus post CONSORT publication or otherwise)?
- Does the evaluation examine the influence of the CONSORT checklist on the completeness of reporting of RCTs?
- Does the evaluation report any of the following: a) 22 items on the CONSORT checklist?, b) any type of overall quality indicators/score? c) adherence to CONSORT checklist?

We approached authors of evaluations that were not comparative, or did not report data in a format coinciding with our needs, for

supplementary information. Subsequently any additional evaluations for which a comparison could be drawn, were included (e.g. Dias 2006).

Types of data

We included studies published in biomedical journals, pertaining to any general or medical subspecialty that enabled comparison of the completeness of reporting of RCTs in any of our three main comparison groups.
In addition, this review only includes evaluations of the 1996 and 2001 CONSORT Statements, since publication of the CONSORT 2010 statement coincides with the search dates for this review and so no evaluations could have been conducted and reported in time for inclusion.

Types of methods

Evaluations using any method to identify and evaluate the reporting of RCTs were included in this review. Evaluations may or may not have considered endorsement of CONSORT as the primary 'exposure' of interest. For instance, evaluations that did not specifically assess CONSORT checklist items, but evaluated the reporting of items relating to existing CONSORT items, were included.

Types of outcome measures

Primary outcomes

The primary outcome is the completeness of reporting of RCTs, as measured by adequate or inadequate reporting of any of the following 27 outcomes: 22 items on the 2001 CONSORT checklist, four additional items relating to the reporting of blinding (i.e. blinding of participants, data analyst, outcome assessor, or intervention), or a sum score across aggregate checklist items, as reported in evaluations. The 2001 CONSORT checklist is reproduced in Table 1. We considered the 22 checklist items in the 2001 Statement as the 'core' items and the four additional items on blinding are simply referred to *en masse* as pertaining to the CONSORT item on 'blinding'. All analyses presented are ordered in line with the CONSORT checklist (i.e. allocation concealment is checklist item number 9 and, hence, results are presented as 1.9, 2.9, and 3.9 for the three comparison groups).

Secondary outcomes

1. Methodological quality of RCTs included in evaluations, as reported

In addition to primary and secondary outcomes, we have included and described evaluations which met the inclusion criteria, but were not eligible for inclusion in meta-analyses.

Search methods for identification of studies

We conducted electronic searches of bibliographic databases, known item searching, and reference list scans to identify records published from January 2005 to March 2010, to capture studies reported in the period after the search of the original systematic review (Plint 2006).
It should be noted that the search was purposefully limited to exclude records published after the publication of the CONSORT 2010 Statement (on 25 March 2010), as there was insufficient time for evaluations of CONSORT 2010 to have been carried out. A future update of this systematic review will include evaluations of the 2010 Statement.

Electronic searches

To ensure all possibly relevant evaluations were obtained, we designed the main search strategy to retrieve reports published since the date of the last search of the original review, carried out in July 2005. Specifically, the dates of the search for this review cover publications from January 2005 in order to ensure that articles which may have been published in the first half of 2005, but not indexed at the time of searching during the original review, were identified.
We conducted literature searches in Ovid MEDLINE (January 2005 to 19 March 2010); OVID EMBASE (January 2005 to 2010 Week 10); ISI Web of Knowledge (including citing reference searches) 2005 to 19 March 2010; Cochrane Methodology Register; and the Cochrane Database of Systematic Reviews (*The Cochrane Library* 2010, Issue 1). We searched the Cochrane Methodology Register and the Cochrane Database of Systematic Reviews using the Wiley interface. We searched the Science Citation Index, Social Science Citation Index, and Arts and Humanities Citation Index through the ISI Web of Knowledge interface. Please see Appendix 1 for the full search strategy, which was developed in MEDLINE and tailored to EMBASE.

Searching other resources

Evaluations were also identified by members of the research team when attending conferences, or from discussions with experts in the field.

Data collection and analysis

Selection of evaluations

We conducted all screening using an online data management software, DistillerSR®, a program capable of tracking and managing the progress of records (i.e. abstracts and full-text reports)

through a review. Title and abstract screening were completed independently, in duplicate by two of three authors (LS, LT, LW) using broad screening criteria. All possibly relevant evaluations and those with all conflicting assessments of reports were included for further review.

The full text of all records identified as potentially eligible were retrieved and independently reviewed for eligibility by two authors (LS and LT) using standardised inclusion criteria developed a priori. Full-text screening disagreements were resolved by consensus or by an independent third author (DM). Six non-English language articles were assessed by colleagues fluent in the relevant language, who completed the same standardised inclusion forms as the other assessors.

Potentially eligible studies were either categorised into one of the three main comparisons of this review or needed further information from authors to determine eligibility, such as whether included trials were published in endorsing or non-endorsing journals or, if that information was unavailable, a list of included journals for review authors to follow up with and determine date and status of endorsement. We contacted authors for this information during data extraction so that both eligibility and potentially necessary data could be obtained in one effort.

For the purpose of this review, endorsement is defined as any of the following situations, implying that, in principle, the CONSORT Statement is incorporated into the editorial process for a particular journal: (a) requirement or recommendation in journal's 'Instructions to Authors' to follow CONSORT when preparing their manuscript; (b) journal editorial statement endorsing the CONSORT Statement: either the flow diagram, the checklist or both; or (c) editorial requirement for authors to submit a CONSORT checklist and/or flow diagram with their manuscript. We determined endorsement status by first cross-checking with the CONSORT group's endorser database. If the journal was not listed, we then reviewed the journals' 'Instructions to Authors' for related text and, if unavailable, lastly by searching for an editorial statement or through previous journal issues for such a statement. Finally, we assumed journals determined not to have endorsed CONSORT at the time we checked for this information never to have been endorsers.

For journals identified as CONSORT endorsers at the time of checking, we sought dates of endorsement by contacting their managing editors or other editorial staff. This information was collected to determine whether RCTs were published after a reasonable amount of time following endorsement, such that its effect had sufficient time to be realised in a journal's output. For this review, we considered six months an adequate amount of time. Determining dates of endorsement was a resource-intense process; for evaluations assessing large numbers of RCTs or large numbers of journals it was not feasible to collect this information. For evaluations where endorsement status has not been verified, this has been noted in the Characteristics of included studies. Evaluations were not excluded on this basis; we used this information to conduct sensitivity analyses, as described below.

We did not exclude evaluations based on publication status, language of publication, or validity assessment. When multiple reports of a single evaluation were identified and outcomes were overlapping, only outcome data from the main publication were included. Data on additional outcomes presented in secondary publications were included under their corresponding secondary publications.

Data extraction and management

We completed data extraction using standardised electronic forms, where conflicts, reasons for exclusion, and level of agreement were all automatically and centrally managed in web-based management software, DistillerSR®. One of two authors extracted general characteristics of included evaluations and all data were verified by a second author. Data describing completeness of reporting were extracted by one author using a pre-specified form; a 10% random sample of evaluations was verified by a second author. Any discrepancies were discussed by both authors.

We extracted the following data from included evaluations:

We extracted general characteristics of evaluations including its journal of publication, number of included RCTs, number of journals, country of publication, source of funding, and CONSORT checklist version used and information pertaining to journal 'quality' (i.e. enforcement of the checklist, editorial policy, size of editorial team, volume of publications, impact factor, and other potential determinants) included in the evaluation.

We collected completeness of reporting of RCTs in included evaluations across 27 a priori outcome measures (Primary outcomes). These included adequacy of reporting any of the 22 2001 core CONSORT checklist items, four additional items pertaining to the 2001 CONSORT checklist item on blinding, and/or a 'sum score' of aggregate checklist items.

For simplicity, we used items on the 2001 CONSORT as data extraction items since they were all encompassing of both CONSORT checklist versions; they include all items contained within the 1996 checklist (some with rewording for improved reporting) as well as some additional items. When completeness of reporting using the 1996 CONSORT checklist was reported in an evaluation, we included items from that checklist that were the same as those in the 2001 checklist. However, for those items of the 2001 version which differed from the 1996 version, we conducted subgroup analyses as described below (Subgroup analysis and investigation of heterogeneity).

Again, for simplicity, we refer to core checklist items with abbreviated descriptions according to their 'Paper section and topic' as found on the CONSORT 2001 checklist. For example, when we refer to reporting of 'title and abstract' and/or 'item one', we are addressing whether reports of RCTs in evaluations contained "randomised" in the title or abstract. For full details of associated recommendations for these items (or more appropriately, method-

ological guidance) please see Table 1.

The four items reporting blinding stem from the 2001 CONSORT checklist item recommending that adequate reporting of blinding should detail "whether or not participants, those administering the interventions, and those assessing the outcomes were blinded to the assessment group". Reflective of the original systematic review (Plint 2006), included evaluations, and subsequent changes made to the CONSORT 2010 checklist, we collected reporting of blinding in four distinct items, in addition to the 'composite' item (i.e. blinding by any description) contained in the 2001 CONSORT checklist. These include: blinding of participants, blinding of the intervention, blinding of outcome assessment, and blinding of the data analyst. We sub-categorised analyses for this item, as described in the subgroup analysis section below.

While calculation of a total sum score based on several CONSORT items is potentially misleading as items are of unequal importance, we collected this information if reported in included evaluations.

We abstracted data on assessment of methodological quality of RCTs included within evaluations, if reported. Although a recent study (Dechartres 2011) identified 74 different items and 26 different scales used for assessing quality of RCTs, measurement of methodological quality using any of these means (e.g. Jadad score, Olivo 2008, Schulz allocation concealment, MINCIR, MINCIR Score) was considered and was not pre-specified for this review.

Validity assessment in included evaluations

The validity of included evaluations was assessed by one author (LS) and all assessments were independently verified by one of three authors (LT, AP, LW); we resolved all conflicts by consensus.

We assessed validity using an a priori checklist developed by the research team for the purpose of this review. As no formal checklist for assessing validity of quasi-experimental evaluations of RCTs currently exists, our research team developed a checklist based on principles of internal and external validity (Campbell 1966). We used the Data Collection Checklist developed by the Cochrane Effective Practice and Organisation of Care Review Group and the 'Risk of bias' tool as guides (Cochrane EPOC 2009; Higgins 2008). The resulting criteria used to gauge validity of evaluations in this review were as follows:

1. The RCTs included in the study represented a large cohort (i.e. at least an entire year), or were randomly chosen from a large cohort.
2. The reviewer(s) who assessed CONSORT criteria were blinded to study authors, institutions, sponsorship, and/or journal name.
3. Consideration of potential clustering by journal was reported (if potential for clustering did not exist, the study was deemed 'low risk').
4. There was no evidence of selective outcome reporting.
5. More than one reviewer assessed adherence to CONSORT criteria.
6. If more than one reviewer assessed CONSORT criteria, whether inter-reviewer agreement was greater than or equal to 90% agreement or a kappa statistic of 0.8.
7. If quality of included RCTs was assessed, the reviewer(s) conducted a blinded assessment.

We assigned each criterion a judgement of yes (i.e. low risk of bias/high validity), no (i.e. high risk of bias/low validity), or can't tell (unclear risk of bias). For some criteria, we allowed an additional rating of 'not applicable' if it was irrelevant to a given comparison, or was dependent on the rating of a previous criterion. For instance, there was no potential for clustering by journal (criterion 3) in comparisons 2 and 3. Criterion 6 was dependent on the rating for criterion '5' being 'yes' and therefore was not applicable when the rating was 'no'. Likewise, criterion 7 is dependent on whether assessment of methodological quality was carried out. For these three criteria (3, 6, and 7) we chose not to penalise evaluations with 'not applicable' ratings, nor to rate them as 'unclear', since this is taken to mean 'not reported', which is also incorrect. As such, the only remaining option which would not connote any negative judgement is a rating of 'yes' (i.e. low risk of bias/high validity).

Note, with regards to item three above, we report here the terms used when validity assessment was conducted. For clarification, from here on we refrain from using the term 'clustering' as this potential bias, more aptly, refers to confounding by journal.

Measures of the effect of the methods

Comparison 1 examines the completeness of reporting of RCTs published in endorsing and non-endorsing journals, comparison 2 examines the completeness of reporting of RCTs published in journals before and after endorsement, and comparison 3 examines completeness of reporting of RCTs before and after publication of CONSORT. Where data from a single evaluation were applicable to more than one comparison, the evaluation was included for each comparison. For instance, where data from an evaluation comparing endorsing and non-endorsing journals were available, it was sometimes possible to use data from only the endorsing journals to also compare the reporting before and after endorsement.

For the primary outcome, where data on completeness of reporting were represented by one or more of the 22 CONSORT 2001 checklist items or of the four additional blinding items, we collected dichotomised adherence to each item. Where evaluations used more than two categories to judge adherence to a given checklist item, we collapsed these to create a dichotomy between 'adequately' and 'inadequately' reported RCTs. For instance, where an item was judged as 'partially' reported, it was considered 'inadequate'. As such, within each comparison, for each dichotomous outcome, the proportion of RCTs within each evaluation ade-

quately reporting one or more checklist items in each comparison group was calculated. Using these proportions we compared completeness of reporting between comparison groups (i.e. endorsers versus non-endorsers, before versus after endorsement, pre versus post publication) in each evaluation using a risk ratio (RR) with a 99% confidence interval for each outcome. A RR greater than 1 was taken to indicate relatively increased reporting of any CONSORT item following CONSORT endorsement. Where completeness of reporting of RCTs was represented by a sum score of aggregate checklist items, we collected the mean sum score for each comparison group within an evaluation. We then calculated the standardised mean difference (SMD) with 99% confidence interval to estimate the difference in completeness of reporting between comparison groups in each evaluation. An SMD greater than 0 indicates better overall reporting of items following CONSORT endorsement.

Due to the design of included evaluations and poor availability of data to make necessary adjustments to estimates of effect at the evaluation level, we were unable to adjust for potential confounders (i.e. improvements in completeness of reporting over time and/or by discrepancies in journal editorial 'quality') and we introduced the use of 99% confidence intervals post hoc to ensure conservative estimates of effect are presented throughout this review.

Data collected on the methodological quality of RCTs within evaluations were reported as collected in evaluations. As these were expected to be variable and inconsistently reported across evaluations, we planned no measures of effect to estimate whether groups within each comparison differed on methodological quality.

Issues of potential confounding

There are two potential factors by which the estimates of effect obtained for each evaluation could be confounded. The first is when there may have been an uneven distribution of journal quality (defined in Data extraction and management) between endorsing and non-endorsing journals in comparison 1. Time is considered a second potential confounder of effect estimates for individual evaluations, since the completeness of reporting may have naturally changed over time with or without endorsement or publication of CONSORT. Time potentially affects effect estimates across all three comparisons of this review, however it is not considered a true confounder for comparison 1, since it may only play a role where comparison groups were sampled at different times. Please see Assessment of risk of bias in included studies.

Dealing with missing data

We experienced two types of missing data: endorsement status of journals included in evaluations and date of endorsement of journals determined to be endorsers by either authors of the evaluation or review authors.

Endorsement status of journals publishing RCTs included in each evaluation was needed to determine whether evaluations were eligible for inclusion in comparisons 1 or 2 or not at all. As described in the Selection of studies and Data extraction and management sections, we contacted corresponding authors a maximum of three times via email over an eight-month period to provide us with these data. If data would have been needed to complete the comparative analysis (i.e. adequacy of reporting data for each checklist item for each included RCT), these were requested at the same time.

Where date of endorsement of CONSORT by journals was not explicit, data for RCTs that subsequently could not be identified as published in either an endorsing or non-endorsing journal were not included in the analyses in order to prevent misclassification. In some circumstances, where this would result in a high proportion of data for a given evaluation being excluded, we categorised these reports as published in an endorsing journal, a conservative classification that underestimates the effect of CONSORT endorsement. Similarly, for before and after comparisons, when a number of evaluations were published in 2001 (or 1996, more infrequently), these evaluations would be classified as pre-CONSORT to ensure that any estimate of the effect of CONSORT endorsement would be conservative.

Assessment of heterogeneity

We explored consistency across the included evaluations quantitatively using the I^2 statistic, and by visual inspection (Deeks 2008). Variation in journal policy regarding how CONSORT is implemented, for example whether submission of a completed checklist is 'required' versus 'recommended', will likely contribute to methodological heterogeneity of results across included evaluations. However, ongoing research by the CONSORT group suggests that the means of implementing CONSORT in the editorial process is difficult to determine without speaking to journal editorial staff directly. As our experience with this review has shown, even when in contact, this information is vague and generally no standardised processes are in place. As it was beyond the scope and feasibility of the current review, we were unable to explore this factor meaningfully.

Assessment of reporting biases

Selective reporting of outcomes has been assessed for each included evaluation as a component of validity assessment (Appendix 2). We conducted assessment by searching for a review protocol and, in the absence of a protocol, compared methods and results sections of included evaluations. An advantage to the design of this review is that unpublished data are provided and included by evaluation authors, which would contribute to mitigating the potential issue of selective reporting of CONSORT items.

Although it is possible to generate funnel plots to assess the potential of publication bias for each meta-analysis in each evalua-

tion within included evaluations, the suitability of this method of assessment is unexplored (although the number of included studies may be insufficient). We know of no alternative methods for assessing publication bias in this review of evaluations of RCTs. Moreover, the number of included studies would frequently not allow for this; as such we are unable to determine any failure to report within the literature.

Data synthesis

We used a pooled RR with 99% confidence intervals to estimate the overall difference between groups within each comparison. We used a random-effects model for all analyses. All available data contributed to our main analyses.

Some evaluations totaled adherence to all or a subset of CONSORT checklist items, and reported averages over assessed RCTs. Because these continuous data are on differing scales, we calculated SMDs for this outcome, with 99% confidence intervals. When medians and ranges were reported instead of means and standard deviations, we used suitable approximations (Higgins 2008). When necessary, we imputed standard deviations.

Data from evaluations reporting on, and comparing, CONSORT-endorsing journals' adherence to items of methodological quality, using means not otherwise evaluated in this review, were qualitatively described and not included in meta-analysis.

In addition to our main analyses, we conducted a descriptive analysis of the included evaluations based on general characteristics of the evaluations. For example, we documented the number of RCTs and journals assessed in those evaluations and the validity of those evaluations.

Subgroup analysis and investigation of heterogeneity

We did not pre-specify any subgroups for analysis. However, post hoc, we decided that for five items of the 1996 checklist that underwent substantial modifications (i.e. re-arranging and wording modifications) in the 2001 checklist, analyses would be subgrouped by CONSORT checklist version (i.e. 1996 or 2001). These items are 'title and abstract', 'outcomes', 'sample size', 'participant flow', and 'numbers analysed'.

In addition, because data on adequacy of reporting of blinding were collected in five different outcomes in this review (as described in Data extraction and management), we sub-categorised meta-analyses for this item (blinding) by each of the five outcomes for which we collected data and carried out pooled estimates of effect within each subcategory.

Sensitivity analysis

As previously stated, when CONSORT endorsement status for a subset of journals in an evaluation was not available, we conducted a sensitivity analysis to compare the pooled risk ratios with journals that were and were not strictly compliant with our definition of a CONSORT endorser (i.e. endorsement occurred at least six months prior to publication of RCT). We also conducted sensitivity analysis for effects which we considered to result in outlying effect estimates when the forest plots were inspected.

Other methodological considerations

Review updating

Given the substantial number of new evaluations included in this update, we treated this update as if it were an original review following the original protocol. A full literature search was conducted from six months prior to the end search date of the original review (Plint 2006) to as recent a date as possible. We then screened all retrieved evaluations, at which point inclusion of the original eight evaluations was confirmed. We conducted data extraction for general characteristics, full data extraction, and validity assessment for all included evaluations in the same manner. We then compared data extracted for the original eight evaluations with the original published results as a means of validation. Data provided by authors and modified for inclusion in the original review were not sought again.

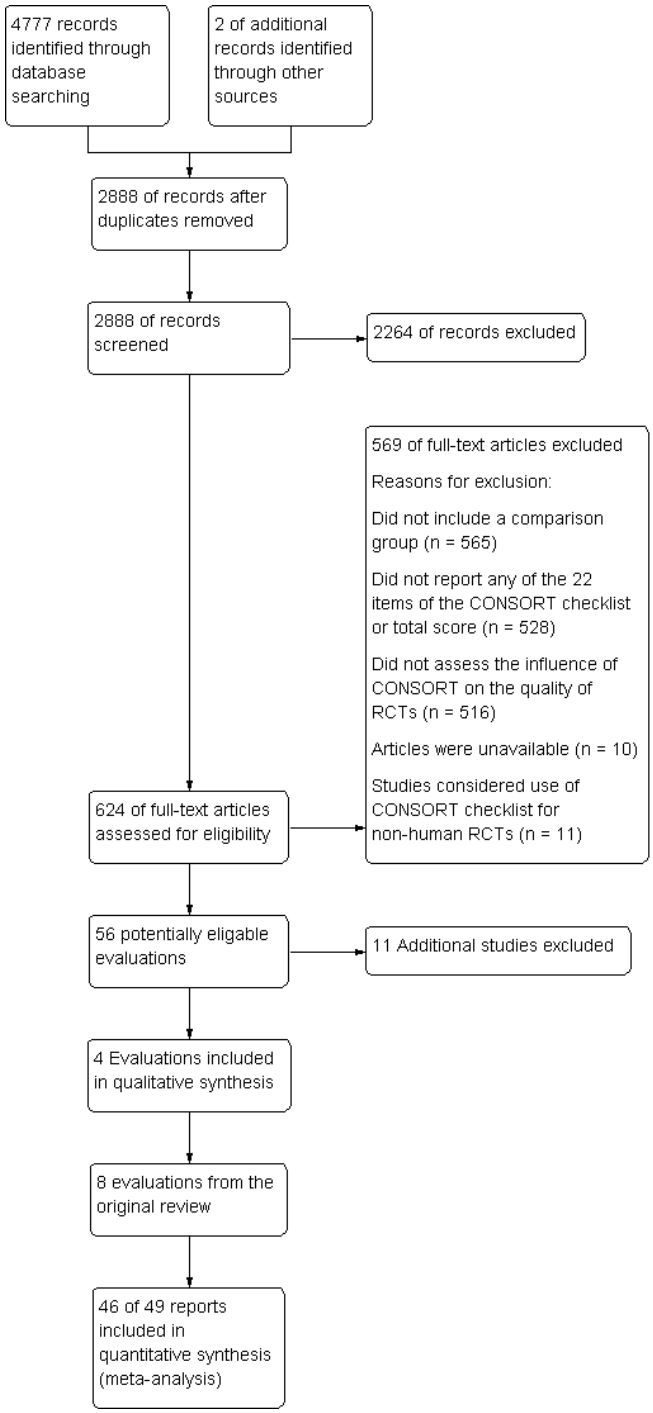
RESULTS

Description of studies

Results of the search

Our electronic search strategy identified a total of 4777 records. Two additional evaluations (Dickinson 2002; Kidwell 2001) were presented as posters at Cochrane Colloquia and identified by members of the research team. We removed duplicates and left the remaining 2888 records as potentially relevant articles. Details about the flow of evaluation records through this review are provided in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram (Moher 2009; Figure 1).

Figure 1. Flow of evaluations through this review



Content experts identified four evaluations before the search was conducted (Agha 2007; Peckitt 2007; Smith 2008; Wang 2007), all of which were also identified through the electronic search. No additional evaluations were identified by screening reference lists of eligible evaluations.

Included studies

After title and abstract screening, we retrieved and reviewed 624 full-text articles. Fifty evaluations, reported in 53 publications, were deemed eligible for inclusion (Figure 1).

We considered three pairs of evaluations to be potential multiple reports of each other as they reported outcomes from the same data set (Hopewell 2010 and Yu 2010; Balasubramanian 2006 and Tiruvoipati 2005; Spring 2007 and Thoma 2006). For 35 evaluations, additional data were needed to determine eligibility or to define the comparative analysis. Of 21 authors who responded, 20 were able to provide additional information to supplement the published data. Some of the information received from authors was not in the necessary format to allow inclusion in meta-analyses. In these cases, only data provided in the evaluation report were included in meta-analyses. One included evaluation was an abstract (Peckitt 2007), for which all necessary data were fully reported; a full-text article for this evaluation was not available at the time of data extraction. One evaluation (Dickinson 2002) was presented as a poster and was not published as a full article. For this evaluation, supplementary details were supplied by the evaluation author. All other included evaluations were journal publications. An additional author (Ellis 2005) provided data that confirmed that their evaluation was ineligible for inclusion.

The total number of included RCTs was 16,604 (median per evaluation (interquartile range, IQR) 123 (77 to 226)). Included evaluations reviewed RCTs published in a median of six (IQR 3 to 26) journals. Two evaluations reported on especially large numbers of RCTs (Hopewell 2010; Wang 2007), with 1135 and 7496 RCTs respectively.

Thirty-five included evaluations used CONSORT checklist items as a means of assessing completeness of reporting of RCTs within a given medical area, from which we could obtain information to form suitable comparisons. Seven evaluations did not list the influence of CONSORT or RCT adherence to the CONSORT checklist as primary or secondary outcomes, but assessed reporting on the basis of self determined methodological outcomes, consistent with the CONSORT checklist, which in turn allowed for a suitable comparison applicable to our review.

All included evaluations were published in English. Seventeen evaluations considered the influence of the 1996 CONSORT checklist, 25 reported data for the 2001 checklist, and the remaining eight evaluations considered outcomes from some form of modified CONSORT checklist. For example, Bian 2006 modified the CONSORT checklist suitable to their field of study or objectives. Forty-one evaluations addressed reporting quality by focusing on trials published within a specific medical field; these fields were broad and diverse, including, for example, behavioural health, urology, drug abuse, and anaesthesiology.

Some evaluations were eligible for more than one of our three comparisons and across the these comparisons, 29 evaluations were included in comparison 1 (CONSORT endorsers versus CONSORT non-endorsers), 11 evaluations were included in comparison 2 (CONSORT-endorsing journals, before and after endorsement), and 21 evaluations were included in comparison 3 (before and after CONSORT publication). Overall, 69 outcomes were quantitatively reported, across the three comparison groups (mean of eight outcomes reported per evaluation).

Evaluations used varying definitions for endorsement. Of the total number of included RCTs, 84% (13,955/16,604) across 85% (45/53) of evaluations were published in journals which endorsed CONSORT at least six months prior to RCT publication (as defined in Selection of studies).

Eight evaluations also assessed RCT quality by proxy, using means of assessing methodological quality; eight assessed quality using the Jadad Score (Jadad 1996); three assessed the completeness of reporting of allocation concealment; two used Schulz allocation concealment (Schulz 1995); and four used other scores or means of quality assessment (Effects of methods).

Excluded studies

We screened 2888 evaluations by title and abstract; we excluded 2264 evaluations as they did not assess completeness of reporting of RCTs. Of the remaining 56 included evaluations, we excluded a further 11 from the review at the data extraction phase due to unavailability of data (Excluded studies).

Risk of bias in included studies

Validity assessment of included studies

Overall, the rated validity of included evaluations was high or unclear (Figure 2; Figure 3). The majority of included evaluations had a large cohort, did not demonstrate selective reporting of outcomes, had more than one rater assessing CONSORT criteria and, if methodological quality was assessed using another tool, blinded assessments were performed. We note, however, that for this latter domain, as well as those pertaining to criteria 3 and 6 (as described in Assessment of risk of bias in included studies), a rating of high validity may appear as a potential overestimate of validity for a given evaluation.

Figure 2. 'Risk of bias' summary: review authors' judgements about each risk of bias item for each included study.

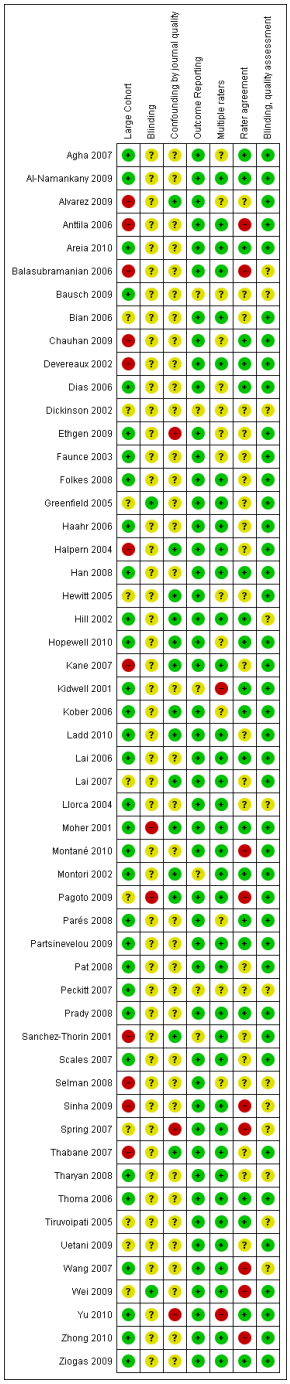
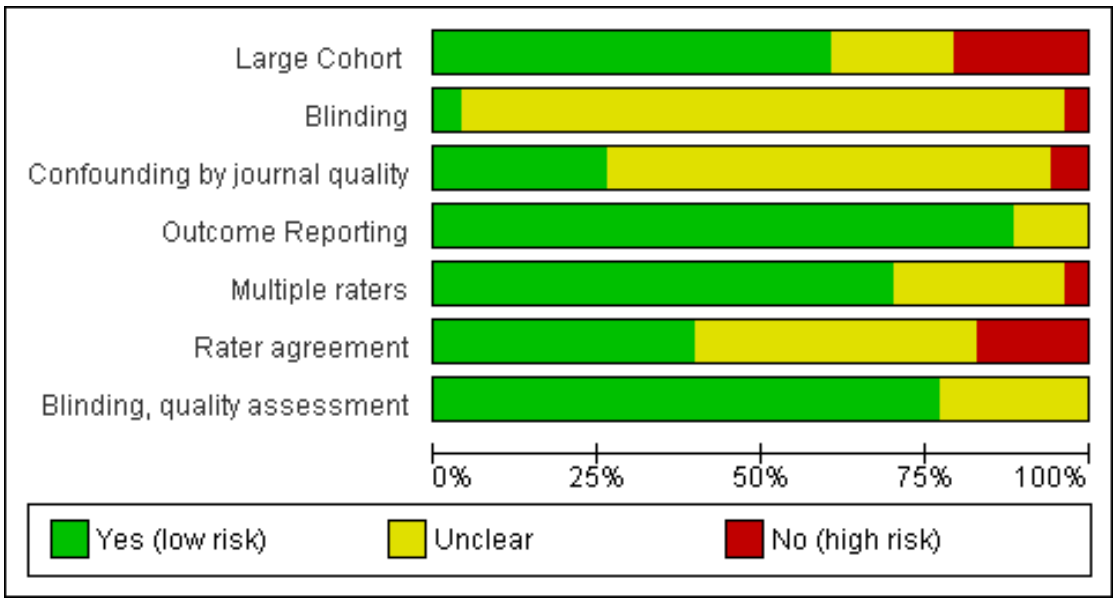


Figure 3. 'Risk of bias' graph: review authors' judgements about each risk of bias item presented as percentages across all included studies.



Across domains, we were uniformly unable to assess validity due to poor reporting of included evaluations, contributing to the large number of 'unclear' ratings. This 'unclear' rating also reflects the need for improvement in the validity assessment tool used in this review. For instance, whether or not confounding by journal occurred was difficult to assess since, for some evaluations, we used data provided by authors to create our own comparisons, thereby nullifying any adjustments for confounding that may have been carried out by authors. Moreover, as frequently discussed with regard to assessing quality of the RCT, the reporting of included evaluations may not reflect their actual conduct; however, information on many of our items was unobtainable from the text, which we thus rated 'unclear'.

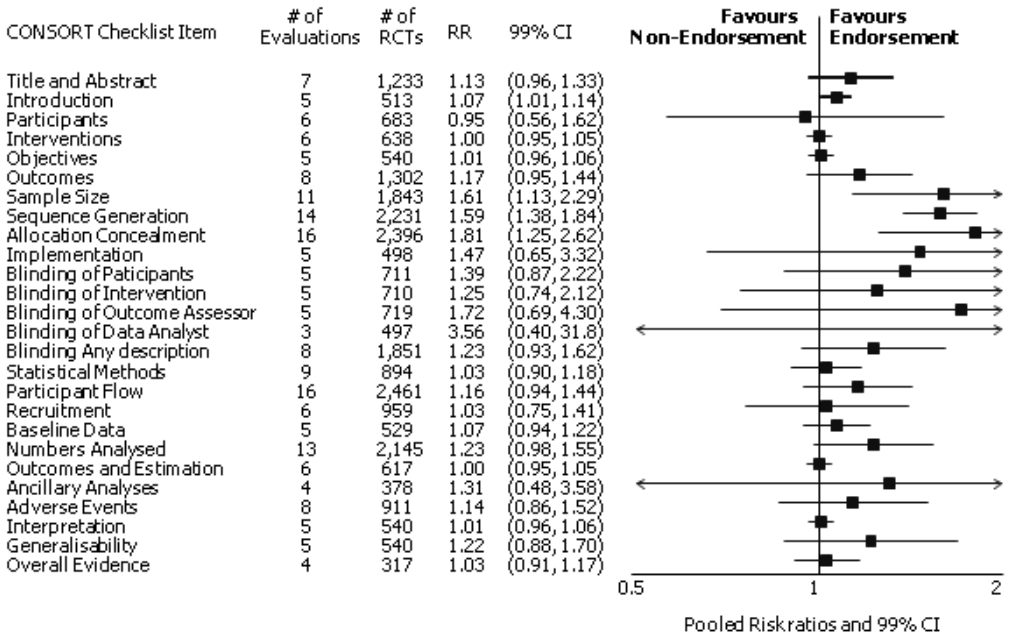
It is important to note that these evaluations were not randomised trials; less than 8% (4/53) of the evaluations reported adjusting for potential confounding factors, for evaluations that did not adjust for confounding (criterion 3), their estimates of effect may potentially be confounded by the natural improvement in completeness of reporting of RCTs over time, or by journal 'quality', as discussed above ('Issues of potential confounding').

Effect of methods

Comparison 1: Completeness of reporting of randomised controlled trials (RCTs) in CONSORT-endorsing journals versus non-endorsing journals

Twenty-nine evaluations were included in this comparison, with RCT level data for at least one of the 2001 CONSORT checklist items, blinding subcategories, or total sum score. Across 27 potential outcomes, the number of evaluations per meta-analysis varied (median (interquartile range, IQR) 6 (5 to 8)). 'Allocation concealment' and 'participant flow' were reported in the largest number of included evaluations: 16 each, with 2396 and 2140 assessed RCTs respectively. 'Ancillary analysis' and 'overall evidence' were reported in the fewest evaluations included in meta-analyses, with four evaluations each, that assessed 378 and 317 RCTs respectively. Results for all outcomes in this comparison are presented in Figure 4.

Figure 4. Pooled risk ratios across assessed 2001 CONSORT checklist items with 99% confidence intervals for primary comparison, adherence of RCTs published in CONSORT-endorsing journals versus RCTs published in CONSORT non-endorsing journalsPlot generated in Comprehensive Meta-analysis Version 2.0 (CMA).



For the 27 outcomes evaluated, five items resulted in statistically significantly more complete reporting in CONSORT-endorsing journals than non-endorsing journals, including complete reporting of: allocation concealment, description of scientific explanation and rationale in the 'Introduction', how 'sample size' was determined, and total sum score. Reporting details of adequate 'allocation concealment' had the largest estimate of effect(risk ratio (RR) 1.81, 99% confidence interval (CI) 1.25 to 2.61) (16 evaluations, 2396 RCTs, I^2 = 75%, Figure 5). For interpretation, this suggests an increase in adequate reporting of allocation concealment of between 25% and 161% in RCTs published in CONSORT-endorsing journals. Allocation concealment was reported adequately in 45% (393/876) of RCTs in CONSORT-endorsing journals and in 22% (329/1520) of RCTs in non-endorsing journals. For all other significant outcomes, which can be interpreted in a similar manner, results are as follows. Description of scientific explanation and rationale in the 'Introduction' was reported 7% more in CONSORT-endorsing journals than non-endorsing journals (RR 1.07, 99% CI 1.01 to 1.14) (five evaluations, 513 RCTs, I^2 = 0%, Figure 6). How 'sample size' was determined was reported between 13% and 129% more in RCTs of CONSORT-endorsing journals (RR 1.61, 99% CI 1.13 to 2.29) (11 evaluations, 1843 RCTs, I^2 = 76%, Figure 7). Description of the method used for 'sequence generation' was reported between 38% and 84% more in CONSORT-endorsing RCTs (RR 1.59, 99% CI 1.38 to 1.84) (14 evaluations, 2231 RCTs, I^2 = 24%, Figure 8). The 'total sum score' item resulted in a significant difference between endorsers and non-endorsers(standardised mean difference (SMD) 0.68, 99% CI 0.38 to 0.98) (seven evaluations, 560 RCTs, I^2 = 0%, Figure 9). This effect estimate suggests that the average reporting of items in RCTs in CONSORT-endorsing journals was more complete than for RCTs in CONSORT non-endorsing journals. For one evaluation (Kidwell 2001), standard deviations were not reported and were imputed from the values reported in other evaluations, using a weighted average.

Figure 5. Forest plot of comparison: 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals, outcome: 1.9 Allocation concealment.

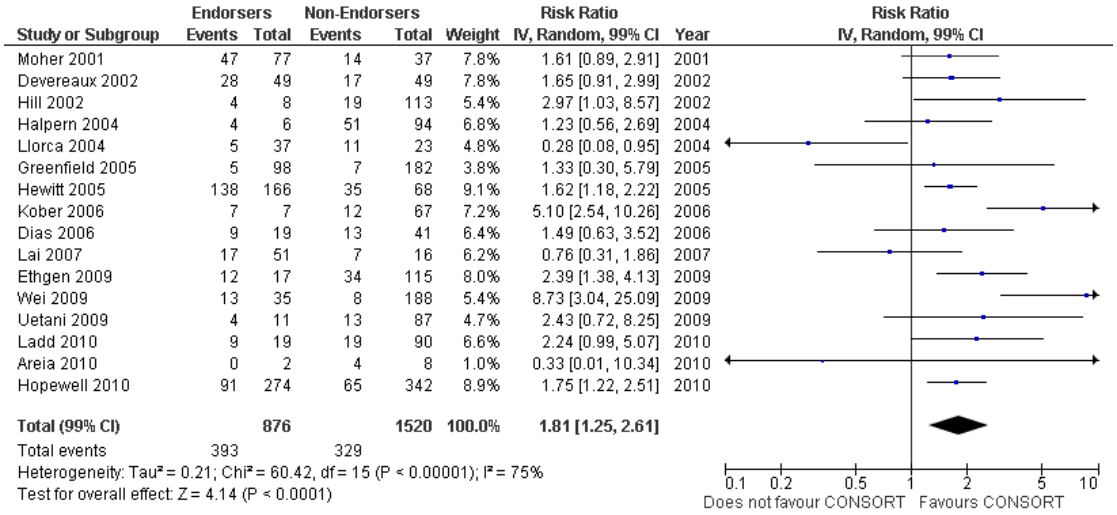


Figure 6. Forest plot of comparison: 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals, outcome: 1.2 Introduction.

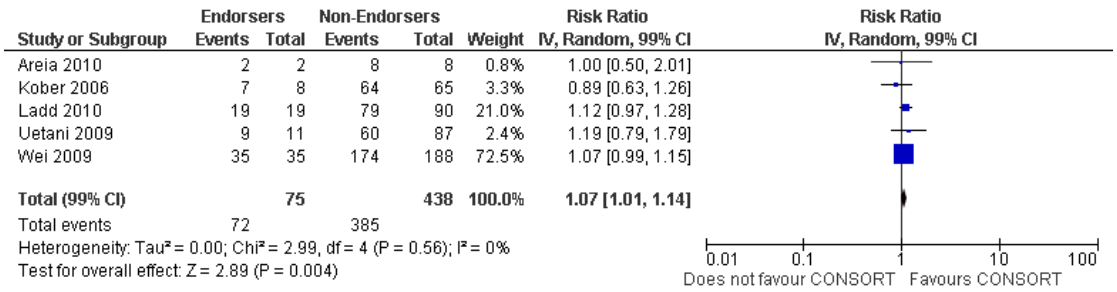


Figure 7. Forest plot of comparison: 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals, outcome: 1.7 Sample size.

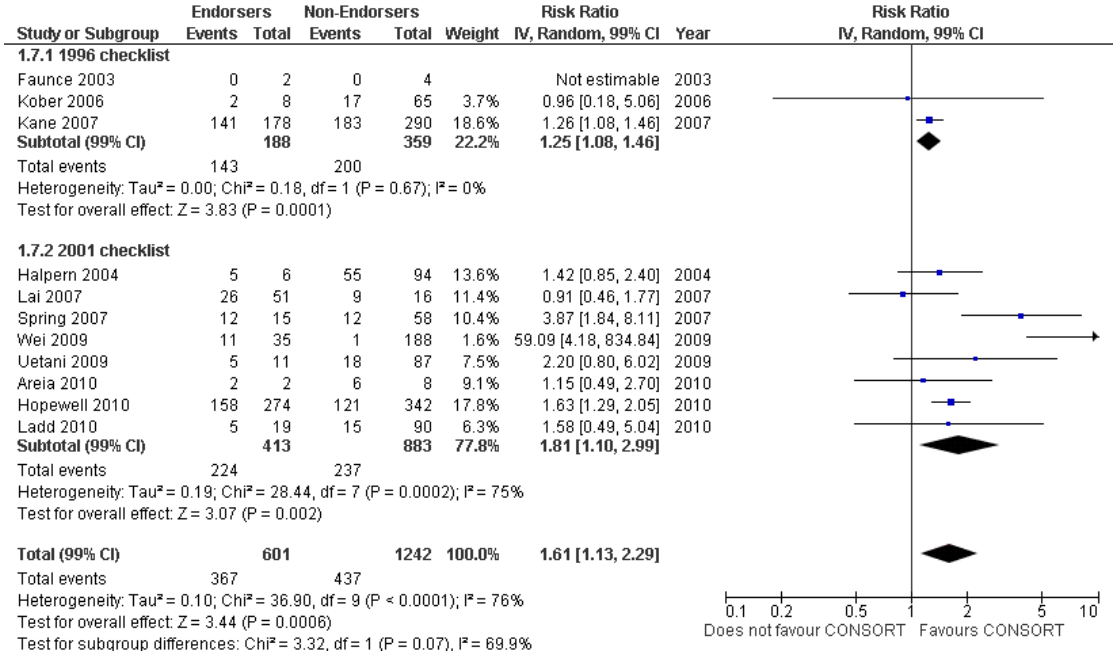


Figure 8. Forest plot of comparison: 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals, outcome: 1.8 Sequence generation.

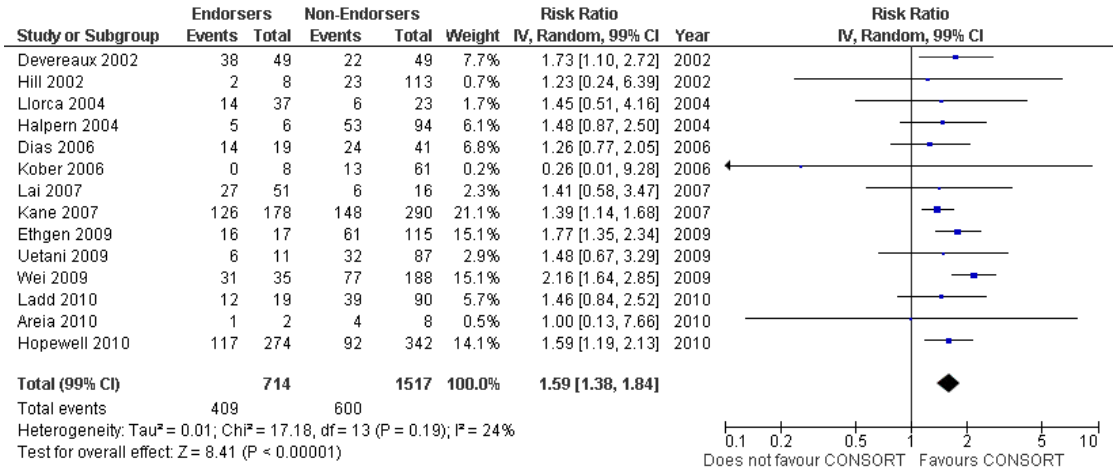
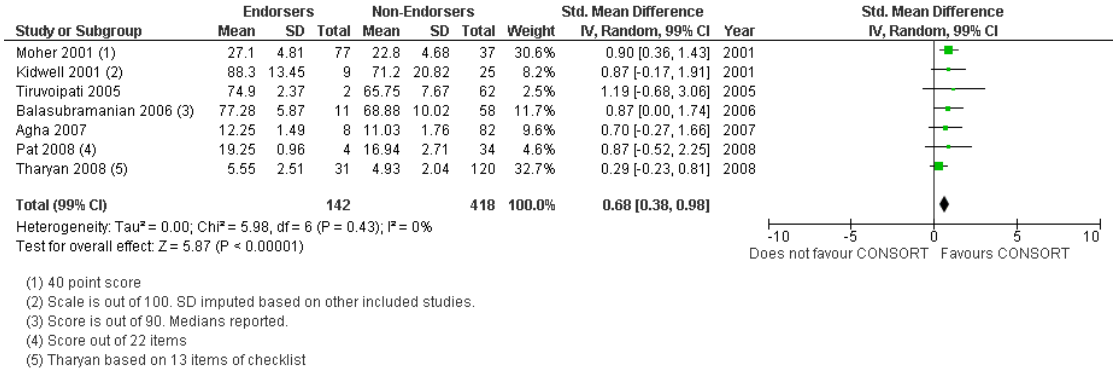


Figure 9. Forest plot of comparison: 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals, outcome: 1.23 Total sum score.



(1) 40 point score
(2) Scale is out of 100. SD imputed based on other included studies.
(3) Score is out of 90. Medians reported.
(4) Score out of 22 items
(5) Tharyan based on 13 items of checklist

For 20 of the 22 remaining outcomes, pooled estimates of effect showed reporting was more complete in a higher proportion of RCT reports for CONSORT-endorsing journals compared to non-endorsing journals (RR > 1.0), but these were not statistically significant. Precise details of 'interventions', item four, were equally well reported in endorsing and non-endorsing journals(RR 1.0, 99% CI 0.95 to 1.05) (six evaluations, 638 RCTs, I^2 = 0%), and eligibility criteria for 'participants', item three, produced a non-significant negative effect (RR 0.95, 99% CI 0.56 to 1.62) (six evaluations, 683 RCTs, I^2 = 91%).

Subgroups for CONSORT 1996 and 2001 checklists

All items resulted in estimates of effect larger in those evaluations assessing reporting using the 2001 checklist than those using the 1996 checklist. Determination of 'sample size' was reported significantly more in CONSORT-endorsing journals in evaluations assessing both the 1996 and 2001 CONSORT checklist versions. The completeness of reporting of 'participant flow' differs between 1996 and 2001 checklist versions. For 'title and abstract', 'outcomes', and 'numbers analysed' comparisons between endorsing and non-endorsing journals were all non-significant for both 1996 and 2001 subgroups. Complete reporting of how 'sample size' was determined yields significant results for CONSORT endorsers for evaluations adhering to either checklist version. This effect is greater in magnitude across evaluations assessing the 2001 checklist version, with RR 1.25 (99% CI 1.08 to 1.46) and RR 1.81 (99% CI 1.25 to 2.61) for 1996 evaluations and 2001 evaluations respectively, but these subgroups did not differ significantly (P = 0.07) (Figure 7). Complete reporting of 'participant flow' also increases in effect, with evaluations assessing the 1996 version, RR 1.01 (99% CI 0.99 to 1.02) and the 2001 evaluations, RR 1.35 (995 CI 1.00 to 1.82). Six evaluations were included in the 1996 subgroup and 10 evaluations in the 2001 subgroup; the latter considered inclusion of a flow diagram or otherwise to describe patient flow in 548 RCTs in CONSORT-endorsing journals and 1088 RCTs in CONSORT non-endorsing journals; testing for differences between subgroups demonstrates a statically significant difference between 1996 and 2001 checklist version groups (P = 0.01).

Complete reporting of randomisation in the 'title and abstract' was reviewed in one evaluation subject to the 1996 checklist, and six evaluations according to the 2001 checklist. Across all evaluations for this outcome, the pooled effect suggests an increase in reporting of 13% (RR 1.13, 99% CI 0.96 to 1.33). Estimates of effect did not differ greatly between checklist versions (P = 0.14), with effect estimates, RR 0.93 (99% CI 0.65 to 1.32) and RR 1.16 (99% CI 0.97 to 1.39) for 1996 and 2001 checklist versions respectively.

Overall, complete reporting of 'outcomes' is not significantly different in CONSORT-endorsing journals compared to non-endorsing(RR 1.17, 99% CI 0.95 to 1.43). The test for subgroup differences did not result in a difference between groups (P = 0.52), where one evaluation saw an effect of RR 1.02 (99% CI 0.58 to 1.78) in 1996 and seven evaluations saw an effect of RR1.18 (99% CI 0.94 to 1.48) in 2001.

Complete reporting of 'numbers analysed' did not differ between the 1996 and the 2001 checklist versions. Across all 13 evaluations in this outcome assessing 2145 RCTs, the estimate of effect was not significant(RR 1.23, 99% CI 0.98 to 1.55). The 1996 version evaluations did not yield more complete reporting in endorsers when pooled(RR 0.99, 99% CI 0.83 to 1.19). The magnitude of effect increases according to the 2001 checklist definition (RR 1.23, 99% CI 0.98to 1.55); testing for differences between subgroups suggests that assessments subject to the two versions differ (P = 0.03) (Figure 10; Figure 11; Figure 12; Figure 13).

Figure 10. Forest plot of comparison: 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals, outcome: 1.13 Participant flow.

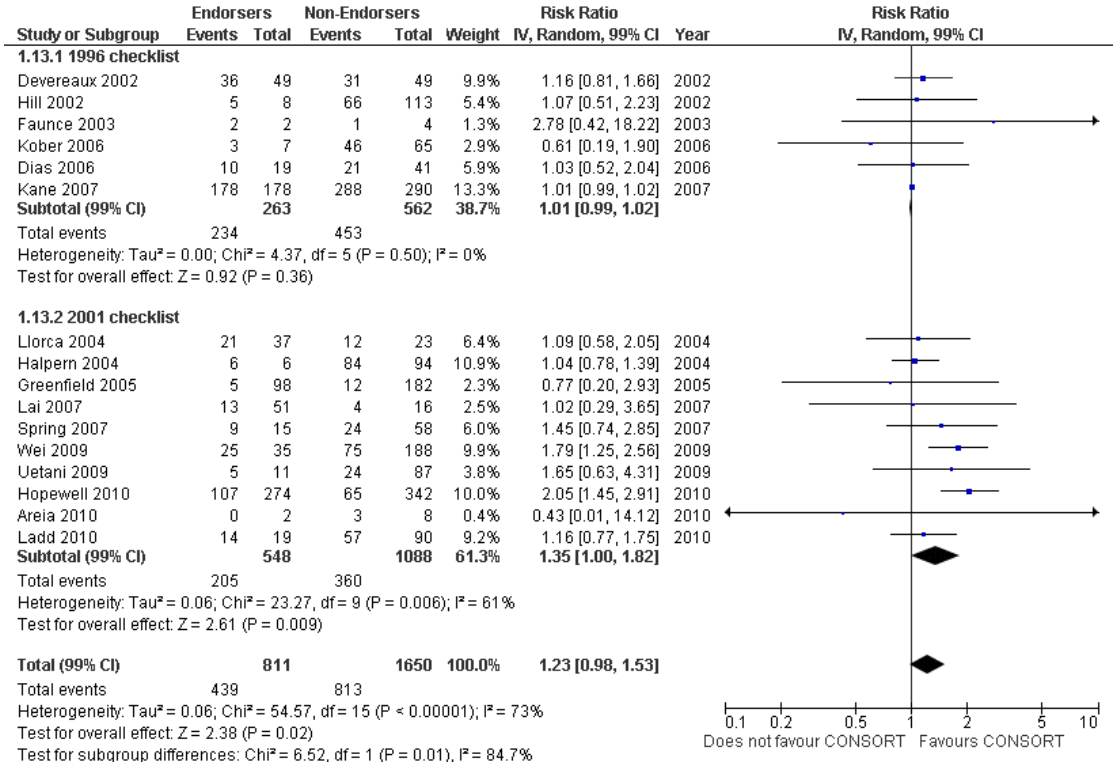


Figure 11. Forest plot of comparison: 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals, outcome: 1.1 Title and abstract.

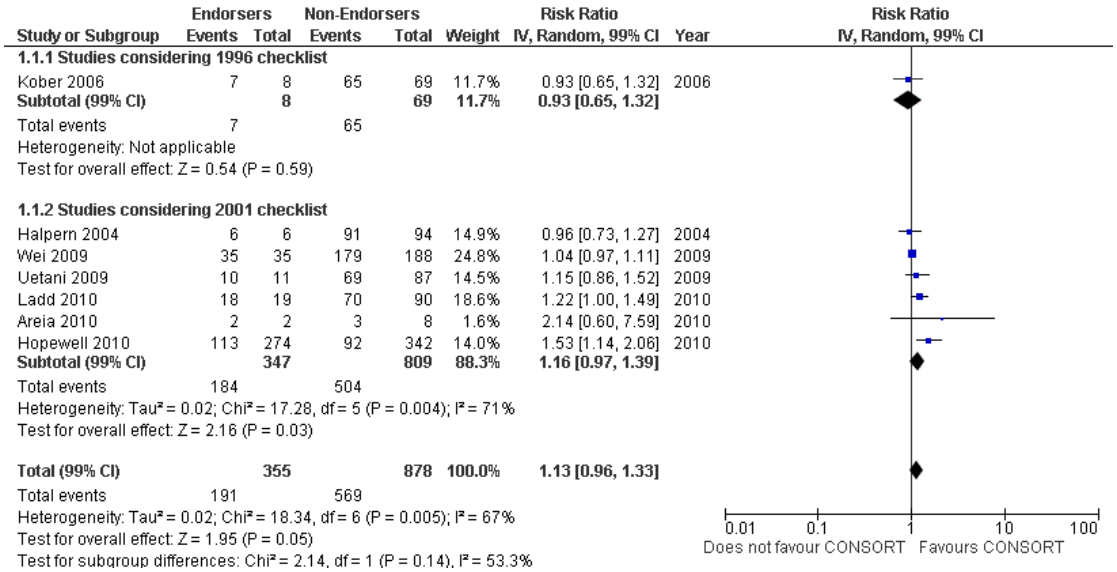


Figure 12. Forest plot of comparison: 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals, outcome: 1.6 Outcomes.

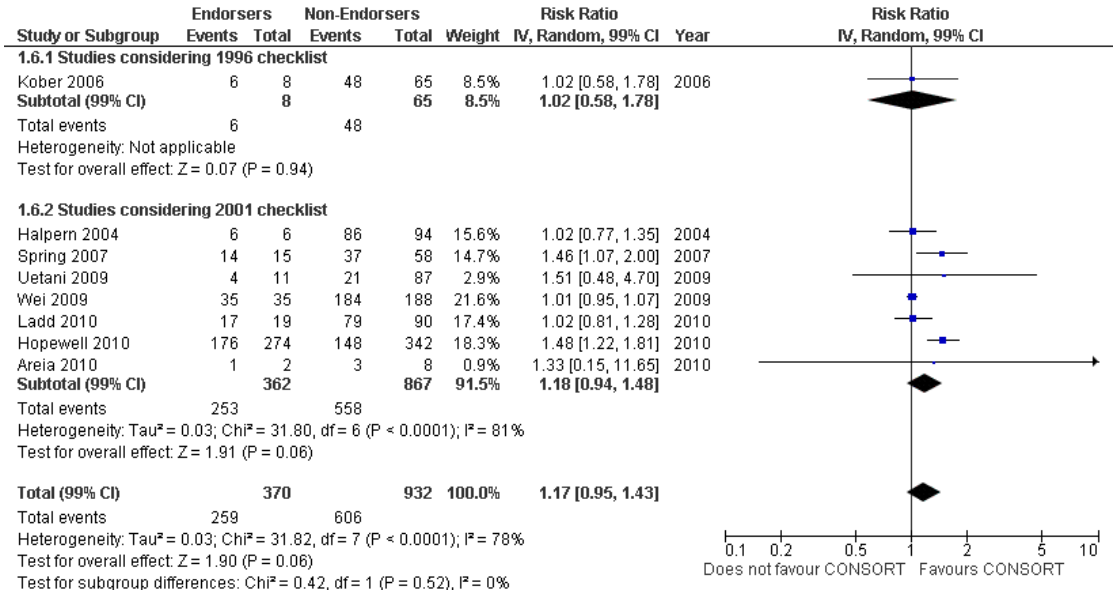
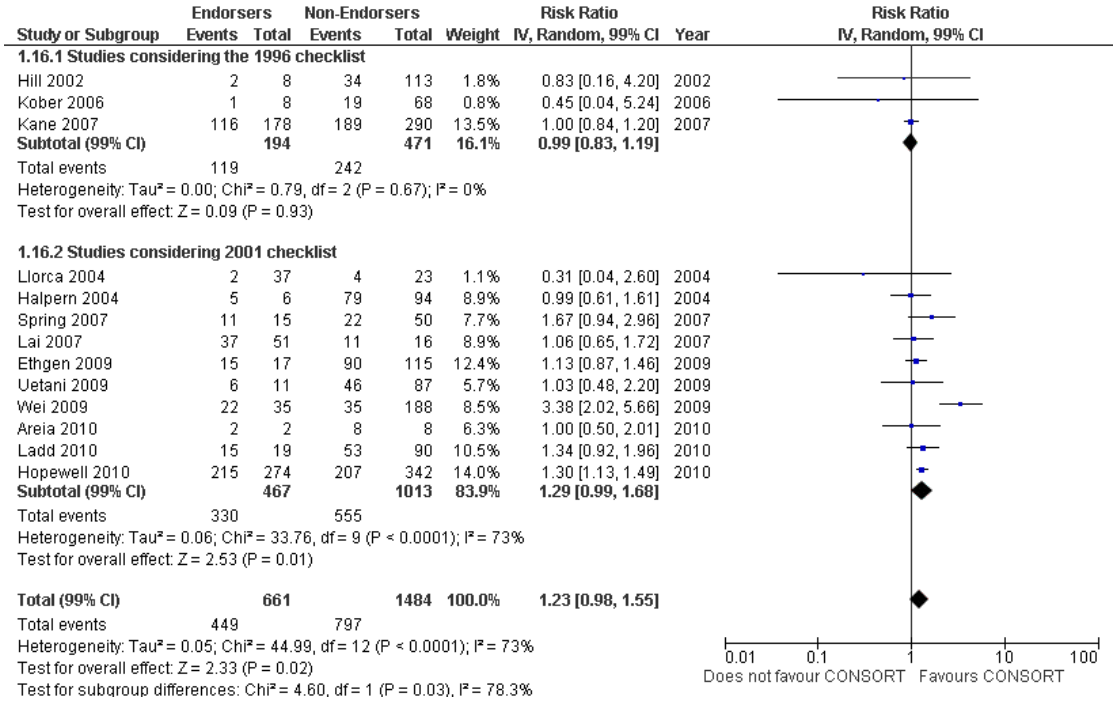


Figure 13. Forest plot of comparison: 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals, outcome: 1.16 Numbers analysed.



Sensitivity analysis

Eight included evaluations (Ethgen 2009; Hopewell 2010; Kidwell 2001; Tharyan 2008; Tiruvoipati 2005; Uetani 2009; Wei 2009; Yu 2010) were not strictly compliant with our definition of CONSORT-endorsing journal (Objectives). Of these eight evaluations, three did not report how a CONSORT endorser was defined, one evaluation categorised endorsing journals as those listed on the CONSORT website, and the remaining four referred to the online journal 'Instructions to authors' to determine if RCTs in a given journal were associated with a journal that endorsed the CONSORT checklist. Although this met our definition for how journal endorsement information is obtained, it does not confirm the date of publication of each assessed RCT as six months prior to the publishing journal endorsing CONSORT; therefore it has not been confirmed that the journal was endorsing CONSORT at the time of manuscript writing. It is important to note that, for all known definitions, such misclassification would lead to underestimates of the relative effect of adherence to the CONSORT items by RCTs in journals which endorse the CONSORT Statement.

We conducted sensitivity analysis across outcomes, excluding the above mentioned evaluations that did not strictly meet our definition of CONSORT endorsement. Only 1/27 outcomes, although only minimally different, differed substantially when evaluations that did not directly meet our definition of endorsement were excluded. Completeness of reporting of the 'Introduction' changed from RR 1.07 (99% CI 1.01 to 1.14) to 1.05 (99% CI 0.87 to 1.27). This suggests that relaxing our criteria for CONSORT endorsement did not alter substantially the estimates for reporting of RCTs published in non-endorsing journals versus those published in journals endorsing CONSORT.

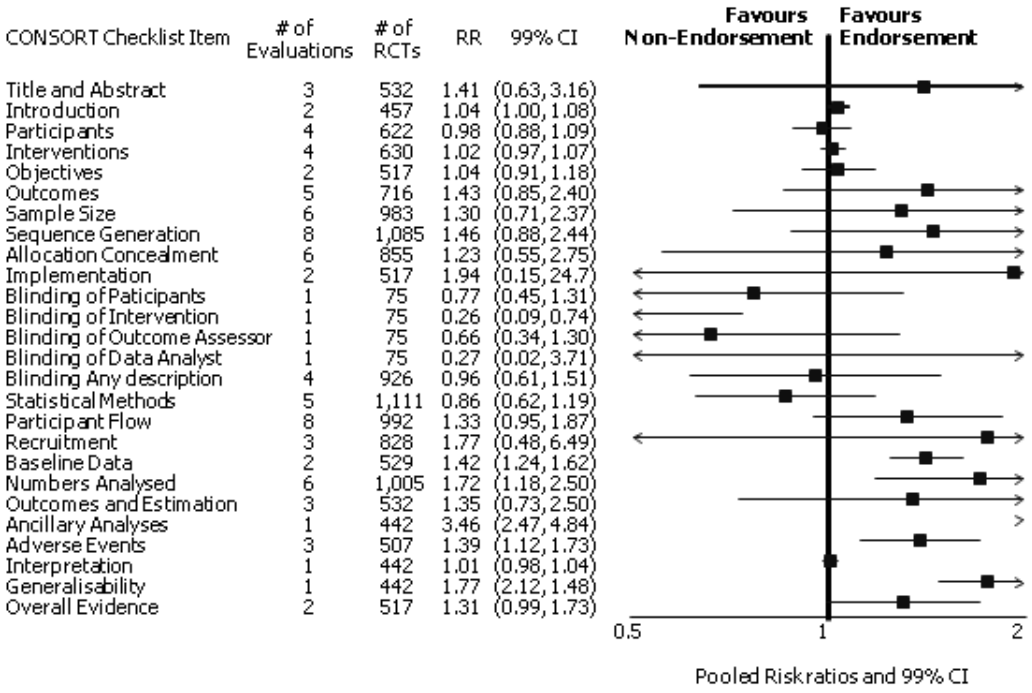
In addition, we considered several point estimates large outliers and we examined these in sensitivity analyses. These include: 'statistical methods', item 12, reported in the Areia 2010 evaluation; 'blinding of data analyst' in the Devereaux 2002 evaluation; 'participants', item three in the Faunce 2003 evaluation; 'blinding of outcome assessor' in Haahr 2006; and 'sample size and allocation concealment' in Wei 2009. Sensitivity analyses excluding these evaluations did not change the significance of completeness of reporting items in RCTs in CONSORT-endorsing journals compared with RCTs published in CONSORT non-endorsing journals.

Comparison 2: Completeness of reporting of RCTs in CONSORT-endorsing journals before and after

endorsement

Eleven evaluations assessed only journals that endorse the CONSORT Statement, but presented RCT completeness of reporting of at least one CONSORT item before and after the journal's date of endorsement of CONSORT. The number of RCTs assessed per outcome had a median (IQR) of 532 (512 to 919). The number of reported CONSORT checklist items varied over evaluations, with a median of 3 (IQR 2 to 5). 'Sequence generation' and 'participant flow' were both reported in eight evaluations. For 15 of 27 outcomes data were reported in fewer than five evaluations. The results across all outcomes in this comparison are presented in Figure 14.

Figure 14. Pooled risk ratios across assessed 2001 CONSORT checklist items with 99% confidence intervals for comparison 2, adherence of RCTs published in CONSORT-endorsing journals before and after endorsement.Plot generated in Comprehensive Meta-analysis Version 2.0 (CMA).



Seven outcomes resulted in statistically significantly more complete reporting in journals after CONSORT endorsement. These include: complete reporting of the scientific rationale and background in the 'Introduction' (RR 1.04, 99% CI 1.00 to 1.08) (two evaluations, 457 RCTs, I^2 = 0%); 'baseline data' (RR 1.42, 99% CI 1.24 to 1.62) (two evaluations, 529 RCTs, I^2 = 0%); 'numbers analysed' (RR 1.72, 99% CI 1.18 to 2.49) (six evaluations, 1005 RCTs, I^2 =76%); 'ancillary analyses' (RR 3.46, 99% CI 2.47 to 4.84) (one evaluation, 442 RCTs); 'adverse events' (RR 1.39, 99% CI 1.12 to 1.73) (three evaluations, 507 RCTs, I^2 = 0%); and 'generalisability' (RR 1.77, 99% CI 1.47 to 2.11) (one evaluation, 442 RCTs). Aggregate scores of items were also significant for this comparison: the total sum score was SMD 0.74 (99% CI 0.30 to

1.18) (one evaluation, 148 RCTs).

Of the remaining outcomes, 13/20 resulted in pooled estimates of effect showing that reporting was more complete in a higher proportion of trial reports for CONSORT-endorsing journals compared to non-endorsing (RR > 1.0), but these were not statistically significant. Overall, completeness of reporting was not optimal either before or after endorsement, even when results have demonstrated a difference when journals have endorsed the statement. For example, only 76% (428/560) of RCTs published after journal endorsement of CONSORT and 38% (171/445) of RCTs published before completely reported 'numbers analysed' as per the CONSORT consolidated standards of reporting trials (CONSORT) and the completeness of reporting guidance.

For seven items, estimates of effect showed less complete reporting in RCTs published in journals after endorsement of CONSORT, but none of the differences were statistically significant. These outcomes include complete reporting of eligibility criteria for participants (RR 0.98, 99% CI 0.88 to 1.09) (four evaluations, 622 RCTs, I^2 = 28%) and complete reporting of statistical methods used (RR 0.86, 99% CI 0.62 to 1.19) (five evaluations, 1111 RCTs, I^2 = 90%). Across all possible blinding subgroups, the relative reporting of blinding decreased in RCTs in CONSORT-endorsing journals after endorsement. Blinding of interventions was reported in one evaluation of 75 RCTs, indicating that reporting is significantly reduced post endorsement (RR 0.26, 99% CI 0.09 to 0.73) (one evaluation, 75 RCTs). All subgroups reflected larger reductions in reporting than the blinding (any description) item, which is considered to be most consistent with the 2001 checklist(RR 0.96, 99% CI 0.61 to 1.50) (four evaluations, 926 RCTs, I^2 = 95%). All blinding subgroups were evaluated by one evaluation assessing 75 RCTs. For all blinding outcomes, RCTs in CONSORT-endorsing journals post endorsement were found to report blinding less completely than in RCTs of CONSORT non-endorsing journals.

Subgroup analyses for 1996 and 2001 checklist version

There were no statistically significant tests for differences in subgroups for the five identified outcomes. Three items saw effects of greater magnitude in the 2001 checklist version group, and two outcomes saw greater effects in the 1996 groups. Three items, the 'title and abstract', 'sample size', and 'numbers analysed' checklist items, were completely reported significantly more in CONSORT-endorsing journals than non-endorsing journals in both subgroups. Despite an increase in effect estimates from 1996 to 2001 checklist versions, 'title and abstract' subgroups did not differ significantly (P = 0.42). There was no statistically significant difference between 'sample size' subgroups, despite the 2001 checklist increasing the magnitude of the effect estimate (P = 0.67). Nor was there a difference between subgroups when assessing 'numbers analysed' (P = 0.26). Two items, 'participant flow' and 'outcomes' had larger effect estimates across evaluations assessing the 1996 checklist version, but neither of these groups differed significantly when subgroups were tested.

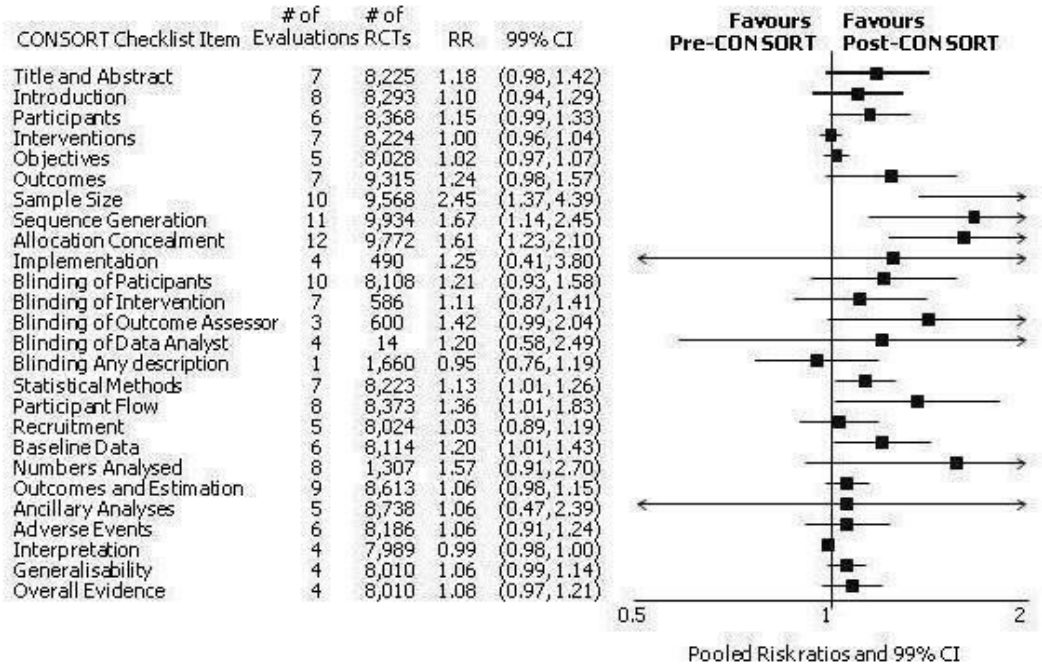
Sensitivity analysis

Two evaluations over three outcomes were considered for sensitivity analyses due to relatively large effects. The Sanchez-Thorin 2001 evaluation reported relatively large effects in favour of CONSORT endorsement for reporting the CONSORT items 'outcomes' and 'participant flow', however, the comparisons remained non-significant at the 1% level when this evaluation was excluded. The Han 2008 evaluation is one of two evaluations reporting on generated and assigned sequence allocation, namely 'implementation'. This evaluation reported a relatively large effect; excluding this evaluation did not change the overall significance of effect for this item.

Comparison 3: Completeness of reporting of RCTs before and after CONSORT publication

This comparison was developed due to the large body of evidence that did not comply fully with our definition of endorsement. Although these data were abundant and consistent with the findings of the other comparisons, evaluations in this comparison did not comply with our prespecified definition of within-journal endorsement (see Objectives). As such the findings may not be as robust and should be interpreted cautiously. The results across all outcomes for this comparison are presented in Figure 15.

Figure 15. Cross-sectional sample of RCTs before and after the publication of CONSORT.



Twenty-one evaluations provided comparisons of completeness of reporting compliant with the CONSORT checklist items, before and after either the 1996 or 2001 publication of CONSORT. Methods for assessing the pre-post intervention were inconsistent across evaluations. Over all outcomes, there were on average 7 (5 to 8) (median, (IQR)) evaluations per checklist item, with an average of 8224 (8017 to 8676) (median (IQR) RCTs per outcome (CONSORT item). 'Allocation concealment' was reported in the largest number of included evaluations: 12 evaluations assessed reporting adherence in 9772 trials.

Six outcomes saw statistically significant results, suggesting that these items were statistically significantly more completely reported after the publication of the CONSORT Statement. These include complete reporting of 'sample size' (RR 2.45, 99% CI 1.37 to 4.39) (10 evaluations, 9568 RCTs, I^2 = 91%), 'sequence generations (RR 1.67, 99% CI 1.14 to 2.45) (11 evaluations, 9934 RCTs, I^2 = 79%), 'allocation concealment' (RR 1.61, 99% CI 1.23 to 2.10) (11 evaluations, 9772 RCTs, I^2 = 13%), 'statistical methods' (RR 1.13, 99% CI 1.01 to 1.25) (seven evaluations, 8223 RCTs, I^2 = 67%), 'participant flow' (RR 1.36, 99% CI 1.01 to 1.83) (eight evaluations, 8373 RCTs, I^2 = 72%), and 'baseline data' (RR 1.20, 99% CI 1.01 to 1.43) (six evaluations, 8114 RCTs, I^2 = 47%).

Of the 21 remaining outcomes, 18 showed completeness of reporting was higher in RCTs published after CONSORT, but the differences were not significantly significant. Complete reporting of the 'intervention' resulted in a neutral effect(RR 1.00, 99% CI 0.97 to 1.04) (seven evaluations, 8224 RCTs, I^2 = 7%) and 'interpretation of the results' had a pooled effect which did not favour the impact of CONSORT on the completeness of reporting(RR 0.99, 99% CI 0.98 to 1.01) (four evaluations, 7989 RCTs, I^2 = 0%).

All subcategories of blinding descriptions resulted in higher proportions of RCTs completely reporting, but the difference before and after publication of CONSORT was not significant. Evaluations providing analyses of any description of blinding showed that fewer RCTs reported a complete description of blinding after the publication of CONSORT (RR 0.95, 99% CI 0.76 to 1.19) (three evaluations, 1660 RCTs, I^2 = 0%). Complete reporting was infrequent for both groups, for example, in total less than 18% (1041/5891) post CONSORT publication RCTs, and less than 9% (345/4043) of pre-CONSORT RCTs, completely report their method of 'sequence generation' as per the CONSORT guidance.

Subgroup analyses for 1996 and 2001 checklist versions

There were no differences between subgroup analyses for the five outcomes specified. Subgroup analyses effect estimates for com-

plete reporting of randomisation in the 'title and abstract' were consistent: the 1996 version saw a relative increase in adequate reporting of 13% (RR 1.13, 99% CI 0.96 to 1.33), while the 2001 version saw a relative increase of 18%(RR 1.18, 99% CI 0.88 to 1.59); the difference between these two groups was not significant (P = 0.73). Complete reporting of derivation of 'sample size' was reported more frequently in assessed RCTs post CONSORT publication, with significant results for both checklist versions assessed. Evaluations considering the 2001 version of the checklist produced a larger pooled effect, suggesting that the percentage of RCTs published after publication of the 2001 CONSORT Statement reporting 'sample size' was greater than those RCTs published before 2001 (RR 2.68, 99% CI 1.00 to 7.16). There was no statistical difference between these groups (P = 0.90). Adequate reporting of 'participant flow' in RCTs published after the publication of the CONSORT Statement saw a larger improvement in evaluations considering the 2001 version of the checklist as the intervention, with 2.14 times more RCTs adequately reporting the flow of participants through the trial (RR 2.14, 99% CI 0.90 to 5.09) than those considering the 1996 evaluation where only 1.16 times more RCTs adequately reported 'participant flow' (RR 1.16, 99% CI 0.87 to 1.53); these differences were not statistically significant (P = 0.08).

Reporting of primary and secondary 'outcomes' saw a greater magnitude of effect across those evaluations assessing the 1996 version(RR 1.47, 99% CI 0.87 to 2.48 and RR 1.15, 99% CI 0.85 to 1.54 for the 1996 and 2001 versions respectively); this difference between subgroups was not significant (P = 0.29). Adequate description of the 'numbers analysed' was non-significantly relatively more frequent in RCTs published after the CONSORT Statement, for both subgroups of evaluations considering the 2001 version and the 1996 version(RR 1.37, 99% CI 0.80 to 2.36 and RR 2.32, 99% CI 0.50 to 10.87 respectively). Over all evaluations, there was a non-significant 57% increase in adequate reporting of denominators for the number of participants analysed in RCTs published after than before the publication of the CONSORT Statement (RR 1.57, 99% CI 0.91 to 2.70); subgroup differences between checklist versions were not significant (P = 0.41).

Sensitivity analysis

The third comparison group was developed to synthesise results of cross-sectional samples of RCTs before and after CONSORT publication, as well as evaluations for which timing of endorsement of CONSORT could not be confirmed as the intervention within journals. As a result, all included evaluations in this comparison have been confirmed to have RCTs pre- and post CONSORT publication of the CONSORT Statement. No sensitivity analysis could be conducted in relation confirmation of endorsement.

Five evaluations (Parés 2008; Partsinevelou 2009; Peckitt 2007; Scales 2007; Wang 2007) report effects that were relatively large. As such we performed sensitivity analyses to assess the difference in pooled effects when these evaluations were not included.

Across all outcomes, evaluations with large effects were not included in pooled effect estimates and discrepancies were observed. Peckitt 2007 and Wang 2007 were simultaneously excluded from the 'sample size' outcome, with a reduction in effect from RR 2.45 (99% CI 1.37 to 4.39) to RR 1.80 (99% CI 1.10 to 2.93). Parés 2008 and Scales 2007 were simultaneously removed from the 'participant flow' outcome, with a reduction from RR 1.36 (99% CI 1.01 to 1.83) to RR 1.20 (99% CI 0.95 to 1.50). When the Partsinevelou 2009 results were removed from the reporting of dates for the 'recruitment' outcome, the effect remained non-significant; and from the adequacy of reporting of which 'numbers [were] analysed' (RR 1.57, 99% CI 0.91 to 2.70 to RR 1.52, 99% CI 0.88 to 2.61).

Qualitative reports on the influence of reporting

Four evaluations that met inclusion criteria were not included in the three quantitative comparisons for this review (Al-Namankany 2009; Chauhan 2009; Montané 2010; Sinha 2009). Relatively few trials were assessed in these reports (n = 305 RCTs). Each provided qualitative descriptions of the influence of endorsement of CONSORT on the completeness of reporting, as detailed below. Three of the four evaluations reported that there was no difference in reporting subject to CONSORT endorsement.

Al-Namankany 2009 aimed to assess the reporting of published RCTs in paediatric dental journals between 1985 and 2006, and to assess whether completeness of reporting had improved since the introduction of CONSORT as a secondary outcome. Although data for inclusion in meta-analysis in this review were not available, the evaluation reported that "overall quality of reporting has not substantially improved since the publication of CONSORT".

The Chauhan 2009 evaluation modified the CONSORT checklist to 50 outcomes to assess the quality of obstetric practice bulletins after the publication of the 1996 CONSORT Statement. The results were not reported or provided upon request, leaving insufficient information for quantitative inclusion in our review. An interesting finding of the evaluation was that regressions conducted to determine if a number of variables could predict reporting based on CONSORT criteria resulted in only multicentre trials proving to be significant, suggesting that for this sample of RCTs completeness of reporting was 'better' in trials conducted in multiple centres. Another result of the evaluation is that even for the RCTs published after the CONSORT Statement, the adherence is variable and lacking at times. This evaluation reported finding no difference before and after publication of the 1996 Statement.

Montané 2010 assessed reports of RCTs assessing analgesics in postoperative pain after traumatic or orthopaedic surgery. The quality of reports was assessed using the CONSORT checklist (scoring range from 0 to 22). The publication year and the impact factor of journals were recorded, but we were unable to obtain

additional information for quantitative inclusion in this review. The authors reported a comparison over time: "The mean (SD) CONSORT scores for RCTs published after 2001 was higher than the mean CONSORT scores for those published previously (14.4 and 10.3 respectively; p<0.0001)."

Sinha 2009 used the Jadad score and eight other methodological items (sequence generation, allocation concealment, implementation of randomisation, blinding status of outcomes, blinding of data analysts, sample size, numbers analysed, and participant flow diagram) to assess quality of reporting in high impact factor surgical journal RCTs, and compared the quality of RCTs from CONSORT-endorsing journals with non-endorsers. In a sample of 42 RCTs, they observed: "There was no significant difference in the number of high-quality RCTs published in CONSORT-endorsing journals compared with non endorsers. This difference did not reach statistical significance suggesting that CONSORT endorsement by surgical journals does not appear to increase quality of reporting, although our study might not be adequately powered to detect such a difference because only one of the three journals studied did not endorse CONSORT".

Other means used to assess the influence of CONSORT on the quality of trials

Seven evaluations assessed the completeness of reporting using CONSORT checklist items in conjunction with another means of assessment. An additional evaluation considered the influence of endorsement of the CONSORT Statement by comparing Jadad scores only. All eight evaluations assessed quality by Jadad score; three evaluations also assessed quality by clear or unclear reporting of allocation concealment (attributed to Schulz); four evaluations also assessed quality by another means, namely, using the MINCIR score, 'quality score', modified Chalmers Score or 'Analytic Quality Elements' score. Four evaluations were of pre-post design with the publication of the CONSORT Statement as the intervention, three evaluations compared CONSORT endorsers and CONSORT non-endorsers, and the final evaluation considered both pre-post and post intervention designs.

Four evaluations compared quality of reporting between CONSORT endorsers and non-endorsers, or listed sufficient data to draw this comparison. Two of the four evaluations (Sinha 2009; Tiruvoipati 2005) found no significant difference between Jadad scores for RCTs in endorsing and non-endorsing journals, where the median for both groups of both evaluations was reported to be 2.0. Aggregate assessments were made in a total of 76 trials in non-endorsing journals and 25 trials in endorsing journals. Two evaluations (Balasubramanian 2006; Tharyan 2008), with a total of 220 RCTs, reported differences in Jadad score means of 0.27 and 0.20 respectively, between RCTs published in endorsing and non-endorsing journals. The mean scores were higher in endorsing journal publications, but these results were not significant.

Four evaluations assessed the pre-post influence on RCT quality according to CONSORT items, as well as the Jadad score. One evaluation did not provide sufficient data or description for comparison of RCT quality according to the Jadad score, before and after publication of CONSORT. The remaining three evaluations reported that there was a difference in quality, assessed by the Jadad score, of RCTs published before and after the publication of CONSORT. Moher 2001 detailed that "Over time, 3 of the 4 journals improved the quality of reports of RCTs as assessed by the Jadad scale, which was statistically significant for 1 journal (Lancet) and across the adopter journals pre-CONSORT, 2.7; mean change, 0.4; 95% CI, 0.1-0.8)." In a total of 2380 trials, Wang 2007 reported the mean (SD) Jadad score was 0.85 (0.53) in 1999 (746 RCTs) and 1.20 (0.62) in 2004 (1634 RCTs); and Parés 2008 reported a median (range) of 3 (0 to 5) in 2001 and after and 2 (0 to 4) before 2001, P = 0.046.

Sufficient reporting of allocation concealment was considered in three evaluations (Balasubramanian 2006; Moher 2001; Tiruvoipati 2005), the first of which did not provide enough information to abstract this data. The two evaluations that could be compared quantitatively suggest the difference in the Jadad scores of RCTs published before and after the endorsement of CONSORT was significant. Tiruvoipati 2005 reported 21% of RCTs with adequate reporting of allocation concealment pre-CONSORT and 50% in RCTs published after the Statement. Similarly, Moher 2001 describes "the proportion of RCTs with unclear reporting of allocation concealment decreased over time in all 4 journals and was statistically significant for adopter journals (pre-CONSORT, 61%; mean change, −22%; 95% CI, −38% to −6%)."

Four evaluations assessed quality of the included RCTs using an author-developed tool or assessment scale. Two evaluations assessed quality, but did not categorise this in relation to RCTs published in CONSORT-endorsing and non-endorsing journals. These evaluations reported RCT quality to a modified Chalmers score and an 'analytic quality elements score' developed for the paper. MINCIR is a methodological scaling tool consisting of three domains with subcategories, where a sum across the three outcomes can total between six and 36 'points'. The Parés 2008 evaluation reported significant differences in the quality of 40 RCTs subject to a MINCIR score assessment, between pre- and post CONSORT-endorsing journals, pre and post respectively, mean (range), 19 (13 to 25), 23 (13 to 36) P = 0.016. Llorca 2004 assessing 37 RCTs, also developed a 'quality score' to assess RCT quality, with a maximum score of 21. No significant differences in scores were found between RCTs published before and after the publication of the CONSORT Statement. For both groups mean scores were < 5/21.

DISCUSSION

Summary of main results

A substantial number of new evaluations have been published and were eligible for inclusion in this review since the last search in July 2005 and the publication of the original systematic review (Plint 2006). We included 50 quasi-experimental evaluations in 53 evaluation reports, examining 16,604 reports of randomised controlled trials (RCTs) in this update; eight evaluations were included in the original CONSORT systematic review. Across the 50 evaluations, a mean of eight CONSORT items were reported. Across the three comparisons, 29 evaluations were included in comparison 1 (CONSORT endorsers versus non-endorsers), 11 evaluations were included in comparison 2 (CONSORT-endorsing journals, before and after endorsement), and 21 evaluations were included in comparison 3 (before and after CONSORT publication).

The number of evaluations per meta-analysis (median (interquartile range, IQR)) were: comparison 1, 6 (5 to 8), comparison 2, 3 (2 to 5), and comparison 3, 7 (5 to 8). Overall, the results demonstrate an improvement in the completeness of reporting when journals endorse the CONSORT Statement. These results are consistent across the three comparison groups, with the exception of outcomes related to blinding, which are inconsistent (Figure 4; Figure 14; Figure 15).

For comparison 1, five of 27 outcomes pertaining to CONSORT items were found to be significantly more completely reported in studies published in CONSORT-endorsing journals than in non-endorsing journals: 'allocation concealment', 'introduction', 'sample size', 'sequence generation', and 'total sum score'. While not statistically significant, completeness of reporting for 18 items favoured CONSORT endorsement. Endorsement was not found to be beneficial for two outcomes (non statistically significantly less complete reporting): 'participants' and 'interventions'. We consider comparison 1 to be the most robust comparison in this review, because it is closest to the RCT design since it compares an intervention (endorsement) to a control (non-endorsement) in a cross-section of time. Within comparison 2, six of 27 outcomes evaluated had estimates of effect demonstrating significant improvement in reporting following CONSORT endorsement: 'introduction', 'baseline data', 'numbers analysed', 'ancillary analyses', 'adverse events', 'generalisability', and 'total sum score'. In contrast to comparison 1, comparison 2 included few evaluations per meta-analysis.

For comparison 3, six of 27 outcomes pertaining to CONSORT items demonstrate statistically significant improvement in reporting following the publication of CONSORT in both 1996 and 2001: 'sample size', 'sequence generation', 'allocation concealment', 'statistical methods', 'participant flow', and 'baseline data'. Completeness of reporting for all other items demonstrated non-significant improvements following publication of the CONSORT Statement.

Quality of the evidence

Like the first review on this topic in 2006 (Plint 2006), assessment of validity of included evaluations indicates that weaknesses regarding the design of evaluations still exist and there remains considerable room for improvement in the quality of the evidence base. Across evaluations, we were uniformly unable to appraise validity due to unclear reporting of methods and findings by evaluation authors; this resulted in largely unclear ratings across all pre-specified domains (Figure 2). This 'unclear' rating may also reflect the need for improvement, validation, and standardisation in a tool to assess aspects of quality (i.e. validity) in future methodological reviews. For instance, whether or not included evaluations determined whether RCTs were clustered within journals of better or worse 'quality' in each comparison arm was assessed in item 3 of our validity assessment, but because we sometimes artificially created comparison arms where none existed, for the purpose of this review, this item can not be interpreted as an informative measure of validity of included evaluations.

None of the eight evaluations included in the original review or the 45 additional included evaluations were prospective in nature. An experimental design such as an RCT, arguably the strongest design that could be used, would help to control for many confounding variables, such as improvement due to the passage of time and variable editorial policies across journals. Such an RCT might target non-endorsing journals with an endorsement 'intervention' that might include a request to endorse CONSORT, evidence of its impact (i.e. the results of this review) and offer explicit wording to insert in a journal's 'Instructions to Authors'; with a control group not receiving any intervention. Future evaluations of the impact of endorsement of CONSORT (or other reporting guidelines) on completeness of reporting should utilise methodologically stronger designs than have been used to date, such as rigorous experimental designs.

Agreements and disagreements with other studies or reviews

Five of 22 items of the 2001 CONSORT checklist were significantly better when endorsement was present and similar positive effects are exhibited for another 15 items. However, there is no evidence to suggest that use of the CONSORT checklist is associated with reduced completeness of reporting of RCTs for some checklist items (i.e. reporting eligibility criteria for 'participants', risk ratio (RR) 0.95, 99% CI 0.56 to 1.62). The findings of this review are consistent with several other evaluations including the original review (Plint 2006) and the two largest evaluations included in this review (Hopewell 2010; Wang 2007).

This update extends the results reported by Plint and colleagues, which is the only previous systematic review of evaluations of the CONSORT checklist. The Plint 2006 review included eight evaluations. The main results demonstrated that CONSORT en-

dorsers had significantly better reporting of the method of sequence generation (RR 1.67, 95% CI 1.19 to 2.33), allocation concealment (RR 1.66, 95% CI 1.37 to 2.00), and overall number of CONSORT items (standardised mean difference 0.83, 95% CI 0.46 to 1.19) than non-endorsers. CONSORT endorsement had a weaker association with participant flow and blinding of participants. For before and after endorsement evaluations, good reporting of sequence generation, participant flow, and total CONSORT items were all associated with the endorsement of CONSORT.

Although our review uses confidence intervals at the 1% significance level (compared to the original review, which used 5%), all but one of the significant results in the original review remained statistically significant in this review; sequence generation was no longer significant for the before and after endorsement evaluations (RR 1.46, 99% CI 0.99 to 2.16). In addition to all other outcomes remaining significant, where there were sparse data per outcome in the original review, the inclusion of results of additional evaluations has seen that additional outcomes (title and abstract, introduction, sample size, participant flow, numbers analysed for endorsing versus non-endorsing journals, and introduction, baseline data, numbers analysed, ancillary analyses, adverse events, generalisability, and overall evidence) have all been influenced when comparing endorsing journals before and after endorsement.

Wang 2007 aimed to assess the quality of Traditional Chinese Medicine (TCM) RCTs published in 13 journals in mainland China, and assessed 20/22 items of the CONSORT checklist in 7422 trials. This evaluation was included as a non-strict comparison before and after the publication of the 2001 version of the CONSORT Statement as we were unable to verify whether all journals were endorsing, with corresponding dates of endorsement to classify each of the 7422 RCTs. Of the 20 items, 13 outcomes resulted in statistically significant effects for higher completeness of reporting after CONSORT publication. These were consistent with the six results deemed significant over all evaluations in this review, and with the findings of the original review.

Hopewell 2010 assessed quality of reporting of trials and also directly compared trials before and after the publication of CONSORT in 2001. This evaluation, assessing 1135 RCTs, was eligible for inclusion in both CONSORT-endorsing versus CONSORT non-endorsing, and pre-post publication of CONSORT comparison groups. Significant increases between 2000 and 2006 in the proportion of trial reports that included details of the primary outcome, sample size calculation, and the methods of random sequence generation and allocation concealment were reported. All of these were found to be significant in this review, for the comparison of completeness of reporting before and after the publication of the CONSORT Statement. Moreover, comparing RCTs of endorsing and non-endorsing journals, reporting of "randomised" in the title and abstract, reporting of the primary outcome, sample size calculation, sequence generation, allocation concealment, blinding, participant flow, and loss to follow-up all yielded significant increases in reporting for CONSORT-endorsing journals. All of these were also found to be significant results when comparing endorsing and non-endorsing journals in this review.

This review assessed the impact of endorsement of CONSORT by biomedical journals, however, evaluations assessing adherence to CONSORT (i.e. not just endorsement) may provide more meaningful insight into its impact on completeness of reporting when used at different stages of the editorial process. One such evaluation, carried out recently (Cobo 2011), incorporated these concepts by comparing use and non-use of reporting guidelines (including CONSORT) during peer review on author-revised manuscript quality. Findings indicate that manuscript quality was higher following peer review using reporting guidelines, including CONSORT.

This review is, itself, reported following the recommendations of the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) Statement (Moher 2009).

Limitations

While the CONSORT checklist aims to provide guidance on a minimum set of items to be reported in trials, during the review process we noted that what constitutes 'complete' reporting for each checklist item appeared to be variable between evaluations, depending on author interpretation. While it is the intention of the CONSORT group that complete reporting of a single checklist item means that all concepts contained within an individual checklist item be reported in order to be considered adequately (or completely) reported, some authors may have considered reporting as complete when at least one concept was reported in a given RCT; whether or not this was done is not identifiable or quantifiable in this review. For some evaluations, authors were more explicit in their interpretation of what constituted complete reporting by including ratings of 'partially reported'; for the purposes of this review, we took 'partial' ratings as 'incomplete'. We recommend that future evaluations assess the completeness of reporting of each checklist item in a dichotomous fashion (i.e. 'complete' versus 'incomplete') and moreover generally suggest to trial authors that items are only 'complete' when adhered to in their entirety.

This review does not assess the most current version of the CONSORT checklist (Moher 2010). To address problems with interpretation of checklist items, when the CONSORT Statement was revised in 2010, some items of the 2001 checklist that covered multiple concepts were purposefully split out into two or more sub-items. For instance, item 3 of the 2001 checklist, which addresses both the reporting of participant eligibility criteria and setting and location where data were collected, became two sub-items in the 2010 checklist (items 4a and 4b). When comparing RCTs published in endorsing and non-endorsing journals, two items 'participants' and 'interventions', although not statistically significant, resulted in effects which did not favour the endorse-

ment of CONSORT. These items have since been divided into two sub-items in the 2010 CONSORT Statement.

The objectives and methods of included evaluations varied considerably. Specification of items assessed pertaining to methods differed from evaluation to evaluation, some of which did not coincide specifically with CONSORT items. For example, some checklists used in evaluations to assess completeness of reporting contained modifications to the native wording of the CONSORT checklist(s) and/or sub-categorised items (and these modifications differed across evaluations); some evaluations assessed additional methodological items. Although we consider these aspects to have had little impact on the overall results, they were, nevertheless, a limitation. Moreover, data that were excluded to prevent potential misclassification as described in the Dealing with missing data section, or RCTs published in 2001 for which endorsement status could not be confirmed and thus were classified as non-endorsers, should be noted. All such classifications were made to ensure that any effect was underestimated rather than inflated.

Within the included evaluations, only four reported data regarding potential confounders. Some evaluations considered broad time intervals over which completeness of reporting was assessed, before and after CONSORT publication, or within endorsing journals before and after endorsement. Unfortunately, as there is insufficient information to adjust for confounding by improvement in reporting quality over time, and as this potential confounding factor impacts results at the evaluation level, we were unable to adjust for it. Similarly, confounding by journal quality, addressing whether those journals that endorse the CONSORT Statement are perhaps of higher 'quality' than those that do not, should also be considered when assessing our findings. This aspect was assessed for each evaluation, with results detailed in the validity assessment tables. Validity assessment was conducted based on pre-specified criteria developed specifically for this review. In particular, some items of the tool are more rigorous than others and quality assessment results should be interpreted cautiously, in particular, there is no evidence to suggest that blinding of assessors to trialists and institutions would improve the validity of the evaluations in this study.

One practical and important implication that could not be assessed when designing or carrying out this review, was the level at which the endorsement of CONSORT was implemented. This review assessed endorsement of the CONSORT Statement at the journal level, but not all journals may enforce CONSORT endorsement in the same manner, which could lead to a different impact of CONSORT on completeness of reporting. As suggested by Cobo et al (Cobo 2011), when and how CONSORT is implemented within the editorial process and who takes responsibility for ensuring adherence to CONSORT policies could impact on RCT reporting. It is reasonable to expect that a recommendation in a journal's 'Instructions to Authors' without any further editorial checks might have less of an impact on completeness of reporting as compared to a requirement to complete a CONSORT checklist and/or flow diagram before a manuscript is considered for peer review. These issues require further prospective study to understand better the impact of CONSORT on completeness of reporting. In the absence of such understanding, however, we believe our handling of reporting data in this review has resulted in an underestimate of the impact of the CONSORT Statement on the completeness of reporting. When authors do not adhere to a journal's recommendations to use CONSORT, endorsement does not achieve its full potential. Alternatively, some journals may not endorse the CONSORT Statement, but authors may use the checklist under their own volition. Again, this would result in an underestimate of the impact of the CONSORT Statement on the completeness of reporting.

It should be noted that comparisons 1 and 3 yield results in favour of CONSORT endorsement for the 'total sum score' item. This result is inclusive of evaluations that reported mean data for RCT adherence over all checklist items. Such scores give equal weighting to all checklist items, which may not be appropriate, although there is no sound basis on which to use unequal weights. Additionally, some evaluations scored an aggregate over CONSORT Statement modifications and included more or fewer than the 22 recommended items. In addition, one evaluation did not report all necessary data for inclusion in meta-analyses (e.g. median and range rather than mean, or not reporting standard deviations). For one evaluation the standard deviation was imputed. These differences between evaluations present some challenges when interpreting the significance of results for total sum scores.

AUTHORS' CONCLUSIONS

Implication for methodological research

While it is gratifying that approximately 600 health journals endorse CONSORT, this is still only a small proportion of all journals in existence. Even among those journals that mention CONSORT, the data suggest that there is considerable room for improvement in how it is endorsed. Hopewell and colleagues (Hopewell 2008) examined the 'Instructions to Authors' in 165 journals for any mention of CONSORT. These researchers observed that 38% mentioned CONSORT, although the language used varied across journals. This figure is an improvement on the 22% reported by Altman a few years earlier (Altman 2005).

We need to better understand barriers and facilitators to introducing CONSORT to the editorial process, and to develop and evaluate different implementation strategies that will increase CONSORT endorsement and adherence. The CONSORT group is currently undertaking further explorations in this area.

ACKNOWLEDGEMENTS

We thank Raymond Daniel and Andra Morrison for developing the search strategy, and Erin Crouchman for her assistance with screening and determining CONSORT endorsement status of many journals.

We would like to thank all authors who provided supplemental data for inclusion in this review. In particular, we thank Andy Vail and Roseanne NcNamee, co-authors of the Dias 2006 paper, who provided a lot of support to enable us to include their data in a suitable format for our review. Dr. Sven Trelle also provided support to ensure the data published in Kober 2006 could be included in this review. Dr. Javier Llorca, also provided extensive data. We also thank Drs. John Hill, Mette Haahr, Heather Dickinson, Stephanie Prady, Carle Paul, and Sally Hopewell who also made contributions.

We would also like to thank Jordi Pardo Pardo, Tinghua Zhang, and Taixiang Wu who provided screening, abstraction, or translation assistance.

REFERENCES

References to studies included in this review

Agha 2007 *{published and unpublished data}*
Agha R, Cooper D, Muir G. The reporting quality of randomised controlled trials in surgery: a systematic review. *International Journal of Surgery* 2007;**5**(6):413–22.

Al-Namankany 2009 *{published and unpublished data}*
Al-Namankany AA, Ashley P, Moles DR, Parekh S. Assessment of the quality of reporting of randomized clinical trials in paediatric dentistry journals. *International Journal of Paediatric Dentistry* 2009;**19**(5):318–24.

Alvarez 2009 *{published and unpublished data}*
Alvarez F, Meyer N, Gourraud PA, Paul C. CONSORT adoption and quality of reporting of randomized controlled trials: a systematic analysis in two dermatology journals. *British Journal of Dermatology* 2009;**161**(5):1159–65.

Anttila 2006 *{published and unpublished data}*
Anttila H, Malmivaara A, Kunz R, Autti-Ramo I, Makela M. Quality of reporting of randomized, controlled trials in cerebral palsy. *Pediatrics* 2006;**117**(6):2222–30.

Areia 2010 *{published and unpublished data}*
Areia M, Soares M, nis-Ribeiro M. Quality reporting of endoscopic diagnostic studies in gastrointestinal journals: where do we stand on the use of the STARD and CONSORT statements?. *Endoscopy* 2010;**42**(2):138–47.

Balasubramanian 2006 *{published and unpublished data}*
Balasubramanian SP, Wiener M, Alshameeri Z, Tiruvoipati R, Elbourne D, Reed MW. Standards of reporting of randomized controlled trials in general surgery: can we do better?. *Annals of Surgery* 2006;**244**(5):663–7.

Bausch 2009 *{published and unpublished data}*
Bausch B, Spaar A, Kleijnen J, Puhan MA. Quality of randomised trials in COPD. *European Respiratory Journal* 2009;**34**(5):1060–5.

Bian 2006 *{published and unpublished data}*
Bian ZX, Moher D, Dagenais S, Li YP, Wu TX, Liu L, et al. Improving the quality of randomized controlled trials in Chinese herbal medicine, part IV: applying a revised CONSORT checklist to measure reporting quality. *Zhong Xi Yi Jie He Xue Bao/Journal of Chinese Integrative Medicine* 2006;**4**(3):233–42.

Chauhan 2009 *{published and unpublished data}*
Chauhan SP, Berghella V, Sanderson M, Siddiqui D, Hendrix NW, Magann EF. Randomized clinical trials behind level A recommendations in obstetric practice bulletins: compliance with CONSORT statement. *American Journal of Perinatology* 2009;**26**(1):69–80.

Devereaux 2002 *{published and unpublished data}*
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. *Controlled Clinical Trials* 2002;**23**(4):380–8.

Dias 2006 *{published and unpublished data}*
Dias S, McNamee R, Vail A. Evidence of improving quality of reporting of randomized controlled trials in subfertility. *Human Reproduction* 2006;**21**(10):2617–27.

Dickinson 2002 *{published and unpublished data}*
Dickinson H, Lock K, Lecouturier J, Campbell F. Quality of trial of lifestyle interventions. Cochrane Colloquium Poster 2006.

Ethgen 2009 *{published and unpublished data}*
Ethgen M, Boutron L, Steg PG, Roy C, Ravaud P. Quality of reporting internal and external validity data from randomized controlled trials evaluating stents for percutaneous coronary intervention. *BMC Medical Research Methodology* 2009;**9**:24.

Faunce 2003 *{published and unpublished data}*
Faunce TA, Buckley NA. Of consents and CONSORTs: reporting ethics, law, and human rights in RCTs involving monitored overdose of healthy volunteers pre and post the "CONSORT" guidelines. *Journal of Toxicology - Clinical Toxicology* 2003;**41**(2):93–9.

Folkes 2008 *{published and unpublished data}*
Folkes A, Urquhart R, Grunfeld E. Are leading medical journals following their own policies on CONSORT reporting?. *Contemporary Clinical Trials* 2008;**29**(6):843–6.

Greenfield 2005 *{published and unpublished data}*
Greenfield ML, Rosenberg AL, O'Reilly M, Shanks AM, Sliwinski MJ, Nauss MD. The quality of randomized

controlled trials in major anesthesiology journals. *Anesthesia and Analgesia* 2005;**100**(6):1759–64.

Haahr 2006 ***{published and unpublished data}***
Haahr MT, Hrobjartsson A. Who is blinded in randomized clinical trials? A study of 200 trials and a survey of authors. *Clinical Trials* 2006;**3**(4):360–5.

Halpern 2004 ***{published and unpublished data}***
Halpern SH, Darani R, Douglas MJ, Wight W, Yee J. Compliance with the CONSORT checklist in obstetric anaesthesia randomised controlled trials. *International Journal of Obstetric Anesthesia* 2004;**13**(4):207–14.

Han 2008 ***{published and unpublished data}***
Han C, Kwak KP, Marks DM, Pae CU, Wu LT, Bhatia KS, et al. The impact of the CONSORT statement on reporting of randomized clinical trials in psychiatry. *Contemporary Clinical Trials* 2009;**30**(2):116–22.

Hewitt 2005 ***{published and unpublished data}***
Hewitt C, Hahn S, Torgerson DJ, Watson J, Bland JM. Adequacy and reporting of allocation concealment: review of recent trials published in four general medical journals. *BMJ* 2005;**330**(7499):1057–8.

Hill 2002 ***{published and unpublished data}***
Hill CL, LaValley MP, Felson DT. Secular changes in the quality of published randomized clinical trials in rheumatology. *Arthritis and Rheumatism* 2002;**46**(3): 779–84.

Hopewell 2010 ***{published and unpublished data}***
Hopewell S, Dutton S, Yu LM, Chan AW, Altman DG. The quality of reports of randomised trials in 2000 and 2006: comparative study of articles indexed in PubMed. *BMJ* 2010;**340**:723.
Yu LM, Chan AW, Hopewell S, Deeks JJ, Altman DG. Reporting on covariate adjustment in randomised controlled trials before and after revision of the 2001 CONSORT statement: a literature review. *Trials* 2010;**18**(11):59.

Kane 2007 ***{published and unpublished data}***
Kane RL, Wang J, Garrard J. Reporting in randomized clinical trials improved after adoption of the CONSORT statement. *Journal of Clinical Epidemiology* 2007;**60**(3): 241–9.

Kidwell 2001 ***{published and unpublished data}***
Kidwell CS, Liebeskind DS, Starkman S, Saver JL. Trends in acute ischemic stroke trials through the 20th century. *Stroke* 2001;**32**(6):1349–59.

Kober 2006 ***{published and unpublished data}***
Kober T, Trelle S, Engert A. Reporting of randomized controlled trials in Hodgkin lymphoma in biomedical journals. *Journal of the National Cancer Institute* 2006;**98** (9):620–5.

Ladd 2010 ***{published and unpublished data}***
Ladd BO, McCrady BS, Manuel JK, Campbell W. Improving the quality of reporting alcohol outcome studies: effects of the CONSORT statement. *Addictive Behaviors* 2010;**35**(7):660–6.

Lai 2006 ***{published and unpublished data}***
Lai R, Chu R, Fraumeni M, Thabane L. Quality of randomized controlled trials reporting in the primary treatment of brain tumors. *Journal of Clinical Oncology* 2006;**24**(7):1136–44.

Lai 2007 ***{published and unpublished data}***
Lai TY, Wong VW, Lam RF, Cheng AC, Lam DS, Leung GM. Quality of reporting of key methodological items of randomized controlled trials in clinical ophthalmic journals. *Ophthalmic Epidemiology* 2007;**14**(6):390–8.

Llorca 2004 ***{published and unpublished data}***
Llorca J, Martinez-Sanz F, Prieto-Salceda D, Farinas-Alvarez C, Chinchon MV, Quinones D, et al. Quality of controlled clinical trials on glaucoma and intraocular high pressure. *Journal of Glaucoma* 2005;**14**(3):190–5.

Moher 2001 ***{published and unpublished data}***
Moher D, Jones A, Lepage L. Use of the CONSORT statement and quality of reports of randomized trials: a comparative before-and-after evaluation. *JAMA* 2001;**285** (15):1992–5.

Montané 2010 ***{published and unpublished data}***
Montane E, Vallano A, Vidal X, Aguilera C, Laporte JR. Reporting randomised clinical trials of analgesics after traumatic or orthopaedic surgery is inadequate: a systematic review. *BMC Clinical Pharmacology* 2010;**10**:2.

Montori 2002 ***{published and unpublished data}***
Montori VM, Bhandari M, Devereaux PJ, Manns BJ, Ghali WA, Guyatt GH. In the dark - the reporting of blinding status in randomized controlled trials. *Journal of Clinical Epidemiology* 2002;**55**(8):787–90.

Pagoto 2009 ***{published and unpublished data}***
Pagoto SL, Kozak AT, John P, Bodenlos JS, Hedeker D, Spring B, et al. Intention-to-treat analyses in behavioral medicine randomized clinical trials. *International Journal of Behavioral Medicine* 2009;**16**(4):316–22.
Spring B, Pagoto S, Knatterud G, Kozak A, Hedeker D. Examination of the analytic quality of behavioral health randomized clinical trials. *Journal of Clinical Psychology* 2007;**63**(1):53–71.

Parés 2008 ***{published and unpublished data}***
Pares D, Norton C, Chelvanayagam S. Fecal incontinence: the quality of reported randomized, controlled trials in the last ten years. *Diseases of the Colon and Rectum* 2008;**51**(1): 88–95.

Partsinevelou 2009 ***{published and unpublished data}***
Partsinevelou A, Zintzaras E. Quality of reporting of randomized controlled trials in polycystic ovary syndrome. *Trials* 2009;**10**:106.

Pat 2008 ***{published and unpublished data}***
Pat K, Dooms C, Vansteenkiste J. Systematic review of symptom control and quality of life in studies on chemotherapy for advanced non-small cell lung cancer: how CONSORTed are the data?. *Lung Cancer* 2008;**62**(1): 126–38.

Peckitt 2007 *{published and unpublished data}*
Peckitt C, Ireland E, Kilburn LS, Bliss JM. Has the quality of early breast cancer randomized controlled trials publications improved since CONSORT? A systematic review. *Breast Cancer Research and Treatment* 2007, (106): S258.

Prady 2008 *{published and unpublished data}*
Prady SL, Richmond SJ, Morton VM, Macpherson H. A systematic evaluation of the impact of STRICTA and CONSORT recommendations on quality of reporting for acupuncture trials. *PLoS ONE* 2008;**3**(2):e1577.

Sanchez-Thorin 2001 *{published and unpublished data}*
Sanchez-Thorin JC, Cortes MC, Montenegro M, Villate N. The quality of reporting of randomized clinical trials published in ophthalmology. *Ophthalmology* 2001;**108**(2): 410–5.

Scales 2007 *{published and unpublished data}*
Scales CD, Norris RD, Keitz SA, Peterson BL, Preminger GM, Vieweg J, et al. A critical assessment of the quality of reporting of randomized, controlled trials in the urology literature. *Journal of Urology* 2007;**177**(3):1090–5.

Selman 2008 *{published data only}*
Selman TJ, Johnson NP, Zamora J, Khan KS. Gynaecologic surgery from uncertainty to science: evolution of randomized control trials. *Human Reproduction* 2008;**23** (4):827–31.

Sinha 2009 *{published and unpublished data}*
Sinha S, Sinha S, Ashby E, Jayaram R, Grocott MP. Quality of reporting in randomized trials published in high-quality surgical journals. *Journal of the American College of Surgeons* 2009;**209**(5):565–71.

Spring 2007 *{published and unpublished data}*
Spring B, Pagoto S, Knatterud G, Kozak A, Hedeker D. Examination of the analytic quality of behavioral health randomized clinical trials. *Journal of Clinical Psychology* 2007;**63**(1):53–71.

Thabane 2007 *{published and unpublished data}*
Thabane L, Chu R, Cuddy K, Douketis J. What is the quality of reporting in weight loss intervention studies? A systematic review of randomized controlled trials. *International Journal of Obesity* 2007;**31**(10):1554–9.
Thoma A, Chew RT, Sprague S, Veltri K. Application of the CONSORT statement to randomized controlled trials comparing endoscopic and open carpal tunnel release. *Canadian Journal of Plastic Surgery* 2006;**14**(4):205–10.

Tharyan 2008 *{published and unpublished data}*
Tharyan P, Premkumar TS, Mathew V, Barnabas JP, Manuelraj. Editorial policy and the reporting of randomized controlled trials: a survey of instructions for authors and assessment of trial reports in Indian medical journals (2004-05). *National Medical Journal of India* 2008;**21**(2):62–8.

Thoma 2006 *{published and unpublished data}*
Thoma A, Chew RT, Sprague S, Veltri K. Application of the CONSORT statement to randomized controlled trials comparing endoscopic and open carpal tunnel release. *Canadian Journal of Plastic Surgery* 2006;**14**(4):205–10.

Tiruvoipati 2005 *{published and unpublished data}*
Tiruvoipati R, Balasubramanian SP, Atturu G, Peek GJ, Elbourne D. Improving the quality of reporting randomized controlled trials in cardiothoracic surgery: the way forward. *Journal of Thoracic & Cardiovascular Surgery* 2006;**132**(2): 233–40.

Uetani 2009 *{published and unpublished data}*
Uetani K, Nakayama T, Ikai H, Yonemoto N, Moher D. Quality of reports on randomized controlled trials conducted in Japan: evaluation of adherence to the CONSORT statement. *Internal Medicine* 2009;**48**(5): 307–13.

Wang 2007 *{published data only}*
Wang G, Mao B, Xiong ZY, Fan T, Chen XD, Wang L, et al. The quality of reporting of randomised controlled trials of Traditional Chinese Medicine: a survey of 13 randomly selected journals from mainland China. *Clinical Theapeutics* 2007;**29**:1456–67.

Wei 2009 *{published and unpublished data}*
Wei X, Tiejun L, Cheng W. Current situation on the reporting quality of randomized controlled trials in 5 leading Chinese medical journals. *Journal of Medical Colleges of PLA* 2009;**24**(2):105–11.

Yu 2010 *{published data only}*
Yu LM, Chan AW, Hopewell S, Deeks JJ, Altman DG. Reporting on covariate adjustment in randomised controlled trials before and after revision of the 2001 CONSORT statement: a literature review. *Trials* 2010;**11**:59.

Zhong 2010 *{published and unpublished data}*
Zhong YQ, Fu JJ, Liu XM, Diao X, Mao B, Fan T, et al. The reporting quality, scientific rigor, and ethics of randomized placebo-controlled trials of Traditional Chinese Medicine compound formulations and the differences between Chinese and non-Chinese trials. *Current Therapeutic Research, Clinical and Experimental* 2010;**71**(1):30–49.

Ziogas 2009 *{published and unpublished data}*
Ziogas DC, Zintzaras E. Analysis of the quality of reporting of randomized controlled trials in acute and chronic myeloid leukemia, and myelodysplastic syndromes as governed by the CONSORT statement. *Annals of Epidemiology* 2009;**19** (7):494–500.

References to studies excluded from this review

Albavera-Hernández 2009 *{published data only}*
Albavera-Hernandez C, Rodriguez JM, Idrovo AJ. Safety of botulinum toxin type A among children with spasticity secondary to cerebral palsy: a systematic review of randomized clinical trials. *Clinical Rehabilitation* 2009;**23** (5):394–407.

Berwanger 2009 *{published data only}*
Berwanger O, Ribeiro RA, Finkelsztejn A, Watanabe M, Suzumura EA, Duncan BB, et al. The quality of reporting of trial abstracts is suboptimal: survey of major general medical journals. *Journal of Clinical Epidemiology* 2009;**62** (4):387–92.

Chowers 2009 *{published data only}*
Chowers MY, Gottesman BS, Leibovici L, Pielmeier U, Andreassen S, Paul M. Reporting of adverse events in randomized controlled trials of highly active antiretroviral therapy: systematic review. *Journal of Antimicrobial Chemotherapy* 2009;**64**(2):239–50.

Ellis 2005 *{published data only}*
Ellis C, Hall JL, Khalil A, Hall JC. Evolution of methodological standards in surgical trials. *ANZ Journal of Surgery* 2005;**75**(10):874–7.

Li 2009 *{published data only}*
Li JY, Zhang YF, Smith GS, Xue CJ, Luo YN, Chen WH, Skinner CJ, Finkelstein J. Quality of reporting of randomized clinical trials in Tai Chi interventions - a systematic review. *Evidence Based Complementary and Alternative Medicine* 2009.

Mills 2005 *{published data only}*
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. *Journal of Clinical Epidemiology* 2005;**58**(7): 662–7.

Norton-Mabus 2008 *{published data only}*
Norton-Mabus JC, Nelson DL. Reporting of randomized controlled trials in occupational therapy and speech therapy: evaluation using an expansion of the CONSORT statement. *OTJR Occupation, Participation and Health* 2008;**28**(2): 64–71.

Smith 2008 *{published data only}*
Smith BA, Lee HJ, Lee JH, Choi M, Jones DE, Bausell RB, et al. Quality of reporting randomized controlled trials (RCTs) in the nursing literature: application of the consolidated standards of reporting trials (CONSORT). *Nursing Outlook* 2008;**56**(1):31–7.

Taghinia 2008 *{published data only}*
Taghinia AH, Liao EC, May JW. Randomized controlled trials in plastic surgery: a 20-year review of reporting standards, methodologic quality, and impact. *Plastic and Reconstructive Surgery* 2008;**122**(4):1253–63.

Xu 2008 *{published data only}*
Xu L, Li J, Zhang M, Ai C, Wang L. Chinese authors do need CONSORT: reporting quality assessment for five leading Chinese medical journals. *Contemporary Clinical Trials* 2008;**29**(5):727–31.

Yu 2009 *{published data only}*
Yu S, Zhong B, Zheng M, Xiao F, Dong Z, Zhang H. The quality of randomized controlled trials on DanShen in the treatment of ischemic vascular disease. *Journal of Alternative and Complementary Medicine* 2009;**15**(5):557–65.

References to studies awaiting assessment

Siegfried 2008 *{published and unpublished data}*
Siegfried N, Clarke M, Volmink J, Van der Merwe L. African HIV/AIDS trials are more likely to report adequate allocation concealment and random generation than North American trials. *PLoS ONE* 2008;**3**(10):e3491.

Additional references

Altman 2001
Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D, et al. for the CONSORT Group. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Annals of Internal Medicine* 2001;**134**:663–94.

Altman 2005
Altman DG. Endorsement of the CONSORT statement by high impact medical journals: survey of instructions for authors. *BMJ* 2005;**330**(7499):1056–7.

Begg 1996
Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I, et al. Improving the quality of reporting of randomized controlled trials: the CONSORT statement. *JAMA* 1996; **276**:637–9. [PUBMED: 8773637]

Birnie 2008
Birnie AJ, Bath-Hextall FJ, Ravenscroft JC, Williams HC. Interventions to reduce Staphylococcus aureus in the management of atopic eczema. *Cochrane Database of Systematic Reviews* 2008, Issue 3. [DOI: 10.1002/14651858.CD003871.pub2

Campbell 1966
Campbell DC, Stanley JT. *Experimental and Quasi-Experimental Design for Research*. Boston: Houghton Mifflin Company, 1966.

Chan 2005
Chan AW, Altman DG. Epidemiology and reporting of randomised trials published in PubMed journals. *Lancet* 2005;**365**:1159–62.

CMA [Computer program]
Borenstein M, Hedges L, Higgins J, Rothstein H. Comprehensive Meta-analysis Version 2. Englewood, NJ: Biostat, 2005.

Cobo 2011
Cobo E, Cortés J, Ribera JM, Cardellach F, Selva-O'Callaghan A, Kostov B, et al. Effect of using reporting guidelines during peer review on quality of final manuscripts submitted to a biomedical journal: masked randomised trial. *BMJ* 2011;**343**:d6783.

Cochrane EPOC 2009
Cochrane Effective Practice and Organisation of Care Review Group. Data Collection Checklist. http://epoc.cochrane.org/sites/epoc.cochrane.org/files/uploads/Suggested%20risk%20of%20bias%20criteria%20for%20EPOC%20reviews.pdf (accessed 6 July 2009):1–30.

CONSORT Group 2009
CONSORT Group. CONSORT: Transparent Reporting of Trials. www.consort-statement.org (accessed 27 May 2009).

Dechartres 2011
Dechartres A, Charles P, Hopewell S, Ravaud P, Altman DG. Reviews assessing the quality or the reporting of

randomized controlled trials are increasing over time but raised questions about how quality is assessed. *Journal of Clinical Epidemiology* 2011;**64**(2):136–44.

Deeks 2008
Deeks JJ, Higgins JPT, Altman DA. Chapter 9: Analysing data and undertaking meta-analyses. In: Higgins JPT, Green S (eds). Cochrane Handbook for Systematic Reviews of Interventions Version 5.0.1 [updated September 2008], The Cochrane Collaboration, 2008. Available from www.cochrane-handbook.org.

Glasziou 2008
Glasziou P, Meats E, Heneghan C, Shepperd S. What is missing from descriptions of treatments in trials and reviews?. *BMJ* 2008;**336**:1472–4.

Higgins 2008
Higgins JPT, Altman DG. Chapter 8: Assessing risk of bias in included studies. In: Higgins JPT, Green S (eds). Cochrane Handbook for Systematic Reviews of Interventions Version 5.0.1 [updated September 2008], The Cochrane Collaboration, 2008. Available from www.cochrane-handbook.org.

Hopewell 2008
Hopewell S, Altman D, Moher D, Schulz K. Endorsement of the CONSORT Statement by high impact factor medical journals: a survey of journal editors and journal 'instruction to authors'. *Trials* 2008;**9**:20.

Jadad 1996
Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJ, Gavaghan DJ, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary?. *Controlled Clinical Trials* 1996;**17**:1–12.

Llorca 2005
Llorca J, Martínez-Sanz F, Prieto-Salceda D, Fariñas-Alvarez C, Chinchón MV, Quinones D, et al. Quality of controlled clinical trials on glaucoma and intraocular high pressure. *Journal of Glaucoma* 2005;**14**(3):190–5.

MINCIR Score
Manterola C, Busquets J, Pascual M, Grande L. Cuál es lacalidad metodológica de los artículos sobre procedimientosterapéuticos publicados en Cirugía Española?. *Cirugía Española* 2006;**79**:95-100.

Moher 2001a
Moher D, Schulz KF, Altman D. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *JAMA* 2001;**285**:1987–91.

Moher 2009
Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Medicine* 2009;**6**(7):e1000097.

Moher 2010
Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trial. *BMJ* 2010;**340**:c869.

Nolte 2004
Nolte S, Wong D, Latchford G, Boyle O, Anaenugwu A. Amphetamines for schizophrenia. *Cochrane Database of Systematic Reviews* 2004, Issue 3. [DOI: 10.1002/14651858.CD004964

Olivo 2008
Olivo SA, Macedo LG, Gadotti IC, Fuentes J, Stanton T, Magee DJ. Scales to assess the quality of randomized controlled trials: a systematic review. *Physical Therapy* 2008; **88**(2):156–75.

Plint 2006
Plint AC, Moher D, Morrison A, Schulz K, Altman DG, Hill C, et al. Does the CONSORT checklist improve the quality of reports of randomised controlled trials? A systematic review. *Medical Journal of Australia* 2006;**185**: 263–7. [PUBMED: 16948622]

Schulz 1995
Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias: dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995;**273**:408–12.

Schulz 2010
Schulz KF, Altman DG, Moher D, for the CONSORT Group. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. *Annals of Internal Medicine* 2010;**152**:726–32.

* *Indicates the major publication for the study*

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

Agha 2007

Methods	Evaluates the degree to which RCTs involving urological surgical techniques (as the intervention) published in the years 2000-2003 complied with the CONSORT Statement, and assesses trends and patterns of compliance The study was then extended to a number of other specialties to assess whether our findings in urology could be generalised to other surgical disciplines
Data	90 RCTs from 35 journals, 22 items unweighted CONSORT score recorded, unable to obtain dates of endorsement for all journals, included what was readily available
Comparisons	CONSORT-endorsing and non-endorsing journals, quality of RCTs before and after CONSORT publication
Outcomes	Total sum score
Included number of RCTs, Journals	88, 33
Checklist version used	2001
Field of Study	Surgical medicine
Notes	Author was contacted, additional item data no longer available

Risk of bias

Item	**Authors' judgement**	**Description**
Large Cohort ?	Yes	2 electronic databases were searched over 3 years
Blinding?	Unclear	Not reported
Confounding by journal quality?	Unclear	Not reported
Outcome Reporting?	Yes	No difference between planned and reported outcomes/analyses
Multiple raters?	Unclear	Not reported
Rater agreement?	Yes	Not applicable
Blinding, quality assessment?	Yes	Not applicable

Al-Namankany 2009

Methods	This evaluation assesses completeness of reporting in published RCTs in paediatric dental journals, as a secondary outcome to see whether quality of reporting has improved since the introduction of the CONSORT guidelines Trials published from 1985 to 1997, and from 1998 to 2006 were compared
Data	173 RCTs from 8 journals, 22 CONSORT items converted into 34 questions
Comparisons	Qualitatively synthesised based on data in the text considering quality of RCTs before and after the publication of CONSORT
Outcomes	Sequence generation, allocation concealment, blinding of participants, blinding of administer of interventions, outcome assessor blinding
Included number of RCTs, Journals	173, 8
Checklist version used	1996
Field of Study	Pediatric dentistry
Notes	Data sent by author, but was not consistent with our needs for inclusion, so used as readily available in the text; as the denominator for comparison groups is not included this study is included for qualitative synthesis

Risk of bias

Item	**Authors' judgement**	**Description**
Large Cohort ?	Yes	Searched PubMed over 2 decades
Blinding?	Unclear	Not reported
Confounding by journal quality?	Unclear	Not reported
Outcome Reporting?	Yes	No evidence of selective reporting
Multiple raters?	Yes	Quote: "...all items were considered together for each paper and a good agreement between the two reviewers..."
Rater agreement?	Yes	Quote: "...was found with κ = 0.92 (0.88-0.96)"
Blinding, quality assessment?	Yes	Not applicable

Alvarez 2009

Methods	Assesses the effect of the adoption of CONSORT on the reporting quality of RCTs by systematic evaluation of RCTs published in 2 dermatology journals pre- and post CONSORT adoption; RCTs were published in 1997 and 2006 6 CONSORT checklist items were evaluated by equal weight
Data	98 RCTs from 2 journals
Comparisons	CONSORT-endorsing journals before and after CONSORT endorsement
Outcomes	Interventions, methods, blinding, outcomes, sample size and sequence generation
Included number of RCTs, Journals	98, 2
Checklist version used	2001
Field of Study	Dermatology
Notes	Author provided all raw data and gave permission to be adapted for inclusion in our study. As such, the endorsement definition holds

Risk of bias

Item	**Authors' judgement**	**Description**
Large Cohort ?	No	From 2 years in 2 journals
Blinding?	Unclear	Not reported
Confounding by journal quality?	Yes	Not applicable
Outcome Reporting?	Yes	No evidence of selective reporting of outcomes
Multiple raters?	Unclear	Not reported
Rater agreement?	Unclear	Not reported
Blinding, quality assessment?	Yes	Not applicable

Anttila 2006

Methods	Evaluates trial reporting by using the CONSORT Statement recommendations for trials published in or after 1990; the checklist was modified to include 33 items Trials published between 1990-1997 and 1998-2002 were compared to see if CONSORT had an influence on the quality of reporting
Data	15 trials from 9 journals, only 1 journal deemed to be an endorsing journal

Anttila 2006 (*Continued*)

Comparisons	Before and after CONSORT publication
Outcomes	Title and abstract, background, participants, interventions, objectives, outcomes, sample size, sequence generation, allocation concealment, implementation, blinding of: participants, data analyst and outcome assessor, statistical methods, participant flow, recruitment, baseline data, numbers analysed, outcomes and estimation, ancillary analyses, adverse events
Included number of RCTs, Journals	14, 9
Checklist version used	1996
Field of Study	Cerebral palsy
Notes	Data needed provided in the appendix; recategorised data to be compliant with our comparison

Risk of bias

Item	**Authors' judgement**	**Description**
Large Cohort ?	No	15 included RCTs
Blinding?	Unclear	Not reported
Confounding by journal quality?	Unclear	Not reported
Outcome Reporting?	Yes	No difference between planned and reported outcomes/analyses
Multiple raters?	Yes	Quote: "Two researchers (R.K. and H.A.) independently evaluated the quality of reporting in the identified trials by using this modified checklist."
Rater agreement?	No	Quote: "The evaluators disagreed in 23% of the evaluations."
Blinding, quality assessment?	Yes	Not applicable

Areia 2010

Methods	This study evaluated quality in recently published endoscopic articles in articles published from 1998 to 2008 by assessing STARD and CONSORT
Data	10 RCTs of 120 articles, 2 endorsing journals
Comparisons	CONSORT-endorsing journals versus CONSORT non-endorsing journals

Areia 2010 (*Continued*)

Outcomes	Title and abstract, background, participants, interventions, objectives, outcomes, sample size, sequence generation, allocation concealment, implementation, blinding of: participants, data analyst and outcome assessor, statistical methods, participant flow, recruitment, baseline data, numbers analysed, outcomes and estimation, ancillary analyses, adverse events, interpretation, generalisability, overall evidence
Included number of RCTs, Journals	10, 5
Checklist version used	2001
Field of Study	Endoscopy
Notes	Author provided full data set; endorsement was confirmed and meets our definition

Risk of bias

Item	Authors' judgement	Description
Large Cohort ?	Yes	Sampled over a decade, large number of trials in study
Blinding?	Unclear	Not reported
Confounding by journal quality?	Unclear	Not reported
Outcome Reporting?	Yes	No evidence of selective outcome reporting
Multiple raters?	Yes	2 reviewers independently reviewed
Rater agreement?	Yes	Interobserver agreement was 97.3%
Blinding, quality assessment?	Yes	Not applicable

Balasubramanian 2006

Methods	Evaluates the quality of reporting of surgical randomised controlled trials published in surgical and general medical journals in 2003 using Jadad score, allocation concealment, and adherence to CONSORT guidelines and to identify factors associated with good quality
Data	CONSORT score is reported as a median across all 30 items scored from 1 to 3 where 1 was no description and 3 corresponded to adequate description
Comparisons	CONSORT endorsers versus non-endorsers
Outcomes	Total sum score
Included number of RCTs, Journals	69, 10

Balasubramanian 2006 (*Continued*)

Checklist version used	2001 (modified to 30 items)
Field of Study	General surgery
Notes	To be an endorser of the journal had to have such guidance in their 'instructions to authors' which meets the definition in this review Unable to obtain scores for each RCT which would have allowed inclusion across all items This study also assessed quality using the Jadad score and Schulz allocation concealment

Risk of bias

Item	Authors' judgement	Description
Large Cohort ?	No	RCTs published in 10 top journals over 1-year period
Blinding?	Unclear	Not reported
Confounding by journal quality?	Unclear	Not reported
Outcome Reporting?	Yes	No difference between planned and reported outcomes/analyses
Multiple raters?	Yes	Quote: "Each article was then assessed for every item on the checklist and scored independently by 2 observers (S.P.B. and R.T.)"
Rater agreement?	No	Quote: "The agreement between the pair of observers who independently assessed the RCTs using the CONSORT checklist was good (ICC 0.85; 95% CI 0.77-0.91; P 0.001) "
Blinding, quality assessment?	Unclear	Not reported

Bausch 2009

Methods	Assessed trial quality in COPD RCTs by key items; quality of reporting was compared over several comparisons, of which CONSORT endorsement was one RCTs published in 1957-2000 versus after 2000
Data	As individual RCT data were available, data were extracted to compare 239 RCTs pre-2001 versus 105 RCTs from 2001 onwards
Comparisons	Before and after CONSORT publication

Bausch 2009 (*Continued*)

Outcomes	Allocation concealment, sequence generation, participants, blinding: participants, intervention, outcome assessor, outcomes and estimation, numbers analysed
Included number of RCTs, Journals	344, 110
Checklist version used	Used pre-specified criteria which coincide with 8 CONSORT 2001 checklist items
Field of Study	Chronic obstructive pulmonary disease
Notes	Author provided RCT level data ensuring this study could be included in our analysis With more resources, this study could potentially be included in the CONSORT endorsers versus non-endorsers comparison 90 RCTs were published before 1990; it is of importance to note potential confounding by improvement in reporting quality over time

Risk of bias

Item	**Authors' judgement**	**Description**
Large Cohort ?	Yes	Multiple databases, large number of trials assessed
Blinding?	Unclear	Not reported
Confounding by journal quality?	Unclear	Not reported
Outcome Reporting?	Unclear	Unsure, no explicit evidence of selective reporting
Multiple raters?	Unclear	Multiple raters, but not specified for CONSORT items assessment
Rater agreement?	Unclear	Not reported
Blinding, quality assessment?	Unclear	Not applicable

Bian 2006

Methods	Evaluated the quality of Chinese herbal medicine RCTs using a modified CONSORT checklist before and after 2000, the 4th of a 4-part series considering the quality of Chinese herbal medicine RCTs
Data	Percentage reported by year, data extracted to form comparison before 2001 and 2001 onwards
Comparisons	Before and after CONSORT publication
Outcomes	Total sum score of 63 items

Bian 2006 (*Continued*)

Included number of RCTs, Journals	167, 35
Checklist version used	63-item modification of the 2001 checklist
Field of Study	Chinese herbal medicine
Notes	Author provided additional information, but this was not all that was necessary to include in a more robust comparison

Risk of bias

Item	Authors' judgement	Description
Large Cohort ?	Unclear	RCTs from 11 systematic reviews on Chinese herbal medicine
Blinding?	Unclear	Not reported
Confounding by journal quality?	Unclear	Not reported
Outcome Reporting?	Yes	No difference between planned and reported outcomes
Multiple raters?	Yes	Independent assessment by 2 reviewers
Rater agreement?	Unclear	Not reported
Blinding, quality assessment?	Yes	Not applicable

Chauhan 2009

Methods	Appraised the compliance of randomised clinical trials (RCTs) cited for level A recommendations in obstetric practice bulletins (OPBs) and published after the CONSORT (Consolidated Standards of Reporting Trials, published 1996) statement
Data	50-item checklist Unweighted median score reported before and after 1997 Compares 58 RCTs before 1997 and 32 RCTs after 1997, described as before and after CONSORT Post
Comparisons	Median 50-item score before and after CONSORT publication
Outcomes	Included as 'primary evidence', synthesised qualitatively Median total sum score
Included number of RCTs, Journals	90, 5

Chauhan 2009 (*Continued*)

Checklist version used	1996
Field of Study	Obstetric practice
Notes	RCT level data unavailable to include with comparison data Desciptive comparison only

Risk of bias

Item	Authors' judgement	Description
Large Cohort ?	No	RCTs from single journal over 8-year period
Blinding?	Unclear	Not reported
Confounding by journal quality?	Unclear	Not reported
Outcome Reporting?	Yes	No difference between planned and reported outcomes/analyses
Multiple raters?	Unclear	Not reported
Rater agreement?	Yes	Not applicable
Blinding, quality assessment?	Yes	Not applicable

Devereaux 2002

Methods	Observational study to determine the quality of reporting key methodological factors in RCTs since the publication of the CONSORT Statement and if CONSORT endorsement by journals of the checklist was associated with superior reporting. 11 key methodological factors Examined the quality of reporting in relation to whether a journal was a 'CONSORT promoter' as defined by inclusion of the CONSORT checklist in a journal's 'information to authors' section or a requirement that authors, manuscript reviewers, or copy editors complete the CONSORT checklist
Data	7 journals were confirmed to meet our definition of CONSORT endorser, versus 19 non-endorsing journals
Comparisons	CONSORT endorsers versus non-endorsers
Outcomes	Allocation concealment, sequence generation, statistical methods, participant flow, baseline data, blinding: outcome assessor, intervention, data analyst, participants
Included number of RCTs, Journals	105, 26
Checklist version used	1996

Devereaux 2002 (*Continued*)

Field of Study	Internal medicine	
Notes	This study was included in the original review	
Risk of bias		
Item	**Authors' judgement**	**Description**
Large Cohort ?	No	3 journals, shorter time period
Blinding?	Unclear	Not reported
Confounding by journal quality?	Unclear	Quote: "We conducted a multivariable analysis (i.e., least squares regression) in which the dependent variable was the number of factors included in each article and the independent variables were the impact factor of the journal"
Outcome Reporting?	Yes	No evidence of selective reporting
Multiple raters?	Yes	Quote: "Two of us (W.G. and G.G.) independently evaluated all summaries"
Rater agreement?	Yes	> 0.8
Blinding, quality assessment?	Yes	Not applicable

Dias 2006

Methods	Aim was to assess whether quality has improved over time, particularly since the publication of CONSORT, and to assess what proportion of trials could be included in the meta-analyses of pregnancy outcomes such as those included in Cochrane Reviews Trials selected were published in 1990, 1996, and 2002; only trials published in English as full journal articles, claiming to be randomised and reporting on pregnancy outcomes, were included
Data	Journal endorsement was verified for compliance with our definition, as such a total of 60 and 53 RCTs were included for the endorsers versus non-endorsers, and before and after endorsement comparisons respectively
Comparisons	CONSORT endorsers versus non-endorsers, CONSORT-endorsing journals before and after CONSORT endorsement
Outcomes	For both comparisons: sequence generation, allocation concealment, participant flow, adverse events

Dias 2006 (*Continued*)

Included number of RCTs, Journals	164, 29
Checklist version used	1996
Field of Study	Subfertility
Notes	Author provided necessary data for our review This study also included data for control comparisons including, pre-CONSORT endorsers versus pre-CONSORT non-endorsers: allocation concealment, sequence generation and adverse events. pre-post consort non-endorsers: sequence generation, allocation concealment, participant flow, adverse events

Risk of bias

Item	**Authors' judgement**	**Description**
Large Cohort ?	Yes	RCTs from Cochrane review group register from which 3455 references were available
Blinding?	Unclear	Not reported
Confounding by journal quality?	Unclear	Not clearly reported
Outcome Reporting?	Yes	No difference between planned and reported outcomes/analyses
Multiple raters?	Unclear	Not reported
Rater agreement?	Yes	Not applicable
Blinding, quality assessment?	Yes	Not applicable

Dickinson 2002

Methods	Assessed the quality of reporting in RCTs of lifestyle interventions
Data	From the provided data, items were sorted into before and after 1996 publication, 10 RCTs were published after 1996 and 72 RCTs were published before, from 1977
Comparisons	Cross-sectional sample before and after CONSORT publication
Outcomes	Allocation concealment, blinding of participants
Included number of RCTs, Journals	166, not reported
Checklist version used	1996

Dickinson 2002 (*Continued*)

Field of Study	Lifestyle interventions
Notes	Author presented poster at Cochrane Colloquium. Author provided data

Risk of bias

Item	Authors' judgement	Description
Large Cohort ?	Unclear	Over long time period, journals unknown
Blinding?	Unclear	Not reported
Confounding by journal quality?	Unclear	Not reported
Outcome Reporting?	Unclear	Not reported
Multiple raters?	Unclear	Not reported
Rater agreement?	Unclear	Not reported
Blinding, quality assessment?	Unclear	Not reported

Ethgen 2009

Methods	Objective of study was to evaluate the quality of reporting internal and external validity data in published reports of RCTs assessing the stents for percutaneous coronary interventions Quality attributed to CONSORT-endorsing journals was also reported in the abstract
Data	Quality was assessed using the CLEAR NPT checklist
Comparisons	CONSORT-endorsing journals versus non-endorsing journals
Outcomes	Interventions, sequence generation, allocation concealment, numbers analysed, blinding: outcome assessor, intervention, participants
Included number of RCTs, Journals	123, 29 (unknown for 9 RCTs)
Checklist version used	2001
Field of Study	Stents for percutaneous coronary interventions
Notes	Author provided data Insufficient resources to confirm endorsement compliance with our definition

Risk of bias

Ethgen 2009 (*Continued*)

Item	Authors' judgement	Description
Large Cohort ?	Yes	Over 5 years, MEDLINE and Cochrane searched, large number of RCTs
Blinding?	Unclear	Not reported
Confounding by journal quality?	No	Quote: "The quality of reporting was better in journals with high impact factors and in journals endorsing the CONSORT statement."
Outcome Reporting?	Yes	No evidence of selective reporting
Multiple raters?	Unclear	Verification of sample conducted
Rater agreement?	Unclear	Not reported
Blinding, quality assessment?	Yes	Not applicable

Faunce 2003

Methods	Reviewed RCTs before and after 1996 publication of CONSORT based on key methodological items
Data	Endorsement of journals was verified, 2 of which endorse the CONSORT checklist
Comparisons	CONSORT endorsers verus non-endorsers, CONSORT-endorsing journals before and after CONSORT endorsement
Outcomes	Participants, sample size, and participant flow for endorsers versus non-endorsers and participants and participant flow before and after endorsement
Included number of RCTs, Journals	13, 7
Checklist version used	1996
Field of Study	Overdoses in health volunteers
Notes	This study was included in the original review

Risk of bias

Item	Authors' judgement	Description
Large Cohort ?	Yes	2-year, multiple journal sample
Blinding?	Unclear	Not reported

Faunce 2003 (*Continued*)

Confounding by journal quality?	Unclear	Not reported
Outcome Reporting?	Yes	No evidence of selective reporting
Multiple raters?	Unclear	Not reported
Rater agreement?	Unclear	Not reported
Blinding, quality assessment?	Yes	Not applicable

Folkes 2008

Methods	Assesses the extent of completeness of reporting to pre-randomisation data reporting in 4 leading general medicine journals, as recommended by CONSORT
Data	Study reports the improvement in reporting from 2004 and 2006, 3 endorsing journals and one journal endorsing in 2005 Data reported for 2004 included only
Comparisons	CONSORT endorsers versus non-endorsers
Outcomes	Participants
Included number of RCTs, Journals	480, 4
Checklist version used	2001
Field of Study	None specified
Notes	Author unable to provide data for RCTs included published by *NEJM* to confirm non-endorser comparison group for 2006 data Including only 2004 data endorsers meets definition for our review

Risk of bias

Item	**Authors' judgement**	**Description**
Large Cohort ?	Yes	Cross-sections of 2 calendar years at 4 top journals; 480 included studies
Blinding?	Unclear	Not reported
Confounding by journal quality?	Unclear	Not reported
Outcome Reporting?	Yes	No difference between planned and reported outcomes/analyses

Folkes 2008 (*Continued*)

Multiple raters?	Yes	Quote: "...two reviewers (AF, RU) independently evaluated the trials' reporting of pre-randomization information (the 'Enrollment' stage), as outlined in the CONSORT statement"
Rater agreement?	Unclear	Not reported
Blinding, quality assessment?	Yes	Not applicable

Greenfield 2005

Methods	To assess the quality of reporting in anesthesiology journals with RCTs published in 2000
Data	A modified version of the Chalmers tool was used to assess quality
Comparisons	CONSORT endorsers versus non-endorser
Outcomes	Allocation concealment, blinding, participant flow, adverse events
Included number of RCTs, Journals	279, 4
Checklist version used	2001 (items coincide with)
Field of Study	Anesthesiology
Notes	Author provided additional data Journal endorsement was verified and consistent with our definition

Risk of bias

Item	**Authors' judgement**	**Description**
Large Cohort ?	Unclear	Large search, 4 journals included over 1 year
Blinding?	Yes	Quote: "These 279 articles were photocopied, and all identifiers were removed from all pages by three investigators (MDN, AS, and MJS) who were not involved in further evaluation"
Confounding by journal quality?	Unclear	Quote: "However, it is important to note that only two of the major general anesthesiology journals reviewed in this article have adopted CONSORT guidelines in their instructions to authors"; no explicit details and no adjustment for clustering

Greenfield 2005 (*Continued*)

Outcome Reporting?	Yes	No evidence of selective reporting of outcomes
Multiple raters?	Yes	Implied. Quote: "Articles were offered in a random order using a computer generated randomization scheme. Both reviewers have had formal training in research design, epidemiology, and biostatistics"
Rater agreement?	Unclear	Not reported
Blinding, quality assessment?	Yes	Chalmers score assessed and blinded

Haahr 2006

Methods	To assess the reporting of blinding in RCTs, sample of 2001 published trials in the Cochrane Central Register of Controlled Trials
Data	15 RCTs included from 10 endorsing journals and 185 RCTs from 61 journals
Comparisons	CONSORT-endorsing journals versus non-endorsing journals
Outcomes	Blinding of: outcome assessor, participants, intervention, data analyst
Included number of RCTs, Journals	200, 171
Checklist version used	2001
Field of Study	None specified
Notes	Author provided data Journal endorsement has been verified and complies with our definition

Risk of bias

Item	**Authors' judgement**	**Description**
Large Cohort ?	Yes	Sample of 200 trials from Cochrane Trials Register Issue 1, 2003
Blinding?	Unclear	Not reported
Confounding by journal quality?	Unclear	Not reported
Outcome Reporting?	Yes	No difference between planned and reported outcomes/analyses

Haahr 2006 (*Continued*)

Multiple raters?	Yes	2 raters
Rater agreement?	Unclear	Not reported
Blinding, quality assessment?	Yes	Not applicable

Halpern 2004

Methods	RCTs pertaining to the practice of obstetric anaesthesia and summarised in *Obstetric Anesthesia Digest* between March 2001 and December 2002 were assessed to compare the quality of reporting to the CONSORT checklist
Data	6 RCTs of one endorsing journal, 77 RCTs from 6 non-endorsing journals
Comparisons	CONSORT-endorsing versus non-endorsing journals
Outcomes	Title and abstract, participants, interventions, objectives, outcomes, sample size, sequence generation, allocation concealment, implementation, blinding: participants, outcome assessor and intervention, statistical methods
Included number of RCTs, Journals	100, 7
Checklist version used	2001
Field of Study	Obstetric anaesthesia
Notes	Included in the original review Author provided data

Risk of bias

Item	**Authors' judgement**	**Description**
Large Cohort ?	No	Only journal articles published in 1 digest magazine
Blinding?	Unclear	Not reported
Confounding by journal quality?	Yes	Not applicable, only 1 journal
Outcome Reporting?	Yes	No evidence of selective outcome reporting
Multiple raters?	Yes	Quote: "Each of the study articles was then scored by two investigators independently"
Rater agreement?	Unclear	Not reported

Blinding, quality assessment?	Yes	Not applicable

Han 2008

Methods	Determined whether the CONSORT recommendations influenced the quality of reporting of randomised controlled trials (RCTs) in the field of psychiatry Evaluated the quality of clinical trial reports before and after the introduction of CONSORT Statement Trials were published from period of 1992-1996 (pre-CONSORT) and 2002-2007 (post CONSORT)
Data	166 pre-CONSORT RCTs were compared across all CONSORT items with 276 post CONSORT items
Comparisons	CONSORT-endorsing journals before and after endorsement
Outcomes	Title and abstract, background, participants, interventions, objectives, outcomes, sample size, sequence generation, allocation concealment, implementation, blinding any, statistical methods, participant flow, recruitment, baseline data, numbers analysed, outcomes and estimation, ancillary analyses, adverse events, interpretation, generalisability, overall evidence
Included number of RCTs, Journals	442, 7
Checklist version used	2001
Field of Study	Psychiatry
Notes	

Risk of bias

Item	**Authors' judgement**	**Description**
Large Cohort ?	Yes	7 journals over 9 years search via PubMed
Blinding?	Unclear	Not reported
Confounding by journal quality?	Unclear	Not reported
Outcome Reporting?	Yes	No evidence of selective outcome reporting
Multiple raters?	Yes	2 raters assessed items
Rater agreement?	Yes	Concordance rate reported of 95%
Blinding, quality assessment?	Yes	Not applicable

Hewitt 2005

Methods	RCTs in general medical journals in 2002 in 4 medical journals assessed for adequacy of reporting of allocation concealment
Data	166 endorsing RCTs and 68 non-endorsing RCTs
Comparisons	CONSORT endorsers versus non-endorsers
Outcomes	Allocation concealment
Included number of RCTs, Journals	234, 4
Checklist version used	Modification
Field of Study	General medical journals
Notes	This study was included in the original review

Risk of bias

Item	**Authors' judgement**	**Description**
Large Cohort ?	Unclear	Not a complete year, but large number of included studies
Blinding?	Unclear	Not reported
Confounding by journal quality?	Yes	Quote: "Our statistical analyses adjusted for clustering effects by journal."
Outcome Reporting?	Yes	No evidence to suggest reported outcomes were selective
Multiple raters?	Unclear	Not reported
Rater agreement?	Unclear	Not reported
Blinding, quality assessment?	Yes	Not applicable

Hill 2002

Methods	RCTs included from 1987-1988 and 1997-1998; quality was assessed by adequate reporting of a modified Jadad scale
Data	119 pre-CONSORT RCTs versus 121 post CONSORT RCTs
Comparisons	CONSORT-endorsing journals before and after endorsement, CONSORT endorsers before and after endorsement

Hill 2002 (*Continued*)

Outcomes	CONSORT-endorsing versus non-endorsing: sequence generation, allocation concealment, statistical methods, participant flow, numbers analysed CONSORT-endorsing journals before and after endorsement: sequence generation, allocation concealment, statistical methods, participant flow
Included number of RCTs, Journals	240, 68
Checklist version used	1996
Field of Study	Adult rheumatological diseases
Notes	This study was included in the original review Author provided data Endorsement of journals has been confirmed and is compliant Used Jadad scaled to assess quality

Risk of bias

Item	Authors' judgement	Description
Large Cohort ?	Yes	Large range of journals over many years
Blinding?	Unclear	Not reported
Confounding by journal quality?	Yes	Quote: "Analyses were undertaken... and comparing RCTs from "high"- and "low"-impact journals"
Outcome Reporting?	Yes	No evidence of selective reporting
Multiple raters?	Yes	Abstracted in duplicate
Rater agreement?	Yes	Quote: "Kappa 0.80 for all features combine"
Blinding, quality assessment?	Unclear	Not reported

Hopewell 2010

Methods	Examines the reporting characteristics and methodological details of randomised trials indexed in PubMed in 2000 and 2006 and assess whether the quality of reporting has improved after publication of CONSORT in 2001
Data	Design: comparison of 2 cross-sectional investigations of indexed trials in PubMed in December 2000 (n = 519) and December 2006 (n = 616)
Comparisons	CONSORT-endorsing versus non-endorsers, cross-sectional sample of before and after CONSORT publication

Hopewell 2010 (*Continued*)

Outcomes	Endorsers versus non-endorsers comparison: blinding any, outcomes, sample size, sequence generation, allocation concealment Before and after publication comparison: outcomes, sequence generation, title and abstract, blinding any, numbers analysed, participant flow, allocation concealment, sample size
Included number of RCTs, Journals	1135, 587
Checklist version used	2001
Field of Study	None specified
Notes	This study is the primary study of the companion Yu 2010 Does not meet definition of endorser defined for our review

Risk of bias

Item	Authors' judgement	Description
Large Cohort ?	Yes	Cross-section of RCTs in PubMed from December 2006; 616 primary RCT reports included
Blinding?	Unclear	Not reported
Confounding by journal quality?	Yes	No potential for clustering by journal
Outcome Reporting?	Yes	No difference between planned and reported outcomes/analyses
Multiple raters?	Unclear	Not reported
Rater agreement?	Yes	Not applicable
Blinding, quality assessment?	Yes	Not applicable

Kane 2007

Methods	Examines the extent to which CONSORT improved clinical trials reporting and subject attrition, which may undermine the credibility of published randomised clinical trials (RCTs) Includes RCTs reported in 2 major medical journals before and after the CONSORT guidelines were endorsed; one used the CONSORT Statement (*JAMA*) and one did not acting as control (*NEJM*)
Data	308 RCTs pre-CONSORT (1993-1995), 88 RCTs published in *JAMA* and 220 in *NEJM*, and 468 RCTs post CONSORT (1999-2002) of which 178 were published in *JAMA* and 290 in *NEJM*

Kane 2007 (*Continued*)

Comparisons	CONSORT-endorsing versus non-endorsing, CONSORT-endorsing journals before and after endorsement
Outcomes	Endorsers versus non-endorsers: sample size, participant flow, numbers analysed, recruitment, blinding any, sequence generation CONSORT endorsers before and after endorsement: participant flow, number analysed, sample size, blinding any, recruitment
Included number of RCTs, Journals	776, 2
Checklist version used	1996
Field of Study	None specified
Notes	This study also includes a number of control comparisons. Data available for: Pre-post CONSORT non-endorsers: sample size, sequence generation, allocation concealment, blinding any description, participant flow, recruitment, numbers analysed Pre-CONSORT endorsers versus non-endorsers: sample size, sequence generation, allocation concealment, blinding any description, participant flow, recruitment, numbers analysed

Risk of bias

Item	**Authors' judgement**	**Description**
Large Cohort ?	No	RCTs in 1 endorsing and 1 non-endorsing journal 3 years prior and post CONSORT publication
Blinding?	Unclear	Not reported
Confounding by journal quality?	Yes	Not applicable
Outcome Reporting?	Yes	No difference between planned and reported outcomes
Multiple raters?	Yes	2 raters
Rater agreement?	Unclear	Not reported
Blinding, quality assessment?	Yes	Not applicable

Kidwell 2001

Methods	Objective of study was to quantitatively characterise developments in clinical trial methodology over time in the field of acute ischaemic stroke A formal 100-point scale was used to rate trial quality and unweighted totals for CONSORT endorsers and non-endorsers was reported in the text
Data	34 RCTs included for our analysis, 9 endorsing journals and 25 non-endorsing journals
Comparisons	CONSORT endorsers versus non-endorsers
Outcomes	Total sum score on 100-point scale
Included number of RCTs, Journals	178, not reported in text
Checklist version used	1996
Field of Study	Stroke
Notes	This study was obtained from an external source, Cochrane Colloquium 2010 CONSORT-endorsing journal was not defined and was not confirmed to coincide with our definition Please note that the standard deviation for this study was imputed

Risk of bias

Item	**Authors' judgement**	**Description**
Large Cohort ?	Yes	178 articles, 40 years
Blinding?	Unclear	Not reported
Confounding by journal quality?	Unclear	Not reported
Outcome Reporting?	Unclear	No apparent difference between planned and reported outcomes
Multiple raters?	No	Full extraction was not verified
Rater agreement?	Yes	Validity assessment conducted, kappa > 0.9
Blinding, quality assessment?	Yes	Not applicable

Kober 2006

Methods	Aims to determine the extent of ambiguity and reporting quality as assessed by completeness of reporting to the CONSORT Statement in published reports of RCTs involving patients with Hodgkin lymphoma from 1966 through 2002 Quality of reporting was assessed using a 14-item questionnaire based on the CONSORT checklist Reporting was studied in 2 pre-CONSORT periods (1966-1988 and 1989-1995) and one post CONSORT period (1996-2002)
Data	77 RCTs eligible for inclusion in our study
Comparisons	CONSORT endorsers versus non-endorsers, CONSORT endorsers before and after CONSORT endorsement
Outcomes	Endorsers versus non-endorsers: title and abstract, introduction, interventions, outcomes, sample size, sequence generation, allocation concealment, statistical methods, participant flow, numbers analysed, outcomes and estimation, adverse events CONSORT endorsers before and after: title and abstract, introduction, interventions, outcomes, sample size, sequence generation, statistical methods, participant flow, numbers analysed, outcomes and estimation, adverse events
Included number of RCTs, Journals	243, 33
Checklist version used	1996
Field of Study	Hodgkin lymphoma
Notes	Author provided data in necessary format Endorsing journals are consistent with our definition

Risk of bias

Item	**Authors' judgement**	**Description**
Large Cohort ?	Yes	Multiple databases over 1-month period (May 2003)
Blinding?	Unclear	Not reported
Confounding by journal quality?	Yes	Quote: "Clustering of articles in a journal or by study group was not taken into account in the analyses."
Outcome Reporting?	Yes	No differences between planned and reported outcomes/analyses
Multiple raters?	Unclear	Not reported
Rater agreement?	Yes	Not applicable

Kober 2006 (*Continued*)

Blinding, quality assessment?	Yes	Not applicable

Ladd 2010

Methods	Aim of this study was to examine if adopting CONSORT standards of reporting improved the quality of reporting of alcohol treatment outcome studies RCTs were identified from 8 journals publishing a substantial number of alcohol treatment outcome studies (n = 127 RCTs) and coded for the quality of reporting according to the CONSORT guidelines
Data	Pre-CONSORT 70 RCTs, post CONSORT 89 RCTs, 1 endorsing journal of 19 RCTs and 108 RCTs from non-endorsing journals
Comparisons	CONSORT endorsers versus non-endorsers, cross-sectional sample before and after CONSORT publication
Outcomes	CONSORT endorsers before and after endorsement: title and abstract, background, interventions, outcomes, sequence generation, allocation concealment, statistical methods, participant flow, numbers analysed, outcomes and estimation, adverse events Endorsers versus non-endorsers: title and abstract, introduction, objectives, outcomes, sample size, sequence generation, blinding any, statistical methods, participant flow, numbers analysed, outcomes and estimation, ancillary analyses, interpretation, generalisability and overall evidence
Included number of RCTs, Journals	127, 8
Checklist version used	2001, 1996 comparison for pre-post
Field of Study	Alcohol outcome studies
Notes	Author provided data for the review; some dates of journal endorsement provided by MEs are vague; these have been conservatively categorised as non-endorsers; in turn, definition is compliant for this study For before and after, 3 time periods reported; to allow for improvement in quality of reporting over time, conservatively, we included 1989-1995 and 1996-2002 in our analysis

Risk of bias

Item	**Authors' judgement**	**Description**
Large Cohort ?	Yes	Large number of trials from 8 journals over a long time period
Blinding?	Unclear	Quote: "It was not feasible to mask year published and author due to high rates of self-citation and dates in reference lists. However, names of the source journals for each article

Ladd 2010 (*Continued*)

		were concealed from coders"
Confounding by journal quality?	Yes	Stratified analysis. Quote: "Studies published pre-CONSORT (1994-1998) did not differ significantly on overall CONSORT score between adopter and nonadopter journals"
Outcome Reporting?	Yes	No evidence of selective outcome reporting
Multiple raters?	Yes	Quote: "Four coders (the four authors) coded the articles for this study. Twenty percent of studies were randomly selected to be double-coded throughout the coding process"
Rater agreement?	Unclear	Not reported
Blinding, quality assessment?	Yes	Not applicable

Lai 2006

Methods	Assesses the reporting quality of randomised controlled trials (RCTs) in the primary treatment of brain tumours and aimed to identify significant predictors of quality in trials published between 1990 and 2004 using items from the CONSORT checklist
Data	23 RCTs pre-CONSORT (1990-1994) and 32 RCTs post CONSORT (2000-2004) Score out of 15
Comparisons	Cross-sectional sample of before and after CONSORT publication
Outcomes	Total sum score
Included number of RCTs, Journals	74, 26
Checklist version used	2001 (1996 intervention)
Field of Study	Brain tumours
Notes	Author provided study data Median overall score

Risk of bias

Item	**Authors' judgement**	**Description**
Large Cohort ?	Yes	Reasonable number of assessed trials, from 4 journals, over 15 years
Blinding?	Unclear	Not reported

Lai 2006 (*Continued*)

Confounding by journal quality?	Unclear	Not reported
Outcome Reporting?	Yes	No evidence of selective outcome reporting
Multiple raters?	Yes	Quote: "Two trained investigators (R.L. and R.C.) who were blinded to each other's ratings abstracted data independently into a standardized data abstraction form, which was pilot tested on 15 studies and subsequently was revised"
Rater agreement?	Yes	Quote: "The overall inter-rater agreement was 0.83"
Blinding, quality assessment?	Yes	Not applicable

Lai 2007

Methods	Evaluates the reporting quality of key methodological items in randomised controlled trials (RCTs) in 4 general clinical ophthalmology journals The reporting of 11 key methodological items in RCTs published in *American Journal of Ophthalmology*, *Archives of Ophthalmology*, *British Journal of Ophthalmology* and *Ophthalmology* in the year 2005 was assessed
Data	51 CONSORT-endorsing RCTs from 3 journals and 16 non-endorsing RCTs from 1 journal
Comparisons	CONSORT endorsers versus non-endorsers
Outcomes	Sample size, sequence generation, allocation concealment, implementation, blinding any, participant flow, numbers analysed, adverse events
Included number of RCTs, Journals	67, 4
Checklist version used	2001
Field of Study	Opthalmology
Notes	

Risk of bias

Item	**Authors' judgement**	**Description**
Large Cohort ?	Unclear	RCTs published in top 4 journals in subspecialty over 1-year period - 67 included
Blinding?	Unclear	Not reported

Lai 2007 (*Continued*)

Confounding by journal quality?	Yes	Top 4 impact factor journals in subspecialty - no potential for clustering
Outcome Reporting?	Yes	No difference between planned and reported outcomes/analyses
Multiple raters?	Yes	Quote: "Each of the eligible RCTs was evaluated by two of the authors independently according to the revised CONSORT statement. "
Rater agreement?	Unclear	Not reported
Blinding, quality assessment?	Yes	Not applicable

Llorca 2004

Methods	To study the quality of controlled clinical trials on glaucoma using 11 key methodological items
Data	37 RCTs published in endorsing journals
Comparisons	CONSORT endorsers versus non-endorsers, CONSORT endorsers before and after CONSORT publication
Outcomes	CONSORT endorsers versus non-endorsers: sequence generation, allocation concealment, blinding any, statistical methods, participant flow, recruitment, numbers analysed CONSORT endorsers before and after CONSORT endorsement: sequence generation, allocation concealment, blinding of participants, blinding any, statistical methods, participant flow, recruitment, numbers analysed
Included number of RCTs, Journals	226, 7
Checklist version used	Modification
Field of Study	Glaucoma and intraocular high pressure
Notes	Author provided data. This study also includes control comparison data, namely: Pre-post CONSORT non-endorsers: blinding of participants, statistical methods, participant flow, recruitment, numbers analysed Pre-CONSORT endorsers versus non-endorsers: sequence generation, allocation concealment, blinding any description and blinding of outcome assessor, statistical methods, participant flow, recruitment, numbers analysed

Risk of bias

Item	**Authors' judgement**	**Description**

Llorca 2004 (*Continued*)

Large Cohort ?	Yes	Large sample of trials from 7 journals over 2 decades
Blinding?	Unclear	Not reported
Confounding by journal quality?	Unclear	Not reported
Outcome Reporting?	Yes	No evidence of selective outcome reporting
Multiple raters?	Yes	Quote: "Each paper was revised by 2 of 4 researchers with epidemiological skills. Discrepancies between reviewers were solved by consensus"
Rater agreement?	Unclear	Not reported
Blinding, quality assessment?	Unclear	Not reported

Moher 2001

Methods	Aims to determine whether or not the CONSORT Statement is associated with improvement in the quality of reports of RCTs RCTs published in 1994 and 1998, with non-endorsing journal acting as a control
Data	71 endorsing RCTs in 1994 from 3 journals and 26 non-endorsing from 1 journal, 77 endorsing RCTs from 3 journals in 1998 and 37 non-endorsing RCTs from one journal; the 3 journals include *BMJ*, *JAMA* and *The Lancet* compared to the *NEJM*
Comparisons	CONSORT endorsers versus non-endorsers, CONSORT-endorsing journals before and after endorsement
Outcomes	Allocation concealment, total sum score based on 40 items
Included number of RCTs, Journals	211, 4
Checklist version used	1996
Field of Study	None specified
Notes	This study was included in the original review Author provided data Jadad scale was also used to assess quality This study also includes control comparison data: allocation concealment for both pre-CONSORT endorsers versus non-endorsers and pre-post CONSORT non-endorsers
Risk of bias	

Moher 2001 (*Continued*)

Item	Authors' judgement	Description
Large Cohort ?	Yes	4 journals from 2 years, samples 4 years apart in large number of trials
Blinding?	No	Quote: "Hard copies of relevant articles were obtained but were not masked because evidence concerning the effect of masking on assessments of trial quality is inconsistent"
Confounding by journal quality?	Yes	Study included control group comparison to assess clustering by quality
Outcome Reporting?	Yes	No evidence of selective outcome reporting
Multiple raters?	Yes	Quote: "Two reviewers (A.J., L.L.) completed all of these evaluations."
Rater agreement?	Yes	Quote: "A k statistic was calculated for each item based on a randomly selected set of 10 RCTs, from 1994 and 1998, and these were not included in this study"
Blinding, quality assessment?	Yes	Not applicable

Montané 2010

Methods	Aimed to examine the quality of reporting RCTs on analgesics for postoperative pain after traumatic or orthopaedic surgery The quality of reports was assessed using the CONSORT checklist (scoring range from 0 to 22)
Data	92 included RCTs
Comparisons	Qualitative description
Outcomes	Insufficient data to include in quantitative synthesis
Included number of RCTs, Journals	92, 46
Checklist version used	2001
Field of Study	Surgery - analgesic drugs post orthopedic surgery
Notes	
Risk of bias	

Montané 2010 (*Continued*)

Item	Authors' judgement	Description
Large Cohort ?	Yes	Multiple databases; 40 years, 92 included studies
Blinding?	Unclear	Not reported
Confounding by journal quality?	Unclear	Not reported
Outcome Reporting?	Yes	No difference between planned and reported outcomes
Multiple raters?	Yes	Quote: "The quality of reporting of each included study in the reports was assessed independently by 3 evaluators (EM, AV, CA) with CONSORT checklist [6]"
Rater agreement?	No	The agreement (ICC) between the 3 evaluators for the overall scores of the CONSORT checklist assessed was 0.77 (95% CI 0.70 to 0.84)
Blinding, quality assessment?	Yes	Not applicable

Montori 2002

Methods	Assessed the quality of reporting in RCTs 4 endorsing and one non-endorsing journal
Data	40 RCTs per journal were sampled, hence, 40 non-endorsing RCTs compared with 160 endorsing RCTs
Comparisons	CONSORT endorsers versus non-endorsers
Outcomes	Blinding: participant, outcome assessor, data analyst, intervention
Included number of RCTs, Journals	200, 5
Checklist version used	Modification
Field of Study	None specified
Notes	This study was included in the original review Endorsement coincides with our definition
Risk of bias	

Montori 2002 (*Continued*)

Item	Authors' judgement	Description
Large Cohort ?	Yes	Wide range of journals searched
Blinding?	Unclear	Not reported
Confounding by journal quality?	Yes	Quote: "We only evaluated very recent RCTs published in five leading general medicine journals. Thus, our findings may not represent reporting in journals with less editorial resources."
Outcome Reporting?	Unclear	No evidence to suggest selective reporting of outcomes
Multiple raters?	Yes	2 authors assessed all criteria
Rater agreement?	Yes	Quote: "Kappa, a measure of interobserver agreement, was between 0.8 and 1.0 for each of the variables assessed."
Blinding, quality assessment?	Yes	Not applicable

Pagoto 2009

Methods	Study aimed to determine whether reporting and correct use of ITT in behavioural medicine randomised clinical trials (RCTs) published in behavioural journals has improved in recent years and since the endorsement of CONSORT
Data	Includes 50 RCTs pre-CONSORT from 3 journals 2000-2003 and 37 post CONSORT RCTs from the same 3 journals 2006-2007
Comparisons	CONSORT-endorsing journals before and after CONSORT endorsement
Outcomes	Outcomes, sample size, baseline data, numbers analysed
Included number of RCTs, Journals	87, 3
Checklist version used	2001
Field of Study	Behavioural medicine
Notes	CONSORT endorsement dates confirmed with journals This is the primary study to the companion, Spring 2007
Risk of bias	

Pagoto 2009 (*Continued*)

Item	Authors' judgement	Description
Large Cohort ?	Unclear	3 journals, over 6 years
Blinding?	No	Quote: "As in other reviews of quality reporting, it was not deemed necessary to mask the articles."
Confounding by journal quality?	Yes	Not applicable
Outcome Reporting?	Yes	No evidence of selective reporting
Multiple raters?	Yes	Quote: "Each article was reviewed independently by two assessors"
Rater agreement?	No	Average of 82% reported
Blinding, quality assessment?	Yes	Not applicable

Partsinevelou 2009

Methods	Purpose of study to assess the reporting quality of RCTs involving patients with polycystic ovary syndrome using a standardised tool based on CONSORT Quality of reporting was assessed using a 24-item questionnaire based on the revised CONSORT checklist Reporting was evaluated overall and for pre- and post CONSORT periods (1990-1995 and 1996-2008)
Data	27 pre-CONSORT RCTs and 237 post CONSORT RCTs
Comparisons	Cross-sectional sample before and after publication
Outcomes	Title and abstract, background, participants, interventions, objectives, outcomes, sample size, sequence generation, allocation concealment, implementation, blinding any, statistical methods, participant flow, recruitment, baseline data, numbers analysed, outcomes and estimation, ancillary analyses, adverse events, interpretation, generalisability, overall evidence
Included number of RCTs, Journals	264, 57
Checklist version used	2001
Field of Study	Polycystic ovary syndrome
Notes	Endorsers listed as those on the CONSORT website 45 journals did not endorse the CONSORT Statement when included in this study, hence, not all post CONSORT are endorsers

Partsinevelou 2009 (*Continued*)

Risk of bias		
Item	**Authors' judgement**	**Description**
Large Cohort ?	Yes	Wide search of PubMed over 18 years
Blinding?	Unclear	Not reported
Confounding by journal quality?	Unclear	Not reported
Outcome Reporting?	Yes	No evidence of selective outcome reporting
Multiple raters?	Yes	2 authors assessed
Rater agreement?	Yes	Kappa reported 0.92
Blinding, quality assessment?	Yes	Not applicable

Parés 2008

Methods	This study was designed to analyse the characteristics and the quality of reporting of RCTs published during the last 10 years on fecal incontinence Quality was assessed by characteristics of reporting, methodology quality assessment using the Jadad scale, and a validated methodology quality score (MINCIR score), evaluation of the items published in the CONSORT Statement, and the journal impact factor Reports were divided into 2 groups: 1996 to 2000 (Group 1) and from 2001 to 2005 (Group 2)
Data	15 RCTs were assessed in group 1 and 27 RCTs in group 2
Comparisons	Cross-sectional sample before and after CONSORT publication
Outcomes	Title and abstract, introduction, interventions, sample size, sequence generation, allocation concealment, statistical methods, participant flow, numbers analysed, outcomes and estimation, adverse events, interpretation
Included number of RCTs, Journals	42, 22
Checklist version used	2001
Field of Study	Fecal Incontinence
Notes	Also considers Jadad score and MINCIR score

Risk of bias		
Item	**Authors' judgement**	**Description**

Parés 2008 (*Continued*)

Large Cohort ?	Yes	RCTs in PubMed from 1996 to 2005
Blinding?	Unclear	Not clearly reported
Confounding by journal quality?	Unclear	Unsure if clustering by CONSORT
Outcome Reporting?	Yes	No difference between planned and reported outcomes
Multiple raters?	Unclear	Not reported
Rater agreement?	Yes	Not applicable
Blinding, quality assessment?	Yes	Not applicable

Pat 2008

Methods	Aims to analyse to what extent the different RCTs with information on PROs adhere to the CONSORT Statement Compliance with the (revised) CONSORT Statement was checked by 2 independent reviewers by making for each study the simple sum of the 22 CONSORT items, or a weighted score with a maximum rating of 31 points
Data	4 CONSORT-endorsing RCTs and 34 non-endorsing
Comparisons	CONSORT endorsers versus non-endorsers
Outcomes	Total sum score
Included number of RCTs, Journals	38, 7
Checklist version used	2001
Field of Study	Chemotherapy for non-small cell lung cancer
Notes	Author provided data of simple score by RCT One journal was not included in analysis as endorsement could not be confirmed for the dates of publication for the included RCTs Complient with our endorser definition

Risk of bias

Item	**Authors' judgement**	**Description**
Large Cohort ?	Yes	Pubmed between 1980 and 2005
Blinding?	Unclear	Not reported

Pat 2008 (*Continued*)

Confounding by journal quality?	Unclear	Not reported
Outcome Reporting?	Yes	No difference between planned and reported outcomes/analyses
Multiple raters?	Yes	2 raters assessed CONSORT score
Rater agreement?	Unclear	Not reported
Blinding, quality assessment?	Yes	Not applicable

Peckitt 2007

Methods	A systematic review comparing early breast cancer (EBC) RCTs pre- and post introduction of CONSORT in systemic treatment was undertaken in part to assess the association between the introduction of CONSORT and the publication quality
Data	0.5 scores given to partially reported items; these frequencies were not included in our analysis
Comparisons	Cross-sectional sample before and after CONSORT publication
Outcomes	Introduction, objectives, sample size, outcomes and estimation
Included number of RCTs, Journals	85, not reported
Checklist version used	1996
Field of Study	Breast cancer
Notes	Data published in abstract, unable to make contact with author

Risk of bias

Item	**Authors' judgement**	**Description**
Large Cohort ?	Yes	Wide database search for RCTs over multiple years
Blinding?	Unclear	Not reported
Confounding by journal quality?	Unclear	Not reported
Outcome Reporting?	Unclear	No evidence of selective reporting, information limited as abstract for poster presentation
Multiple raters?	Unclear	Not reported

Peckitt 2007 (*Continued*)

Rater agreement?	Unclear	Not reported
Blinding, quality assessment?	Unclear	Not reported

Prady 2008

Methods	Conducted a before-and-after study, comparing ratings for quality of reporting following the publication of both STRICTA and CONSORT recommendations 90 peer-reviewed journal articles reporting the results of acupuncture trials were selected at random from a wider sample frame of 266 papers Papers published in 3 distinct time periods (1994-1995, 1999-2000, and 2004-2005) were compared
Data	Pre 2001 groups were collapsed
Comparisons	Cross-sectional sample before and after CONSORT publication
Outcomes	Sequence generation, allocation concealment, blinding: outcome assessor, intervention, participant, baseline data, number analysed
Included number of RCTs, Journals	90, 52
Checklist version used	2001
Field of Study	Acupuncture
Notes	Author provided data Score by journal available, unable to determine endorsement for all journals

Risk of bias

Item	**Authors' judgement**	**Description**
Large Cohort ?	Yes	Multiple databases, 3 1-year cross-sections
Blinding?	Unclear	Quote: "Efforts were made to guard against the possible introduction of systematic bias. In order to assess whether knowledge of publication period, journal type or authorship might affect scoring, all papers given to SJR had this information removed. This was achieved by censoring all pertinent material with a black marker pen or blank paper prior to photocopying. SJR also remained unaware of the three date ranges from which papers were drawn. Blinding of the other assessor (SLP) was not possible due to practical reasons, and

Prady 2008 (*Continued*)

		she was already familiar with the research literature relating to acupuncture."
Confounding by journal quality?	Unclear	Not reported
Outcome Reporting?	Yes	No difference between planned and reported outcomes/analyses
Multiple raters?	Yes	2 reviewers assessed or 1 extractor and verification
Rater agreement?	Yes	Quote: "There was a high degree of concordance (kappa 0.8) between assessors in terms of their scoring for the majority of STRICTA (17/31) and CONSORT (6/8) checklist items."
Blinding, quality assessment?	Yes	Not applicable

Sanchez-Thorin 2001

Methods	Assesses if structured abstract use is associated with improved reporting of RCTs
Data	56 items derived from the CONSORT checklist, comparison of 51 1991-1993 RCTs and 24 RCTs published in 1999
Comparisons	CONSORT endorsers before and after endorsement
Outcomes	Title and abstract, participants, interventions, objectives, outcomes, sample size, sequence generation, allocation concealment, implementation, blinding: interventions, outcome assessor, participants, data analyst, statistical methods, participant flow, numbers analysed, outcomes and estimation, overall evidence
Included number of RCTs, Journals	75, 1
Checklist version used	1996
Field of Study	Opthalmology
Notes	This study was included in the original review

Risk of bias

Item	**Authors' judgement**	**Description**
Large Cohort ?	No	1 volume of 1 journal sampled
Blinding?	Unclear	Not reported

Sanchez-Thorin 2001 (*Continued*)

Confounding by journal quality?	Yes	Not applicable
Outcome Reporting?	Unclear	No evidence of selective reporting of outcomes
Multiple raters?	Yes	Quote: "Each study was evaluated by two independent observers"
Rater agreement?	Unclear	Not reported
Blinding, quality assessment?	Yes	Not applicable

Scales 2007

Methods	Reports a systematic assessment of RCT quality in the urology literature by compliance with CONSORT
Data	87 pre-CONSORT RCTs and 65 post CONSORT RCTs were included
Comparisons	Cross-sectional sample before and after CONSORT publication
Outcomes	Sample size, allocation concealment, implementation, blinding: intervention, participant, outcome assessor, participant flow, number analysed
Included number of RCTs, Journals	152, 4
Checklist version used	2001
Field of Study	Urology
Notes	Endorsement status could not be confirmed for the 4 journals

Risk of bias

Item	**Authors' judgement**	**Description**
Large Cohort ?	Yes	Cross-section of 2 years on MEDLINE
Blinding?	Unclear	Not reported
Confounding by journal quality?	Unclear	Not reported
Outcome Reporting?	Yes	No difference in planned and reported outcomes/analyses
Multiple raters?	Yes	Quote: "As determined by the 2 reviewers, the assessment of each criterion was entered into a dedicated study database."

Scales 2007 (*Continued*)

Rater agreement?	Unclear	Not reported
Blinding, quality assessment?	Yes	Not applicable

Selman 2008

Methods	Assess if there has been progress made in establishing the evidence base for surgical interventions in gynaecology Quality was assessed for pre- and post CONSORT Pre-CONSORT 1974-1996 publication intervention, 1998-2005 post CONSORT
Data	39 pre-CONSORT RCTs compared with 35 post CONSORT RCTs
Comparisons	Cross-sectional sample before and after CONSORT publication
Outcomes	Allocation concealment
Included number of RCTs, Journals	74, not reported
Checklist version used	1996
Field of Study	Gynaecologic surgery
Notes	This study was excluded prior to the inclusion of the cross-sectional sample comparison group and re-included RCTs obtained from 23 reviews published in *The Cochrane Library*

Risk of bias

Item	**Authors' judgement**	**Description**
Large Cohort ?	No	RCTs included in Cochrane systematic reviews 2006(Issue 3); only relevant reviews selected, but no selection criteria given
Blinding?	Unclear	Not reported
Confounding by journal quality?	Unclear	Not reported
Outcome Reporting?	Yes	No differences in planned and reported outcomes/analyses
Multiple raters?	Unclear	Not reported
Rater agreement?	Unclear	Not applicable

Selman 2008 (*Continued*)

Blinding, quality assessment?	Unclear	Not applicable

Sinha 2009

Methods	Assesses the quality of reporting of trial methodology and adverse events in a sample of general surgical RCTs published in high-quality surgical journals using the criteria specified in the CONSORT Statements
Data	Not reported in needed format
Comparisons	Qualitative comparison of CONSORT endorsers versus non-endorsers using the Jadad score detailed in study
Outcomes	Included for qualitative analysis
Included number of RCTs, Journals	42, 3
Checklist version used	2001
Field of Study	Surgery
Notes	Also considered Jadad score to assess quality

Risk of bias

Item	**Authors' judgement**	**Description**
Large Cohort ?	No	Relatively few assessed studies from journals selected by impact factor
Blinding?	Unclear	Not reported
Confounding by journal quality?	Unclear	Not reported
Outcome Reporting?	Yes	No evidence of selective reporting of outcomes
Multiple raters?	Yes	2 authors independently reviewed methodological items
Rater agreement?	No	Quote: "Agreement between the pair of observers who independently assessed the RCTs was good (median K 0.795; range 0.4 to 1)"
Blinding, quality assessment?	Unclear	Not reported

Spring 2007

Methods	Compared analytic quality features of all behavioural health RCTs (n = 73) published in 3 leading behavioural journals and 2 leading medical journals between January 2000 and July 2003
Data	15 endorsing RCTs and 58 non-endorsing RCTs
Comparisons	CONSORT endorsers versus non-endorsers
Outcomes	Outcomes, sample size, participant flow, numbers analysed
Included number of RCTs, Journals	73, 5
Checklist version used	Modification of 2001 version
Field of Study	Behavioural health
Notes	This is the companion study to Pagoto 2009 Provides supplementary outcomes data, included in a different comparison Endorsement of journals confirmed

Risk of bias

Item	**Authors' judgement**	**Description**
Large Cohort ?	Unclear	5 journals over 3 years, judgement based on number of RCTs in endorsing group
Blinding?	Unclear	Quote: "It was not deemed necessary to mask the articles."
Confounding by journal quality?	No	Quote: "Perhaps if mental health had been the outcome, the analytic quality of RCTs reported in psychology journals might have been superior because of the longer history of studying that content area in psychology"
Outcome Reporting?	Yes	No evidence of selective outcome reporting
Multiple raters?	Yes	Quote: "Each article was reviewed and coded by two people, using all possible combinations of pairs of rater"
Rater agreement?	No	Quote: "Average intercoder agreement across the 73 articles was 85% prior to resolving discrepant rating"
Blinding, quality assessment?	Unclear	Not reported

Thabane 2007

Methods	Assesses the quality of reporting of RCTs of weight loss interventions and to identify predictors of reporting quality The RCTs assessed were derived from a published systematic review of trials investigating the efficacy of weight loss interventions Quality based on CONSORT items; 44-item score was detailed
Data	50 pre-CONSORT RCTs from 23 journals and 13 post CONSORT RCTs from 10 journals
Comparisons	Cross-sectional sample before and after CONSORT publication in 1996
Outcomes	Title and abstract, introduction, participants, interventions, objectives, outcomes, sample size, sequence generation, allocation concealment, implementation, blinding of participants, statistical methods, participant flow, recruitment, baseline data, outcomes and estimation, adverse events, interpretation, generalisability, overall evidence
Included number of RCTs, Journals	63, 28
Checklist version used	1996
Field of Study	Weight loss
Notes	Author provided data for this study RCTs published in 2001 were included as pre-CONSORT as a conservative estimate This is the primary study of the companion Thoma 2006

Risk of bias

Item	Authors' judgement	Description
Large Cohort ?	No	All RCTs identified from a single systematic review
Blinding?	Unclear	Not reported
Confounding by journal quality?	Yes	Quote: "GEEs were chosen to account for the possible intrajournal correlation and an exchangeable correlation structure was assumed for these analyses"
Outcome Reporting?	Yes	No difference between planned and reported outcomes/analyses
Multiple raters?	Yes	Quote: "An independent double review of included trials was done by two authors (RC, KC) to assess agreement regarding CONSORT criteria that were satisfied."
Rater agreement?	Unclear	Not reported

Thabane 2007 (*Continued*)

Blinding, quality assessment?	Yes	Not applicable

Tharyan 2008

Methods	Examines the extent to which CONSORT has been adopted by Indian medical journals RCTs published during 2004 and 2005 were assessed against selected CONSORT items and ICMJE requirements, and scored on the Jadad scale
Data	31 endorsing RCTs and 120 non-endorsing RCTs
Comparisons	CONSORT endorsers versus non-endorsers
Outcomes	Total sum score on an unweighted score out of 13
Included number of RCTs, Journals	151, 37
Checklist version used	2001
Field of Study	Indian medical journals
Notes	With additional data, this could have been included by item Instructions to authors were searched by study authors, but we were unable to confirm due to insufficient information This study also reports Jadad scores

Risk of bias

Item	**Authors' judgement**	**Description**
Large Cohort ?	Yes	RCT reports published between 2004-5 in Indian medical journals; 151 included studies
Blinding?	Unclear	Not reported
Confounding by journal quality?	Unclear	Not reported
Outcome Reporting?	Yes	No differences between planned and reported outcomes/analyses
Multiple raters?	Yes	Single extraction with verification
Rater agreement?	Unclear	Not reported
Blinding, quality assessment?	Unclear	Not applicable

Thoma 2006

Methods	Assesses the reporting quality of published RCTs that compare endoscopic carpal tunnel release (ECTR) with open carpal tunnel release (OCTR) using the CONSORT Statement
Data	Studies published between 1989 and 2004 Before and after 1996 comparison with 11 RCTs published before 1997 and 7 after 1996
Comparisons	Cross-sectional sample before and after CONSORT publication
Outcomes	Total sum score
Included number of RCTs, Journals	18, not reported
Checklist version used	2001
Field of Study	ECTR and OCTR
Notes	This is the companion study of Thabane 2007 No journal information provided

Risk of bias

Item	Authors' judgement	Description
Large Cohort ?	Yes	Wide database search over 15 years
Blinding?	Unclear	Not reported
Confounding by journal quality?	Unclear	Not reported
Outcome Reporting?	Yes	No evidence of outcome reporting bias
Multiple raters?	Yes	Quote: "Two investigators (RTC and KV) independently reviewed each articles"
Rater agreement?	Yes	Quote: "...kappa value of 0.90"
Blinding, quality assessment?	Yes	Not applicable

Tiruvoipati 2005

Methods	Evaluates the quality of reporting of RCTs in cardiothoracic surgery, to identify factors associated with good reporting quality, and assesses the awareness of CONSORT and ascertains the views of authors reporting RCTs on the difficulties in conducting RCTs and the possible ways to further improve the reporting quality of randomised controlled trials in cardiothoracic surgery
Data	2 endorsing RCTs and 62 non-endorsing RCTs from 4 journals published in 2003

Tiruvoipati 2005 (*Continued*)

Comparisons	CONSORT endorsers versus non-endorsers
Outcomes	Total sum score
Included number of RCTs, Journals	64, 4
Checklist version used	2001
Field of Study	Cardiothoracic surgery
Notes	Potential for overlap of 2 RCTs from *NEJM* with Balasubramanian 2006. As this is not confirmed, we have not listed these studies as companions. This study has not been confirmed to be compliant with our endorser definition Median score out of 90 reported; Jadad scores were also reported

Risk of bias

Item	**Authors' judgement**	**Description**
Large Cohort ?	Unclear	RCTs in cardiology published in 4 top journals over 1 year period, n = 64
Blinding?	Unclear	Not reported
Confounding by journal quality?	Unclear	Not reported
Outcome Reporting?	Yes	No difference between planned and reported outcomes/analyses
Multiple raters?	Yes	Quote: "Each article included in the study was then assessed for every item on the checklist and scored independently by 2 observers (R.T. and S.P.B.) to arrive at a consensus-modified CONSORT score."
Rater agreement?	Yes	Quote: "The agreement of the pair of observers who independently assessed the RCTs by using the CONSORT checklist was good (intra-class correlation coefficient, 0.85; 95% confidence interval, 0.76-0.90; P .001)."
Blinding, quality assessment?	Unclear	Not reported

Uetani 2009

Methods	Assesses the quality of Japanese RCT reports by conducting a cross-sectional study to examine the extent to which they adhere to the CONSORT Statement Sample of 98 RCTs published in 2004
Data	11 endorsing RCTs and 87 non-endorsing RCTs
Comparisons	CONSORT endorsers versus non-endorsers
Outcomes	Title and abstract, background, participants, interventions, objectives, outcomes, sample size, sequence generation, allocation concealment, implementation, blinding any, statistical methods, participant flow, recruitment, baseline data, numbers analysed, outcomes and estimation, ancillary analyses, adverse events, interpretation, generalisability, overall evidence
Included number of RCTs, Journals	98, not reported
Checklist version used	2001
Field of Study	Japanese trials
Notes	Journal endorsement determined from CONSORT website, however these have not been checked against RCT publication dates to ensure that our definition of CONSORT endorser coincides

Risk of bias

Item	**Authors' judgement**	**Description**
Large Cohort ?	Unclear	Sample of journals not reported
Blinding?	Unclear	Not reported
Confounding by journal quality?	Unclear	Not reported
Outcome Reporting?	Yes	No evidence of selective reporting
Multiple raters?	Yes	Quote: "We adopted the "checked by a second" method, which is recognized as a systematic review methodology"
Rater agreement?	Unclear	Not reported
Blinding, quality assessment?	Yes	Not applicable

Wang 2007

Methods	Evaluates the quality of reporting of RCTs in TCM journals published in 1999 and 2004
Data	Reported by years 1999 to 2004 and CONSORT item in the paper This has been sorted into pre-CONSORT (1999-2001) and post CONSORT (2002-2004), with 2930 and 4492 RCTs respectively
Comparisons	Cross-sectional sample of RCTs before and after CONSORT publication
Outcomes	Title and abstract, background, participants, interventions, objectives, outcomes, sample size, sequence generation, allocation concealment, blinding of participants, statistical methods, participant flow, recruitment, baseline data, outcomes and estimation, ancillary analyses, adverse events, interpretation, generalisability, overall evidence
Included number of RCTs, Journals	7422, 13
Checklist version used	Modification of the 2001 checklist
Field of Study	Traditional Chinese Medicine
Notes	This study also assesses quality based on the Jadad score This study was found externally to the search Score out of 30 items It should be noted that this study was conducted on behalf of the CONSORT group for TCM

Risk of bias

Item	**Authors' judgement**	**Description**
Large Cohort ?	Yes	RCTs published in 13 journals over 5 years (1999-2004)
Blinding?	Unclear	Not reported
Confounding by journal quality?	Unclear	Not explicitly discussed
Outcome Reporting?	Yes	No differences between planned and reported outcomes and analyses
Multiple raters?	Yes	Quote: "Data extraction and the evaluation of methodologic quality were performed independently by 2 reviewers"
Rater agreement?	No	Agreement was high (> 0.70), indicating low interobserver and intraobserver variability
Blinding, quality assessment?	Unclear	Not reported

Wei 2009

Methods	Evaluates the reporting quality of RCTs on papers published in 5 leading Chinese medical journals by assessing adherence to CONSORT
Data	35 endorsing RCTs from 1 journal and 188 non-endorsing RCTs from 4 journals
Comparisons	CONSORT endorsers versus non-endorsers
Outcomes	Title and abstract, background, participants, interventions, objectives, outcomes, sample size, sequence generation, allocation concealment, implementation, blinding any, statistical methods, participant flow, recruitment, baseline data, numbers analysed, outcomes and estimation, ancillary analyses, adverse events, interpretation, generalisability
Included number of RCTs, Journals	123, 5
Checklist version used	2001
Field of Study	Chinese medical journals
Notes	Author provided data for post 2004 data from which a comparison group compliant with our definition could be formed

Risk of bias

Item	**Authors' judgement**	**Description**
Large Cohort ?	Unclear	Only one journal in endorsing arm
Blinding?	Yes	Assessors were blinded
Confounding by journal quality?	Unclear	Potential clustering based on study design; this was not discussed by the author
Outcome Reporting?	Yes	No evidence of selective outcome reporting
Multiple raters?	Yes	2 independently assessed items
Rater agreement?	No	0.61 reported interobserver agreement
Blinding, quality assessment?	Yes	Not applicable

Yu 2010

Methods	Evaluates the use and reporting of adjusted analysis in randomised controlled trials (RCTs) and compares the quality of reporting before and after the revision of the CONSORT Statement in 2001

Yu 2010 (*Continued*)

Data	Journal articles sampled from 2000 and 2006 355 RCTs pre-CONSORT and 422 RCTs post CONSORT 113 RCTs described as endorsing journals and 48 described as non-endorsing
Comparisons	CONSORT endorsers versus non-endorsers, cross-sectional sample before and after publication
Outcomes	Ancillary analyses
Included number of RCTs, Journals	777, not reported
Checklist version used	2001
Field of Study	None specified
Notes	This is a companion study to Hopewell 2010 Journal endorsement was not confirmed 6 months prior to the publication of each RCT, hence, this does not strictly comply with our definition of CONSORT-endorsing journal

Risk of bias

Item	**Authors' judgement**	**Description**
Large Cohort ?	Yes	Large number of included trials from wide sample
Blinding?	Unclear	Not reported
Confounding by journal quality?	No	Quote: "We identified slightly better reporting of key methodological items in CONSORT endorsing as opposed to non CONSORT endorsing journals. However, because there was a time-lag between article publication (December 2006) and when the journal 'Instructions to Authors' were assessed (June 2008) these results should be viewed with some caution."
Outcome Reporting?	Yes	No evidence of selective reporting of outcomes
Multiple raters?	No	Quote: "Data regarding trial characteristics were extracted by two reviewers (LY and SH) , while outcome and adjusted analysis information were extracted by a single reviewer"
Rater agreement?	Yes	Not applicable
Blinding, quality assessment?	Yes	Not applicable

Zhong 2010

Methods	Assessed the reporting quality, scientific rigour, and ethics of randomised placebo-controlled trials of TCM compound formulations and compared these differences between Chinese and non-Chinese trials
Data	52 pre-CONSORT RCTs and 227 post CONSORT RCTs published before 1999 and from 2005-2009 respectively
Comparisons	Cross-sectional sample before and after CONSORT publication
Outcomes	Total sum score
Included number of RCTs, Journals	279, not reported
Checklist version used	2001
Field of Study	Traditional Chinese Medicine articles
Notes	Author provided additional data but no journal information

Risk of bias

Item	**Authors' judgement**	**Description**
Large Cohort ?	Yes	Large databases searched over many years
Blinding?	Unclear	Not reported
Confounding by journal quality?	Unclear	Not reported
Outcome Reporting?	Yes	No evidence of selective reporting
Multiple raters?	Yes	Quote: "Two authors assessed each included trial independently."
Rater agreement?	No	Quote: "Interrater reliability was used to test values from each reviewer and Cohen's K was 0.721"
Blinding, quality assessment?	Yes	Not applicable

Ziogas 2009

Methods	Evaluates the reporting quality of published RCTs concerning myeloid haematologic malignancies according to the CONSORT Statement
Data	74 pre-CONSORT RCTs compared with 187 post CONSORT RCTs

Ziogas 2009 (*Continued*)

Comparisons	Title and abstract, background, participants, interventions, objectives, outcomes, sample size, sequence generation, allocation concealment, implementation, blinding any, statistical methods, participant flow, recruitment, baseline data, numbers analysed, outcomes and estimation, ancillary analyses, adverse events, interpretation, generalisability, overall evidence
Outcomes	Cross-sectional sample before and after CONSORT publication
Included number of RCTs, Journals	261, not reported
Checklist version used	1996
Field of Study	Myeloid leukaemia and myelodysplastic syndromes
Notes	

Risk of bias

Item	**Authors' judgement**	**Description**
Large Cohort ?	Yes	Over 2 decades, large number of studies, wide database search
Blinding?	Unclear	Not reported
Confounding by journal quality?	Unclear	Quote: "The RCTs of major IF journals have adhered better to the CONSORT statement. "
Outcome Reporting?	Yes	No evidence of outcome reporting
Multiple raters?	Yes	Quote: "No pilot training of the data extraction was performed."
Rater agreement?	Yes	Not reported
Blinding, quality assessment?	Yes	Not applicable

BMJ: *British Medical Journal*
CI: confidence interval
CONSORT: Consolidated Standards of Reporting Trials
COPD: chronic obstructive pulmonary disease
ECTR: endoscopic carpal tunnel release
ITT: intention-to-treat
ICC: Intra-Class Correlation Coefficient
ICJME: Iternational Commitee of Medical Journal Editors
JAMA: *Journal of the American Medical Association*
ME: Managing Editor

NEJM: *New England Journal of Medicine*
NPT: Non-pharmocological Trials
OCTR: open carpal tunnel release
PRO: Patient reported outcomes
RCT: randomised controlled trial
STRICTA: Standards for Reporting Interventions in Controlled Trials of Acupuncture
TCM: Traditional Chinese Medicine

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
Albavera-Hernández 2009	No information available to form before and after or endorsers versus non-endorsers comparison groups
Berwanger 2009	Adherence limited to abstracts. This study reports data for 4 journals in 2006 all of which are endorsers of CONSORT, which would prevent inclusion based on establishing reporting by comparison group
Chowers 2009	Unable to form comparison group
Ellis 2005	Author provided data to enable our team to determine if BJS could potentially be included in a before and after comparison; given the BJS endorsement was in August 2006 and the study period runs through 2003, no comparison could be formed
Li 2009	Information was not available by journal, hence no comparison could be formed. It should be noted that although this report was e-published in advance in 2009, it was not formally published until 2011
Mills 2005	Author investigated but was unable to find data file, as journal information was categorised by study we were unable to determine endorsement status and form a comparison
Norton-Mabus 2008	Investigated, but unable to form an endorser comparison group
Smith 2008	Partial information reported in tables 1 and 2 of the text; as author was unable to provide additional information we did not have sufficient information to include
Taghinia 2008	No information was obtained from author, hence no comparison group could be established
Xu 2008	Emailed authors in attempt to obtain data by journal; we did not receive a response, so unable to form a comparison group
Yu 2009	Unable to determine comparison group

BJS: British Journal of Surgery

DATA AND ANALYSES

Comparison 1. CONSORT-endorsing journals versus CONSORT non-endorsing journals

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Title and abstract	7	1233	Risk Ratio (IV, Random, 99% CI)	1.13 [0.96, 1.33]
1.1 Studies considering 1996 checklist	1	77	Risk Ratio (IV, Random, 99% CI)	0.93 [0.65, 1.32]
1.2 Studies considering 2001 checklist	6	1156	Risk Ratio (IV, Random, 99% CI)	1.16 [0.97, 1.39]
2 Introduction	5	513	Risk Ratio (IV, Random, 99% CI)	1.07 [1.01, 1.14]
3 Participants	6	683	Risk Ratio (IV, Random, 99% CI)	0.95 [0.56, 1.62]
4 Interventions	6	638	Risk Ratio (IV, Random, 99% CI)	1.00 [0.95, 1.05]
5 Objectives	5	540	Risk Ratio (IV, Random, 99% CI)	1.01 [0.96, 1.06]
6 Outcomes	8	1302	Risk Ratio (IV, Random, 99% CI)	1.17 [0.95, 1.43]
6.1 Studies considering 1996 checklist	1	73	Risk Ratio (IV, Random, 99% CI)	1.02 [0.58, 1.78]
6.2 Studies considering 2001 checklist	7	1229	Risk Ratio (IV, Random, 99% CI)	1.18 [0.94, 1.48]
7 Sample Size	11	1843	Risk Ratio (IV, Random, 99% CI)	1.61 [1.13, 2.29]
7.1 1996 checklist	3	547	Risk Ratio (IV, Random, 99% CI)	1.25 [1.08, 1.46]
7.2 2001 checklist	8	1296	Risk Ratio (IV, Random, 99% CI)	1.81 [1.10, 2.99]
8 Sequence generation	14	2231	Risk Ratio (IV, Random, 99% CI)	1.59 [1.38, 1.84]
9 Allocation concealment	16	2396	Risk Ratio (IV, Random, 99% CI)	1.81 [1.25, 2.61]
10 Implementation	5	498	Risk Ratio (IV, Random, 99% CI)	2.90 [0.54, 15.54]
11 Blinding	13		Risk Ratio (IV, Random, 99% CI)	Subtotals only
11.1 Blinding (participants)	5	711	Risk Ratio (IV, Random, 99% CI)	1.39 [0.87, 2.21]
11.2 Blinding (intervenor)	5	710	Risk Ratio (IV, Random, 99% CI)	1.25 [0.74, 2.12]
11.3 Blinding (outcome assessor)	5	719	Risk Ratio (IV, Random, 99% CI)	1.72 [0.69, 4.31]
11.4 Blinding (data analyst)	3	497	Risk Ratio (IV, Random, 99% CI)	3.56 [0.40, 31.99]
11.5 Blinding (any description)	8	1851	Risk Ratio (IV, Random, 99% CI)	1.23 [0.93, 1.62]
12 Statistical methods	9	894	Risk Ratio (IV, Random, 99% CI)	1.03 [0.90, 1.18]
13 Participant flow	16	2461	Risk Ratio (IV, Random, 99% CI)	1.23 [0.98, 1.53]
13.1 1996 checklist	6	825	Risk Ratio (IV, Random, 99% CI)	1.01 [0.99, 1.02]
13.2 2001 checklist	10	1636	Risk Ratio (IV, Random, 99% CI)	1.35 [1.00, 1.82]
14 Recruitment	6	959	Risk Ratio (IV, Random, 99% CI)	1.03 [0.75, 1.40]
15 Baseline data	5	529	Risk Ratio (IV, Random, 99% CI)	1.07 [0.94, 1.22]
16 Numbers analysed	13	2145	Risk Ratio (IV, Random, 99% CI)	1.23 [0.98, 1.55]
16.1 Studies considering the 1996 checklist	3	665	Risk Ratio (IV, Random, 99% CI)	0.99 [0.83, 1.19]
16.2 Studies considering 2001 checklist	10	1480	Risk Ratio (IV, Random, 99% CI)	1.29 [0.99, 1.68]
17 Outcomes and estimation	6	617	Risk Ratio (IV, Random, 99% CI)	1.00 [0.96, 1.05]
18 Ancillary analyses	4	378	Risk Ratio (IV, Random, 99% CI)	1.31 [0.48, 3.58]
19 Adverse events	8	911	Risk Ratio (IV, Random, 99% CI)	1.14 [0.85, 1.51]
20 Interpretation	5	540	Risk Ratio (IV, Random, 99% CI)	1.01 [0.96, 1.06]

21 Generalisability	5	540	Risk Ratio (IV, Random, 99% CI)	1.22 [0.87, 1.69]
22 Overall evidence	4	317	Risk Ratio (IV, Random, 99% CI)	1.03 [0.91, 1.17]
23 Total sum score	7	560	Std. Mean Difference (IV, Random, 99% CI)	0.68 [0.38, 0.98]

Comparison 2. CONSORT-endorsing journals before and after CONSORT endorsement

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Title and abstract	3	532	Risk Ratio (IV, Random, 99% CI)	1.41 [0.63, 3.16]
1.1 Studies considering 1996 checklist	2	90	Risk Ratio (IV, Random, 99% CI)	1.75 [0.30, 10.17]
1.2 Studies considering 2001 checklist	1	442	Risk Ratio (IV, Random, 99% CI)	1.01 [0.97, 1.05]
2 Introduction	2	457	Risk Ratio (IV, Random, 99% CI)	1.04 [1.00, 1.08]
3 Participants	4	622	Risk Ratio (IV, Random, 99% CI)	0.98 [0.88, 1.09]
4 Interventions	4	630	Risk Ratio (IV, Random, 99% CI)	1.02 [0.97, 1.07]
5 Objectives	2	517	Risk Ratio (IV, Random, 99% CI)	1.04 [0.91, 1.17]
6 Outcomes	5	716	Risk Ratio (IV, Random, 99% CI)	1.68 [0.96, 2.96]
6.1 Studies considering 1996 checklist	2	89	Risk Ratio (IV, Random, 99% CI)	2.23 [0.20, 25.38]
6.2 Studies considering 2001 checklist	3	627	Risk Ratio (IV, Random, 99% CI)	1.61 [0.95, 2.72]
7 Sample size	6	983	Risk Ratio (IV, Random, 99% CI)	1.30 [0.71, 2.36]
7.1 studies considering 1996 checklist	3	356	Risk Ratio (IV, Random, 99% CI)	1.19 [0.62, 2.29]
7.2 Studies considering 2001 checklist	3	627	Risk Ratio (IV, Random, 99% CI)	1.50 [0.44, 5.13]
8 Sequence generation	8	1085	Risk Ratio (IV, Random, 99% CI)	1.46 [0.88, 2.45]
9 Allocation concealment	6	855	Risk Ratio (IV, Random, 99% CI)	1.23 [0.55, 2.74]
10 Implementation	2	517	Risk Ratio (IV, Random, 99% CI)	1.94 [0.15, 24.36]
11 Blinding	5		Risk Ratio (IV, Random, 99% CI)	Subtotals only
11.1 Blinding (participants)	1	75	Risk Ratio (IV, Random, 99% CI)	0.77 [0.45, 1.31]
11.2 Blinding (interventions)	1	75	Risk Ratio (IV, Random, 99% CI)	0.26 [0.09, 0.73]
11.3 Blinding (outcome assessors)	1	75	Risk Ratio (IV, Random, 99% CI)	0.66 [0.34, 1.31]
11.4 Blinding (data analyst)	1	75	Risk Ratio (IV, Random, 99% CI)	0.27 [0.02, 3.78]
11.5 Blinding (any description)	4	926	Risk Ratio (IV, Random, 99% CI)	0.96 [0.61, 1.50]
12 Statistical methods	5	1111	Risk Ratio (IV, Random, 99% CI)	0.86 [0.62, 1.19]
13 Participant flow	8	992	Risk Ratio (IV, Random, 99% CI)	1.33 [0.95, 1.87]
13.1 Studies considering 1996 checklist	6	430	Risk Ratio (IV, Random, 99% CI)	1.44 [0.73, 2.87]
13.2 Studies considering 2001 checklist	2	562	Risk Ratio (IV, Random, 99% CI)	1.30 [1.08, 1.57]
14 Recruitment	3	828	Risk Ratio (IV, Random, 99% CI)	1.77 [0.48, 6.46]
15 Baseline data	2	529	Risk Ratio (IV, Random, 99% CI)	1.42 [1.24, 1.62]
16 Numbers analysed	6	1005	Risk Ratio (IV, Random, 99% CI)	1.72 [1.18, 2.49]

16.1 Studies considering 1996 checklist	3	356	Risk Ratio (IV, Random, 99% CI)	1.50 [0.86, 2.62]
16.2 Studies considering 2001 checklist	3	649	Risk Ratio (IV, Random, 99% CI)	1.95 [1.60, 2.37]
17 Outcomes and estimation	3	532	Risk Ratio (IV, Random, 99% CI)	1.35 [0.73, 2.51]
18 Ancillary analyses	1	442	Risk Ratio (IV, Random, 99% CI)	3.46 [2.47, 4.84]
19 Adverse events	3	507	Risk Ratio (IV, Random, 99% CI)	1.39 [1.12, 1.73]
20 Interpretation	1	442	Risk Ratio (IV, Random, 99% CI)	1.01 [0.99, 1.04]
21 Generalisability	1	442	Risk Ratio (IV, Random, 99% CI)	1.77 [1.47, 2.11]
22 Overall evidence	2	517	Risk Ratio (IV, Random, 99% CI)	1.31 [0.99, 1.73]
23 Total sum score	1	148	Std. Mean Difference (IV, Random, 99% CI)	0.74 [0.30, 1.18]

Comparison 3. Sample of RCTs before and after CONSORT publication

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Title and abstract	7	8225	Risk Ratio (IV, Random, 99% CI)	1.18 [0.98, 1.42]
1.1 Studies considering 1996 checklist	4	602	Risk Ratio (IV, Random, 99% CI)	1.13 [0.96, 1.33]
1.2 Studies considering 2001 checklist	3	7623	Risk Ratio (IV, Random, 99% CI)	1.18 [0.88, 1.59]
2 Introduction	8	8293	Risk Ratio (IV, Random, 99% CI)	1.10 [0.94, 1.30]
3 Participants	6	8368	Risk Ratio (IV, Random, 99% CI)	1.15 [0.99, 1.33]
4 Interventions	7	8224	Risk Ratio (IV, Random, 99% CI)	1.00 [0.97, 1.04]
5 Objectives	5	8028	Risk Ratio (IV, Random, 99% CI)	1.02 [0.97, 1.07]
6 Outcomes	7	9315	Risk Ratio (IV, Random, 99% CI)	1.24 [0.98, 1.58]
6.1 Studies considering 1996 checklist	4	602	Risk Ratio (IV, Random, 99% CI)	1.47 [0.87, 2.48]
6.2 Studies considering 2001 checklist	3	8713	Risk Ratio (IV, Random, 99% CI)	1.15 [0.85, 1.54]
7 Sample size	10	9568	Risk Ratio (IV, Random, 99% CI)	2.45 [1.37, 4.39]
7.1 Studies considering 1996 checklist	5	663	Risk Ratio (IV, Random, 99% CI)	2.49 [0.78, 7.95]
7.2 Studies considering 2001 checklist	5	8905	Risk Ratio (IV, Random, 99% CI)	2.68 [1.00, 7.16]
8 Sequence generation	11	9934	Risk Ratio (IV, Random, 99% CI)	1.67 [1.14, 2.45]
9 Allocation concealment	11	9772	Risk Ratio (IV, Random, 99% CI)	1.61 [1.23, 2.10]
10 Implementation	4	490	Risk Ratio (IV, Random, 99% CI)	1.25 [0.41, 3.79]
11 Blinding	10		Risk Ratio (IV, Random, 99% CI)	Subtotals only
11.1 Blinding (participants)	6	8108	Risk Ratio (IV, Random, 99% CI)	1.21 [0.93, 1.58]
11.2 Blinding (intervenor)	3	586	Risk Ratio (IV, Random, 99% CI)	1.11 [0.88, 1.42]
11.3 Blinding (outcome assessor)	4	600	Risk Ratio (IV, Random, 99% CI)	1.42 [0.99, 2.04]
11.4 Blinding (data analyst)	1	14	Risk Ratio (IV, Random, 99% CI)	1.2 [0.58, 2.50]
11.5 Blinding (any description)	3	1660	Risk Ratio (IV, Random, 99% CI)	0.95 [0.76, 1.19]
12 Statistical methods	7	8223	Risk Ratio (IV, Random, 99% CI)	1.13 [1.01, 1.25]
13 Participant flow	8	8373	Risk Ratio (IV, Random, 99% CI)	1.36 [1.01, 1.83]

13.1 Studies considering 1996 checklist	4	602	Risk Ratio (IV, Random, 99% CI)	1.16 [0.87, 1.53]
13.2 Studies considering 2001 checklist	4	7771	Risk Ratio (IV, Random, 99% CI)	2.14 [0.90, 5.09]
14 Recruitment	5	8024	Risk Ratio (IV, Random, 99% CI)	1.03 [0.89, 1.18]
15 Baseline data	6	8114	Risk Ratio (IV, Random, 99% CI)	1.20 [1.01, 1.43]
16 Numbers analysed	8	1307	Risk Ratio (IV, Random, 99% CI)	1.57 [0.91, 2.70]
16.1 Studies considering 1996 checklist	3	539	Risk Ratio (IV, Random, 99% CI)	2.32 [0.50, 10.87]
16.2 Studies considering 2001 checklist	5	768	Risk Ratio (IV, Random, 99% CI)	1.37 [0.80, 2.36]
17 Outcomes and estimation	9	8613	Risk Ratio (IV, Random, 99% CI)	1.06 [0.98, 1.15]
18 Ancillary analysis	5	8738	Risk Ratio (IV, Random, 99% CI)	1.06 [0.47, 2.39]
19 Adverse events	6	8186	Risk Ratio (IV, Random, 99% CI)	1.06 [0.91, 1.24]
20 Interpretation	4	7989	Risk Ratio (IV, Random, 99% CI)	0.99 [0.98, 1.01]
21 Generalisability	4	8010	Risk Ratio (IV, Random, 99% CI)	1.06 [0.99, 1.15]
22 Overall evidence	4	8010	Risk Ratio (IV, Random, 99% CI)	1.08 [0.97, 1.21]
23 Total sum score	5	528	Std. Mean Difference (IV, Random, 99% CI)	0.51 [-0.28, 1.30]

Analysis 1.1. Comparison 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals, Outcome 1 Title and abstract.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals

Outcome: 1 Title and abstract

Study or subgroup	Endorsers n/N	Non-Endorsers n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
1 Studies considering 1996 checklist					
Kober 2006	7/8	65/69		11.7 %	0.93 [0.65, 1.32]
Subtotal (99% CI)	**8**	**69**		**11.7 %**	**0.93 [0.65, 1.32]**
Total events: 7 (Endorsers), 65 (Non-Endorsers)					
Heterogeneity: not applicable					
Test for overall effect: Z = 0.54 (P = 0.59)					
2 Studies considering 2001 checklist					
Halpern 2004	6/6	91/94		14.9 %	0.96 [0.73, 1.27]
Wei 2009	35/35	179/188		24.8 %	1.04 [0.97, 1.11]
Uetani 2009	10/11	69/87		14.5 %	1.15 [0.86, 1.52]
Ladd 2010	18/19	70/90		18.6 %	1.22 [1.00, 1.49]
Areia 2010	2/2	3/8		1.6 %	2.14 [0.60, 7.59]
Hopewell 2010	113/274	92/342		14.0 %	1.53 [1.14, 2.06]

(*Continued* . . .)

(. . . *Continued*)

Study or subgroup	Endorsers n/N	Non-Endorsers n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
Subtotal (99% CI)	**347**	**809**		**88.3 %**	**1.16 [0.97, 1.39]**
Total events: 184 (Endorsers), 504 (Non-Endorsers)					
Heterogeneity: Tau^2 = 0.02; Chi^2 = 17.28, df = 5 (P = 0.004); I^2 =71%					
Test for overall effect: Z = 2.16 (P = 0.030)					
Total (99% CI)	**355**	**878**		**100.0 %**	**1.13 [0.96, 1.33]**
Total events: 191 (Endorsers), 569 (Non-Endorsers)					
Heterogeneity: Tau^2 = 0.02; Chi^2 = 18.34, df = 6 (P = 0.01); I^2 =67%					
Test for overall effect: Z = 1.95 (P = 0.051)					
Test for subgroup differences: Chi^2 = 2.14, df = 1 (P = 0.14), I^2 =53%					

Analysis 1.2. Comparison 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals, Outcome 2 Introduction.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals

Outcome: 2 Introduction

Study or subgroup	Endorsers n/N	Non-Endorsers n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
Areia 2010	2/2	8/8		0.8 %	1.00 [0.50, 2.01]
Kober 2006	7/8	64/65		3.3 %	0.89 [0.63, 1.26]
Ladd 2010	19/19	79/90		21.0 %	1.12 [0.97, 1.28]
Uetani 2009	9/11	60/87		2.4 %	1.19 [0.79, 1.79]
Wei 2009	35/35	174/188		72.5 %	1.07 [0.99, 1.15]
Total (99% CI)	**75**	**438**		**100.0 %**	**1.07 [1.01, 1.14]**
Total events: 72 (Endorsers), 385 (Non-Endorsers)					
Heterogeneity: Tau^2 = 0.0; Chi^2 = 2.99, df = 4 (P = 0.56); I^2 =0.0%					
Test for overall effect: Z = 2.89 (P = 0.0039)					
Test for subgroup differences: Not applicable					

Analysis 1.3. Comparison 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals, Outcome 3 Participants.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals

Outcome: 3 Participants

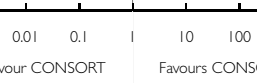
Study or subgroup	Endorsers n/N	Non-Endorsers n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
Faunce 2003	2/2	0/4		2.1 %	8.33 [0.25, 281.33]
Halpern 2004	1/6	7/94		3.8 %	2.24 [0.18, 28.13]
Folkes 2008	42/163	55/83		23.4 %	0.39 [0.26, 0.58]
Uetani 2009	11/11	59/87		25.5 %	1.42 [1.11, 1.81]
Wei 2009	35/35	181/188		26.8 %	1.03 [0.96, 1.09]
Areia 2010	2/2	8/8		18.4 %	1.00 [0.50, 2.01]
Total (99% CI)	**219**	**464**		**100.0 %**	**0.95 [0.56, 1.62]**

Total events: 93 (Endorsers), 310 (Non-Endorsers)

Heterogeneity: $Tau^2 = 0.16$; $Chi^2 = 54.04$, df = 5 ($P<0.00001$); $I^2 = 91\%$

Test for overall effect: Z = 0.24 (P = 0.81)

Test for subgroup differences: Not applicable



Analysis 1.4. Comparison 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals, Outcome 4 Interventions.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals

Outcome: 4 Interventions

Study or subgroup	Endorsers n/N	Non-Endorsers n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
Halpern 2004	5/6	92/94		1.1 %	0.85 [0.53, 1.36]
Kober 2006	6/8	66/67		0.9 %	0.76 [0.45, 1.29]
Wei 2009	35/35	186/188		77.9 %	1.00 [0.95, 1.06]
Ethgen 2009	17/17	108/115		16.5 %	1.04 [0.92, 1.17]
Uetani 2009	10/11	81/87		3.6 %	0.98 [0.76, 1.26]
Areia 2010	1/2	6/8		0.1 %	0.67 [0.10, 4.44]
Total (99% CI)	**79**	**559**		**100.0 %**	**1.00 [0.95, 1.05]**

Total events: 74 (Endorsers), 539 (Non-Endorsers)

Heterogeneity: $Tau^2 = 0.0$; $Chi^2 = 3.59$, df = 5 (P = 0.61); $I^2 = 0.0\%$

Test for overall effect: Z = 0.03 (P = 0.97)

Test for subgroup differences: Not applicable

Analysis 1.5. Comparison 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals, Outcome 5 Objectives.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals

Outcome: 5 Objectives

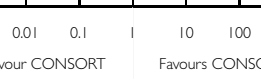
Study or subgroup	Endorsers n/N	Non-Endorsers n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
Halpern 2004	6/6	85/94		2.9 %	1.03 [0.78, 1.37]
Uetani 2009	10/11	82/87		3.6 %	0.96 [0.75, 1.24]
Wei 2009	35/35	188/188		86.3 %	1.00 [0.95, 1.05]
Areia 2010	2/2	8/8		0.5 %	1.00 [0.50, 2.01]
Ladd 2010	18/19	75/90		6.8 %	1.14 [0.95, 1.37]
Total (99% CI)	**73**	**467**		**100.0 %**	**1.01 [0.96, 1.06]**

Total events: 71 (Endorsers), 438 (Non-Endorsers)

Heterogeneity: $Tau^2 = 0.0$; $Chi^2 = 3.21$, df = 4 (P = 0.52); $I^2 = 0.0\%$

Test for overall effect: Z = 0.45 (P = 0.66)

Test for subgroup differences: Not applicable

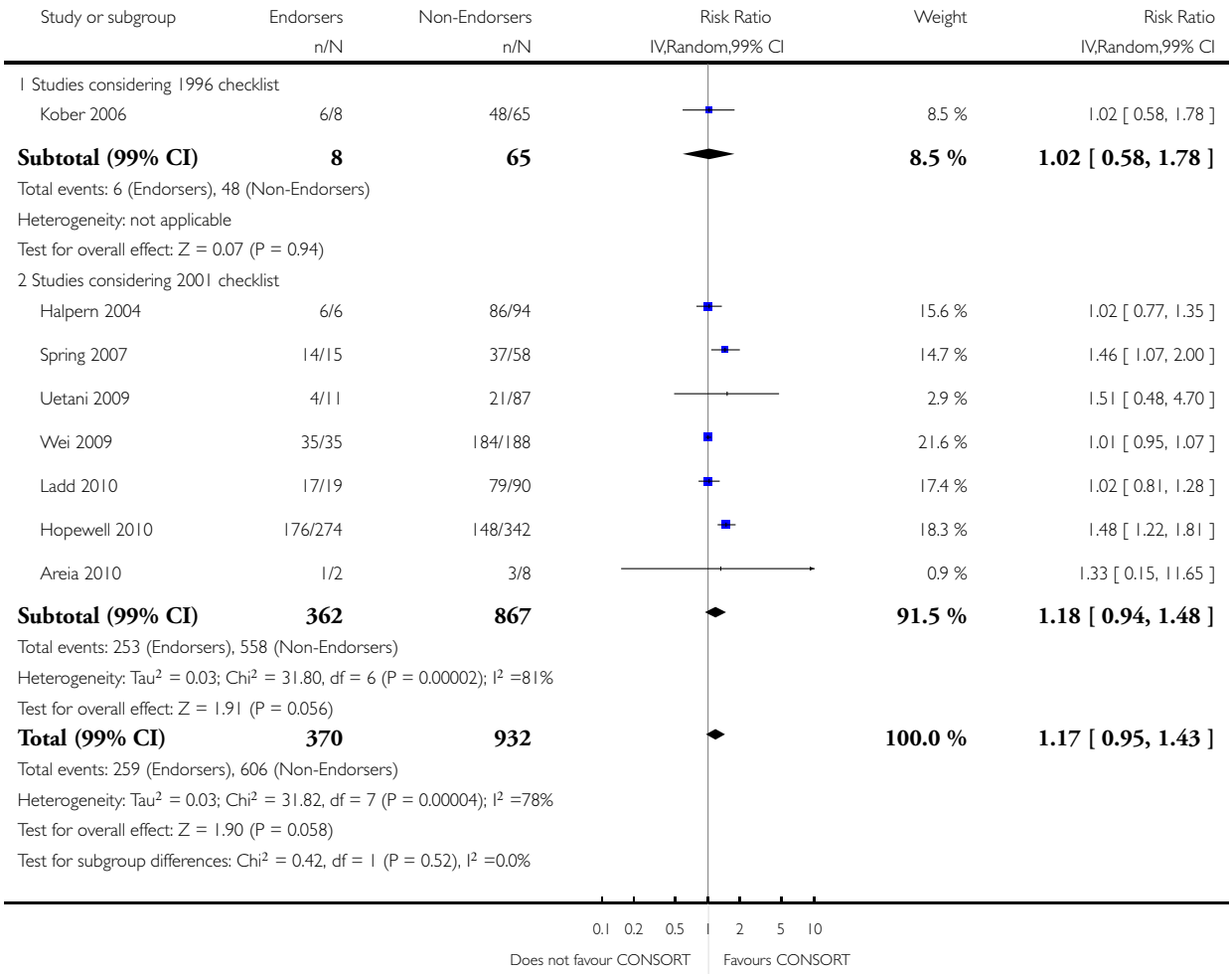


Analysis 1.6. Comparison 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals, Outcome 6 Outcomes.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals

Outcome: 6 Outcomes



Analysis 1.7. Comparison 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals, Outcome 7 Sample Size.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals

Outcome: 7 Sample Size

Study or subgroup	Endorsers n/N	Non-Endorsers n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
1 1996 checklist					
Faunce 2003	0/2	0/4			Not estimable
Kober 2006	2/8	17/65		3.7 %	0.96 [0.18, 5.06]
Kane 2007	141/178	183/290		18.6 %	1.26 [1.08, 1.46]
Subtotal (99% CI)	**188**	**359**		**22.2 %**	**1.25 [1.08, 1.46]**
Total events: 143 (Endorsers), 200 (Non-Endorsers)					
Heterogeneity: Tau^2 = 0.0; Chi^2 = 0.18, df = 1 (P = 0.67); I^2 =0.0%					
Test for overall effect: Z = 3.83 (P = 0.00013)					
2 2001 checklist					
Halpern 2004	5/6	55/94		13.6 %	1.42 [0.85, 2.40]
Lai 2007	26/51	9/16		11.4 %	0.91 [0.46, 1.77]
Spring 2007	12/15	12/58		10.4 %	3.87 [1.84, 8.11]
Wei 2009	11/35	1/188		1.6 %	59.09 [4.18, 834.84]
Uetani 2009	5/11	18/87		7.5 %	2.20 [0.80, 6.02]
Areia 2010	2/2	6/8		9.1 %	1.15 [0.49, 2.70]
Hopewell 2010	158/274	121/342		17.8 %	1.63 [1.29, 2.05]
Ladd 2010	5/19	15/90		6.3 %	1.58 [0.49, 5.04]
Subtotal (99% CI)	**413**	**883**		**77.8 %**	**1.81 [1.10, 2.99]**
Total events: 224 (Endorsers), 237 (Non-Endorsers)					
Heterogeneity: Tau^2 = 0.19; Chi^2 = 28.44, df = 7 (P = 0.00018); I^2 =75%					
Test for overall effect: Z = 3.07 (P = 0.0022)					
Total (99% CI)	**601**	**1242**		**100.0 %**	**1.61 [1.13, 2.29]**
Total events: 367 (Endorsers), 437 (Non-Endorsers)					
Heterogeneity: Tau^2 = 0.10; Chi^2 = 36.90, df = 9 (P = 0.00003); I^2 =76%					
Test for overall effect: Z = 3.44 (P = 0.00058)					
Test for subgroup differences: Chi^2 = 3.32, df = 1 (P = 0.07), I^2 =70%					

Analysis 1.8. Comparison 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals, Outcome 8 Sequence generation.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals

Outcome: 8 Sequence generation

Study or subgroup	Endorsers n/N	Non-Endorsers n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
Devereaux 2002	38/49	22/49		7.7 %	1.73 [1.10, 2.72]
Hill 2002	2/8	23/113		0.7 %	1.23 [0.24, 6.39]
Llorca 2004	14/37	6/23		1.7 %	1.45 [0.51, 4.16]
Halpern 2004	5/6	53/94		6.1 %	1.48 [0.87, 2.50]
Dias 2006	14/19	24/41		6.8 %	1.26 [0.77, 2.05]
Kober 2006	0/8	13/61		0.2 %	0.26 [0.01, 9.28]
Lai 2007	27/51	6/16		2.3 %	1.41 [0.58, 3.47]
Kane 2007	126/178	148/290		21.1 %	1.39 [1.14, 1.68]
Ethgen 2009	16/17	61/115		15.1 %	1.77 [1.35, 2.34]
Uetani 2009	6/11	32/87		2.9 %	1.48 [0.67, 3.29]
Wei 2009	31/35	77/188		15.1 %	2.16 [1.64, 2.85]
Ladd 2010	12/19	39/90		5.7 %	1.46 [0.84, 2.52]
Areia 2010	1/2	4/8		0.5 %	1.00 [0.13, 7.66]
Hopewell 2010	117/274	92/342		14.1 %	1.59 [1.19, 2.13]
Total (99% CI)	**714**	**1517**		**100.0 %**	**1.59 [1.38, 1.84]**

Total events: 409 (Endorsers), 600 (Non-Endorsers)

Heterogeneity: $Tau^2 = 0.01$; $Chi^2 = 17.18$, df = 13 (P = 0.19); $I^2 = 24\%$

Test for overall effect: Z = 8.41 (P < 0.00001)

Test for subgroup differences: Not applicable

Analysis 1.9. Comparison 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals, Outcome 9 Allocation concealment.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals

Outcome: 9 Allocation concealment

Study or subgroup	Endorsers n/N	Non-Endorsers n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
Moher 2001	47/77	14/37		7.8 %	1.61 [0.89, 2.91]
Devereaux 2002	28/49	17/49		7.8 %	1.65 [0.91, 2.99]
Hill 2002	4/8	19/113		5.4 %	2.97 [1.03, 8.57]
Halpern 2004	4/6	51/94		6.8 %	1.23 [0.56, 2.69]
Llorca 2004	5/37	11/23		4.8 %	0.28 [0.08, 0.95]
Greenfield 2005	5/98	7/182		3.8 %	1.33 [0.30, 5.79]
Hewitt 2005	138/166	35/68		9.1 %	1.62 [1.18, 2.22]
Kober 2006	7/7	12/67		7.2 %	5.10 [2.54, 10.26]
Dias 2006	9/19	13/41		6.4 %	1.49 [0.63, 3.52]
Lai 2007	17/51	7/16		6.2 %	0.76 [0.31, 1.86]
Ethgen 2009	12/17	34/115		8.0 %	2.39 [1.38, 4.13]
Wei 2009	13/35	8/188		5.4 %	8.73 [3.04, 25.09]
Uetani 2009	4/11	13/87		4.7 %	2.43 [0.72, 8.25]
Ladd 2010	9/19	19/90		6.6 %	2.24 [0.99, 5.07]
Areia 2010	0/2	4/8		1.0 %	0.33 [0.01, 10.34]
Hopewell 2010	91/274	65/342		8.9 %	1.75 [1.22, 2.51]
Total (99% CI)	**876**	**1520**		**100.0 %**	**1.81 [1.25, 2.61]**

Total events: 393 (Endorsers), 329 (Non-Endorsers)

Heterogeneity: $Tau^2 = 0.21$; $Chi^2 = 60.42$, df = 15 (P<0.00001); $I^2 = 75\%$

Test for overall effect: Z = 4.14 (P = 0.000035)

Test for subgroup differences: Not applicable

Analysis 1.10. Comparison 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals, Outcome 10 Implementation.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals

Outcome: 10 Implementation

Study or subgroup	Endorsers n/N	Non-Endorsers n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
Halpern 2004	1/6	8/94		21.1 %	1.96 [0.16, 24.04]
Lai 2007	19/51	5/16		34.4 %	1.19 [0.41, 3.45]
Wei 2009	11/35	0/188		13.6 %	120.75 [3.01, 4842.78]
Uetani 2009	3/11	12/87		30.8 %	1.98 [0.47, 8.38]
Areia 2010	0/2	0/8			Not estimable
Total (99% CI)	**105**	**393**		**100.0 %**	**2.90 [0.54, 15.54]**

Total events: 34 (Endorsers), 25 (Non-Endorsers)

Heterogeneity: $Tau^2 = 1.06$; $Chi^2 = 9.64$, df = 3 (P = 0.02); $I^2 = 69\%$

Test for overall effect: Z = 1.64 (P = 0.10)

Test for subgroup differences: Not applicable

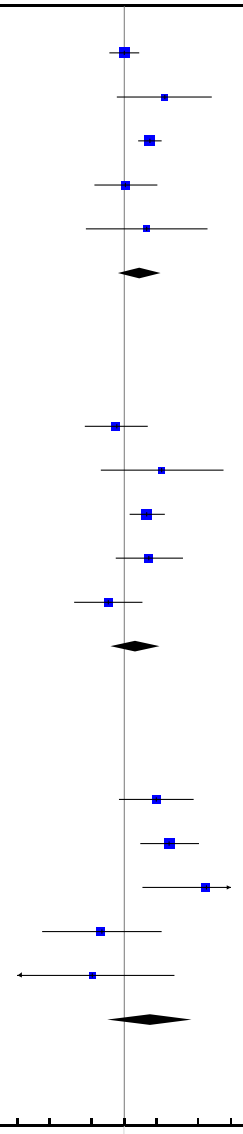
Analysis 1.11. Comparison 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals, Outcome 11 Blinding.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

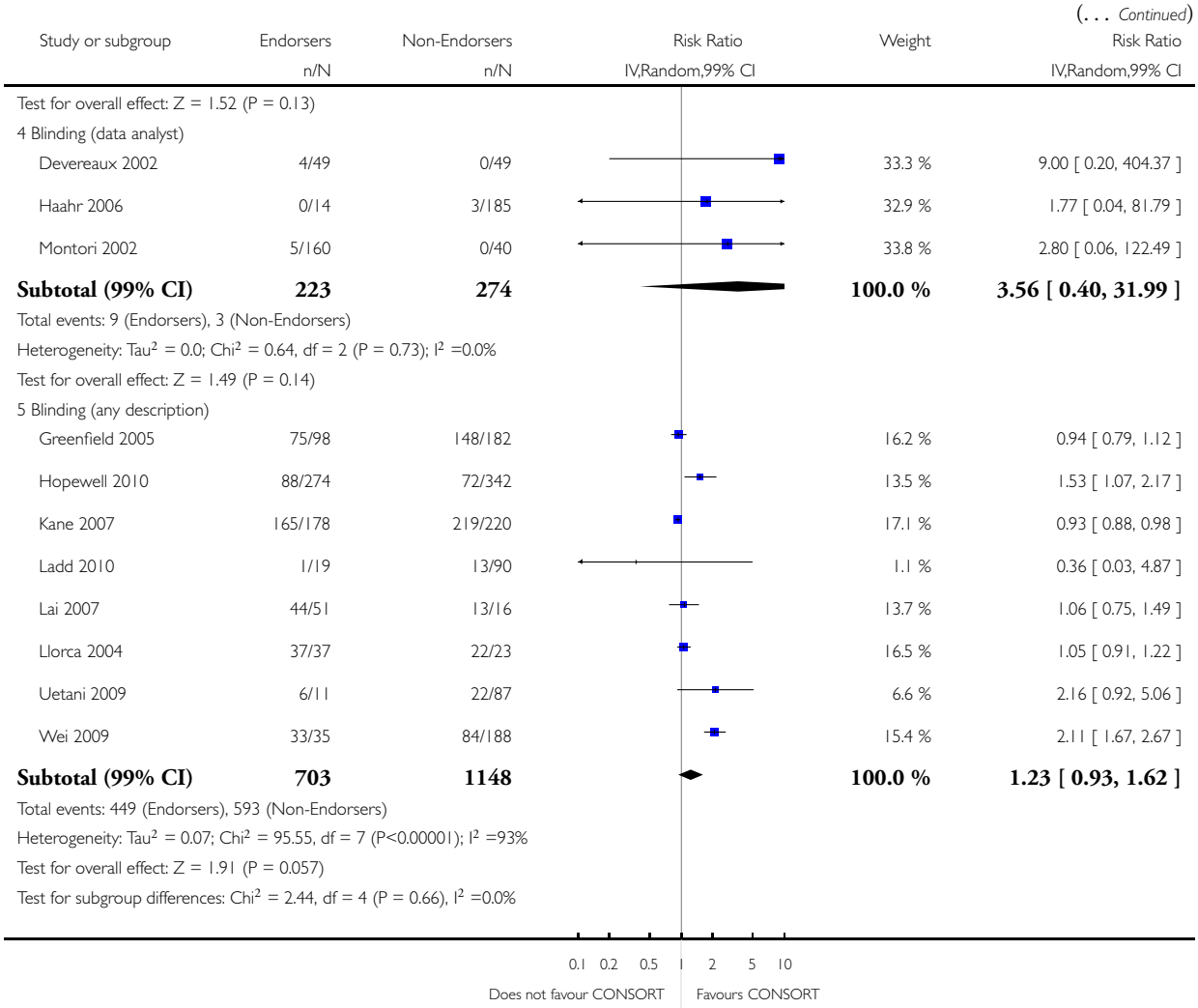
Comparison: 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals

Outcome: 11 Blinding

Study or subgroup	Endorsers n/N	Non-Endorsers n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
1 Blinding (participants)					
Devereaux 2002	36/49	36/49		28.7 %	1.00 [0.73, 1.37]
Ethgen 2009	6/17	17/115		12.8 %	2.39 [0.86, 6.65]
Haahr 2006	13/14	99/185		29.9 %	1.74 [1.34, 2.25]
Halpern 2004	2/2	65/80		19.4 %	1.03 [0.52, 2.03]
Montori 2002	26/160	4/40		9.2 %	1.63 [0.44, 6.00]
Subtotal (99% CI)	**242**	**469**		**100.0 %**	**1.39 [0.87, 2.21]**
Total events: 83 (Endorsers), 221 (Non-Endorsers)					
Heterogeneity: $Tau^2 = 0.10$; $Chi^2 = 15.42$, df = 4 (P = 0.004); $I^2 = 74\%$					
Test for overall effect: Z = 1.80 (P = 0.072)					
2 Blinding (intervenor)					
Devereaux 2002	17/49	20/49		21.4 %	0.85 [0.43, 1.66]
Ethgen 2009	4/17	12/115		10.6 %	2.25 [0.60, 8.51]
Haahr 2006	12/15	91/185		27.9 %	1.63 [1.11, 2.39]
Halpern 2004	2/2	38/78		20.1 %	1.71 [0.83, 3.54]
Montori 2002	34/160	12/40		20.0 %	0.71 [0.34, 1.48]
Subtotal (99% CI)	**243**	**467**		**100.0 %**	**1.25 [0.74, 2.12]**
Total events: 69 (Endorsers), 173 (Non-Endorsers)					
Heterogeneity: $Tau^2 = 0.13$; $Chi^2 = 11.47$, df = 4 (P = 0.02); $I^2 = 65\%$					
Test for overall effect: Z = 1.11 (P = 0.27)					
3 Blinding (outcome assessor)					
Devereaux 2002	22/49	11/49		24.0 %	2.00 [0.90, 4.43]
Ethgen 2009	11/17	28/115		25.8 %	2.66 [1.42, 4.97]
Haahr 2006	4/14	9/185		17.7 %	5.87 [1.49, 23.19]
Halpern 2004	2/4	70/86		18.5 %	0.61 [0.17, 2.24]
Montori 2002	6/160	3/40		14.1 %	0.50 [0.09, 2.92]
Subtotal (99% CI)	**244**	**475**		**100.0 %**	**1.72 [0.69, 4.31]**
Total events: 45 (Endorsers), 121 (Non-Endorsers)					
Heterogeneity: $Tau^2 = 0.43$; $Chi^2 = 15.04$, df = 4 (P = 0.005); $I^2 = 73\%$					



(*Continued . . .*)


Study or subgroup
Endorsers n/N
Non-Endorsers n/N
Risk Ratio IV,Random,99% CI
Weight
Risk Ratio IV,Random,99% CI
Test for overall effect: Z = 1.52 (P = 0.13)
4 Blinding (data analyst)
Devereaux 2002 4/49 0/49 33.3 % 9.00 [0.20, 404.37]
Haahr 2006 0/14 3/185 32.9 % 1.77 [0.04, 81.79]
Montori 2002 5/160 0/40 33.8 % 2.80 [0.06, 122.49]
Subtotal (99% CI) 223 274 100.0 % 3.56 [0.40, 31.99]
Total events: 9 (Endorsers), 3 (Non-Endorsers)
Heterogeneity: Tau² = 0.0; Chi² = 0.64, df = 2 (P = 0.73); I² =0.0%
Test for overall effect: Z = 1.49 (P = 0.14)
5 Blinding (any description)
Greenfield 2005 75/98 148/182 16.2 % 0.94 [0.79, 1.12]
Hopewell 2010 88/274 72/342 13.5 % 1.53 [1.07, 2.17]
Kane 2007 165/178 219/220 17.1 % 0.93 [0.88, 0.98]
Ladd 2010 1/19 13/90 1.1 % 0.36 [0.03, 4.87]
Lai 2007 44/51 13/16 13.7 % 1.06 [0.75, 1.49]
Llorca 2004 37/37 22/23 16.5 % 1.05 [0.91, 1.22]
Uetani 2009 6/11 22/87 6.6 % 2.16 [0.92, 5.06]
Wei 2009 33/35 84/188 15.4 % 2.11 [1.67, 2.67]
Subtotal (99% CI) 703 1148 100.0 % 1.23 [0.93, 1.62]
Total events: 449 (Endorsers), 593 (Non-Endorsers)
Heterogeneity: Tau² = 0.07; Chi² = 95.55, df = 7 (P<0.00001); I² =93%
Test for overall effect: Z = 1.91 (P = 0.057)
Test for subgroup differences: Chi² = 2.44, df = 4 (P = 0.66), I² =0.0%
0.1 0.2 0.5 1 2 5 10
Does not favour CONSORT
Favours CONSORT

Analysis 1.12. Comparison 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals, Outcome 12 Statistical methods.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals

Outcome: 12 Statistical methods

Study or subgroup	Endorsers n/N	Non-Endorsers n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
Hill 2002	6/8	36/113		3.8 %	2.35 [1.25, 4.44]
Devereaux 2002	31/49	32/49		8.2 %	0.97 [0.66, 1.43]
Llorca 2004	27/37	19/23		9.2 %	0.88 [0.62, 1.26]
Halpern 2004	6/6	91/94		12.4 %	0.96 [0.73, 1.27]
Kober 2006	5/8	64/67		3.2 %	0.65 [0.32, 1.33]
Wei 2009	35/35	185/188		24.2 %	1.00 [0.95, 1.06]
Uetani 2009	11/11	87/87		18.9 %	1.00 [0.86, 1.17]
Ladd 2010	19/19	79/90		20.0 %	1.12 [0.97, 1.28]
Areia 2010	2/2	0/8		0.1 %	15.00 [0.41, 550.11]
Total (99% CI)	**175**	**719**		**100.0 %**	**1.03 [0.90, 1.18]**

Total events: 142 (Endorsers), 593 (Non-Endorsers)

Heterogeneity: $Tau^2 = 0.01$; $Chi^2 = 22.67$, df = 8 (P = 0.004); $I^2 = 65\%$

Test for overall effect: Z = 0.53 (P = 0.59)

Test for subgroup differences: Not applicable

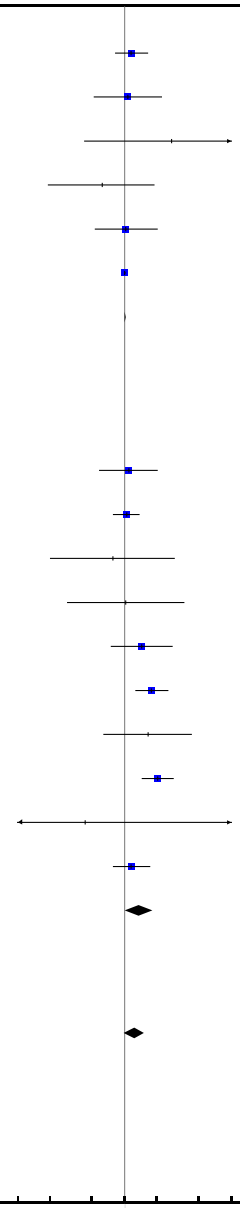
Analysis 1.13. Comparison 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals, Outcome 13 Participant flow.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals

Outcome: 13 Participant flow

Study or subgroup	Endorsers n/N	Non-Endorsers n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
1 1996 checklist					
Devereaux 2002	36/49	31/49		9.9 %	1.16 [0.81, 1.66]
Hill 2002	5/8	66/113		5.4 %	1.07 [0.51, 2.23]
Faunce 2003	2/2	1/4		1.3 %	2.78 [0.42, 18.22]
Kober 2006	3/7	46/65		2.9 %	0.61 [0.19, 1.90]
Dias 2006	10/19	21/41		5.9 %	1.03 [0.52, 2.04]
Kane 2007	178/178	288/290		13.3 %	1.01 [0.99, 1.02]
Subtotal (99% CI)	**263**	**562**		**38.7 %**	**1.01 [0.99, 1.02]**
Total events: 234 (Endorsers), 453 (Non-Endorsers)					
Heterogeneity: $Tau^2 = 0.0$; $Chi^2 = 4.37$, df = 5 (P = 0.50); $I^2 = 0.0\%$					
Test for overall effect: Z = 0.92 (P = 0.36)					
2 2001 checklist					
Llorca 2004	21/37	12/23		6.4 %	1.09 [0.58, 2.05]
Halpern 2004	6/6	84/94		10.9 %	1.04 [0.78, 1.39]
Greenfield 2005	5/98	12/182		2.3 %	0.77 [0.20, 2.93]
Lai 2007	13/51	4/16		2.5 %	1.02 [0.29, 3.65]
Spring 2007	9/15	24/58		6.0 %	1.45 [0.74, 2.85]
Wei 2009	25/35	75/188		9.9 %	1.79 [1.25, 2.56]
Uetani 2009	5/11	24/87		3.8 %	1.65 [0.63, 4.31]
Hopewell 2010	107/274	65/342		10.0 %	2.05 [1.45, 2.91]
Areia 2010	0/2	3/8		0.4 %	0.43 [0.01, 14.12]
Ladd 2010	14/19	57/90		9.2 %	1.16 [0.77, 1.75]
Subtotal (99% CI)	**548**	**1088**		**61.3 %**	**1.35 [1.00, 1.82]**
Total events: 205 (Endorsers), 360 (Non-Endorsers)					
Heterogeneity: $Tau^2 = 0.06$; $Chi^2 = 23.27$, df = 9 (P = 0.01); $I^2 = 61\%$					
Test for overall effect: Z = 2.61 (P = 0.0090)					
Total (99% CI)	**811**	**1650**		**100.0 %**	**1.23 [0.98, 1.53]**
Total events: 439 (Endorsers), 813 (Non-Endorsers)					
Heterogeneity: $Tau^2 = 0.06$; $Chi^2 = 54.57$, df = 15 (P<0.00001); $I^2 = 73\%$					
Test for overall effect: Z = 2.38 (P = 0.017)					
Test for subgroup differences: $Chi^2 = 6.52$, df = 1 (P = 0.01), $I^2 = 85\%$					



Analysis 1.14. Comparison 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals, Outcome 14 Recruitment.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals

Outcome: 14 Recruitment

Study or subgroup	Endorsers n/N	Non-Endorsers n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
Halpern 2004	3/6	12/94		5.4 %	3.92 [1.11, 13.81]
Llorca 2004	3/37	2/23		1.9 %	0.93 [0.10, 8.84]
Kane 2007	147/178	285/290		44.5 %	0.84 [0.77, 0.92]
Uetani 2009	2/11	6/87		2.5 %	2.64 [0.38, 18.25]
Wei 2009	33/35	171/188		43.2 %	1.04 [0.92, 1.17]
Areia 2010	1/2	6/8		2.6 %	0.67 [0.10, 4.44]
Total (99% CI)	**269**	**690**		**100.0 %**	**1.03 [0.75, 1.40]**

Total events: 189 (Endorsers), 482 (Non-Endorsers)

Heterogeneity: $Tau^2 = 0.03$; $Chi^2 = 23.68$, df = 5 (P = 0.00025); $I^2 = 79\%$

Test for overall effect: Z = 0.20 (P = 0.84)

Test for subgroup differences: Not applicable

Analysis 1.15. Comparison 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals, Outcome 15 Baseline data.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals

Outcome: 15 Baseline data

Study or subgroup	Endorsers n/N	Non-Endorsers n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
Devereaux 2002	45/49	46/49		28.1 %	0.98 [0.85, 1.13]
Halpern 2004	6/6	86/94		14.0 %	1.02 [0.77, 1.35]
Wei 2009	34/35	153/188		32.1 %	1.19 [1.06, 1.34]
Uetani 2009	11/11	78/87		23.2 %	1.07 [0.90, 1.29]
Areia 2010	2/2	7/8		2.6 %	1.00 [0.46, 2.16]
Total (99% CI)	**103**	**426**		**100.0 %**	**1.07 [0.94, 1.22]**

Total events: 98 (Endorsers), 370 (Non-Endorsers)

Heterogeneity: $Tau^2 = 0.01$; $Chi^2 = 8.22$, df = 4 (P = 0.08); $I^2 = 51\%$

Test for overall effect: Z = 1.41 (P = 0.16)

Test for subgroup differences: Not applicable

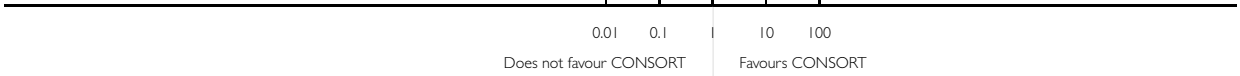
Analysis 1.16. Comparison 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals, Outcome 16 Numbers analysed.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals

Outcome: 16 Numbers analysed

Study or subgroup	Endorsers n/N	Non-Endorsers n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
1 Studies considering the 1996 checklist					
Hill 2002	2/8	34/113		1.8 %	0.83 [0.16, 4.20]
Kober 2006	1/8	19/68		0.8 %	0.45 [0.04, 5.24]
Kane 2007	116/178	189/290		13.5 %	1.00 [0.84, 1.20]
Subtotal (99% CI)	**194**	**471**		**16.1 %**	**0.99 [0.83, 1.19]**
Total events: 119 (Endorsers), 242 (Non-Endorsers)					
Heterogeneity: $Tau^2 = 0.0$; $Chi^2 = 0.79$, df = 2 (P = 0.67); $I^2 = 0.0\%$					
Test for overall effect: Z = 0.09 (P = 0.93)					
2 Studies considering 2001 checklist					
Llorca 2004	2/37	4/23		1.1 %	0.31 [0.04, 2.60]
Halpern 2004	5/6	79/94		8.9 %	0.99 [0.61, 1.61]
Spring 2007	11/15	22/50		7.7 %	1.67 [0.94, 2.96]
Lai 2007	37/51	11/16		8.9 %	1.06 [0.65, 1.72]
Ethgen 2009	15/17	90/115		12.4 %	1.13 [0.87, 1.46]
Uetani 2009	6/11	46/87		5.7 %	1.03 [0.48, 2.20]
Wei 2009	22/35	35/188		8.5 %	3.38 [2.02, 5.66]
Areia 2010	2/2	8/8		6.3 %	1.00 [0.50, 2.01]
Ladd 2010	15/19	53/90		10.5 %	1.34 [0.92, 1.96]
Hopewell 2010	215/274	207/342		14.0 %	1.30 [1.13, 1.49]
Subtotal (99% CI)	**467**	**1013**		**83.9 %**	**1.29 [0.99, 1.68]**
Total events: 330 (Endorsers), 555 (Non-Endorsers)					
Heterogeneity: $Tau^2 = 0.06$; $Chi^2 = 33.76$, df = 9 (P = 0.00010); $I^2 = 73\%$					
Test for overall effect: Z = 2.53 (P = 0.012)					
Total (99% CI)	**661**	**1484**		**100.0 %**	**1.23 [0.98, 1.55]**
Total events: 449 (Endorsers), 797 (Non-Endorsers)					
Heterogeneity: $Tau^2 = 0.05$; $Chi^2 = 44.99$, df = 12 (P = 0.00001); $I^2 = 73\%$					
Test for overall effect: Z = 2.33 (P = 0.020)					
Test for subgroup differences: $Chi^2 = 4.60$, df = 1 (P = 0.03), $I^2 = 78\%$					



Analysis 1.17. Comparison 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals, Outcome 17 Outcomes and estimation.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals

Outcome: 17 Outcomes and estimation

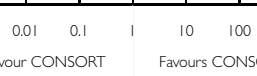
Study or subgroup	Endorsers n/N	Non-Endorsers n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
Halpern 2004	6/6	94/94		3.3 %	1.00 [0.76, 1.31]
Kober 2006	8/8	69/69		5.5 %	1.00 [0.81, 1.23]
Uetani 2009	2/11	6/87		0.1 %	2.64 [0.38, 18.25]
Wei 2009	35/35	188/188		89.7 %	1.00 [0.95, 1.05]
Ladd 2010	13/19	48/90		1.1 %	1.28 [0.80, 2.06]
Areia 2010	2/2	7/8		0.4 %	1.00 [0.46, 2.16]
Total (99% CI)	**81**	**536**		**100.0 %**	**1.00 [0.96, 1.05]**

Total events: 66 (Endorsers), 412 (Non-Endorsers)

Heterogeneity: $Tau^2 = 0.0$; $Chi^2 = 3.46$, df = 5 (P = 0.63); $I^2 = 0.0\%$

Test for overall effect: Z = 0.17 (P = 0.86)

Test for subgroup differences: Not applicable



Analysis 1.18. Comparison 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals, Outcome 18 Ancillary analyses.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals

Outcome: 18 Ancillary analyses

Study or subgroup	Endorsers n/N	Non-Endorsers n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
Uetani 2009	0/11	3/87		6.1 %	1.05 [0.02, 47.43]
Ladd 2010	1/19	12/90		11.5 %	0.39 [0.03, 5.32]
Yu 2010	65/113	11/48		41.0 %	2.51 [1.23, 5.12]
Areia 2010	2/2	8/8		41.4 %	1.00 [0.50, 2.01]
Total (99% CI)	**145**	**233**		**100.0 %**	**1.31 [0.48, 3.58]**

Total events: 68 (Endorsers), 34 (Non-Endorsers)

Heterogeneity: $Tau^2 = 0.29$; $Chi^2 = 7.51$, df = 3 (P = 0.06); $I^2 = 60\%$

Test for overall effect: Z = 0.70 (P = 0.48)

Test for subgroup differences: Not applicable

Analysis 1.19. Comparison 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals, Outcome 19 Adverse events.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals

Outcome: 19 Adverse events

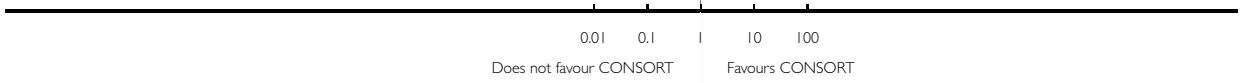
Study or subgroup	Endorsers n/N	Non-Endorsers n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
Halpern 2004	6/6	89/94		18.9 %	0.99 [0.75, 1.30]
Greenfield 2005	68/98	120/182		20.2 %	1.05 [0.84, 1.31]
Dias 2006	3/19	6/41		2.6 %	1.08 [0.20, 5.76]
Kober 2006	6/8	54/65		12.5 %	0.90 [0.52, 1.56]
Lai 2007	39/51	10/16		12.7 %	1.22 [0.71, 2.09]
Wei 2009	30/35	93/188		19.3 %	1.73 [1.34, 2.25]
Uetani 2009	7/11	46/87		10.6 %	1.20 [0.63, 2.29]
Areia 2010	1/2	8/8		3.1 %	0.53 [0.12, 2.38]
Total (99% CI)	**230**	**681**		**100.0 %**	**1.14 [0.85, 1.51]**

Total events: 160 (Endorsers), 426 (Non-Endorsers)

Heterogeneity: $Tau^2 = 0.05$; $Chi^2 = 22.59$, df = 7 (P = 0.002); $I^2 = 69\%$

Test for overall effect: Z = 1.16 (P = 0.25)

Test for subgroup differences: Not applicable



Analysis 1.20. Comparison 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals, Outcome 20 Interpretation.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals

Outcome: 20 Interpretation

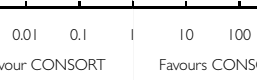
Study or subgroup	Endorsers n/N	Non-Endorsers n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
Halpern 2004	4/6	67/94		0.4 %	0.94 [0.44, 2.00]
Wei 2009	35/35	188/188		85.9 %	1.00 [0.95, 1.05]
Uetani 2009	6/11	45/87		0.4 %	1.05 [0.49, 2.25]
Areia 2010	2/2	8/8		0.5 %	1.00 [0.50, 2.01]
Ladd 2010	19/19	80/90		12.8 %	1.10 [0.96, 1.26]
Total (99% CI)	**73**	**467**		**100.0 %**	**1.01 [0.96, 1.06]**

Total events: 66 (Endorsers), 388 (Non-Endorsers)

Heterogeneity: $Tau^2 = 0.0$; $Chi^2 = 3.13$, df = 4 (P = 0.54); $I^2 = 0.0\%$

Test for overall effect: Z = 0.67 (P = 0.51)

Test for subgroup differences: Not applicable



Analysis 1.21. Comparison 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals, Outcome 21 Generalisability.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals

Outcome: 21 Generalisability

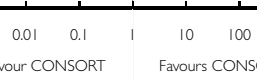
Study or subgroup	Endorsers n/N	Non-Endorsers n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
Halpern 2004	4/6	24/94		10.2 %	2.61 [1.09, 6.24]
Wei 2009	26/35	149/188		27.9 %	0.94 [0.71, 1.23]
Uetani 2009	9/11	51/87		21.7 %	1.40 [0.90, 2.15]
Ladd 2010	16/19	64/90		26.6 %	1.18 [0.87, 1.61]
Areia 2010	2/2	8/8		13.6 %	1.00 [0.50, 2.01]
Total (99% CI)	**73**	**467**		**100.0 %**	**1.22 [0.87, 1.69]**

Total events: 57 (Endorsers), 296 (Non-Endorsers)

Heterogeneity: $Tau^2 = 0.05$; $Chi^2 = 11.16$, df = 4 (P = 0.02); $I^2 = 64\%$

Test for overall effect: Z = 1.53 (P = 0.13)

Test for subgroup differences: Not applicable



Analysis 1.22. Comparison 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals, Outcome 22 Overall evidence.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals

Outcome: 22 Overall evidence

Study or subgroup	Endorsers n/N	Non-Endorsers n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
Halpern 2004	6/6	94/94		21.5 %	1.00 [0.76, 1.31]
Uetani 2009	10/11	86/87		25.8 %	0.92 [0.72, 1.18]
Ladd 2010	18/19	77/90		49.5 %	1.11 [0.93, 1.32]
Areia 2010	2/2	8/8		3.2 %	1.00 [0.50, 2.01]
Total (99% CI)	**38**	**279**		**100.0 %**	**1.03 [0.91, 1.17]**

Total events: 36 (Endorsers), 265 (Non-Endorsers)

Heterogeneity: $Tau^2 = 0.0$; $Chi^2 = 2.57$, df = 3 (P = 0.46); $I^2 = 0.0\%$

Test for overall effect: Z = 0.59 (P = 0.55)

Test for subgroup differences: Not applicable

Analysis 1.23. Comparison 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals, Outcome 23 Total sum score.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 1 CONSORT-endorsing journals versus CONSORT non-endorsing journals

Outcome: 23 Total sum score

Study or subgroup	Endorsers N	Mean(SD)	Non-Endorsers N	Mean(SD)	Std. Mean Difference IV,Random,99% CI	Weight	Std. Mean Difference IV,Random,99% CI
Moher 2001 (1)	77	27.1 (4.81)	37	22.8 (4.68)		30.6 %	0.90 [0.36, 1.43]
Kidwell 2001 (2)	9	88.3 (13.45)	25	71.2 (20.82)		8.2 %	0.87 [-0.17, 1.91]
Tiruvoipati 2005	2	74.9 (2.37)	62	65.75 (7.67)		2.5 %	1.19 [-0.68, 3.06]
Balasubramanian 2006 (3)	11	77.28 (5.87)	58	68.88 (10.02)		11.7 %	0.87 [0.00, 1.74]
Agha 2007	8	12.25 (1.49)	82	11.03 (1.76)		9.6 %	0.70 [-0.27, 1.66]
Pat 2008 (4)	4	19.25 (0.96)	34	16.94 (2.71)		4.6 %	0.87 [-0.52, 2.25]
Tharyan 2008 (5)	31	5.55 (2.51)	120	4.93 (2.04)		32.7 %	0.29 [-0.23, 0.81]
Total (99% CI)	**142**		**418**			**100.0 %**	**0.68 [0.38, 0.98]**

Heterogeneity: $Tau^2 = 0.0$; $Chi^2 = 5.98$, df = 6 (P = 0.43); $I^2 = 0.0\%$

Test for overall effect: Z = 5.87 (P < 0.00001)

Test for subgroup differences: Not applicable

(1) 40 point score

(2) Scale is out of 100. SD imputed based on other included studies.

(3) Score is out of 90. Medians reported.

(4) Score out of 22 items

(5) Tharyan based on 13 items of checklist

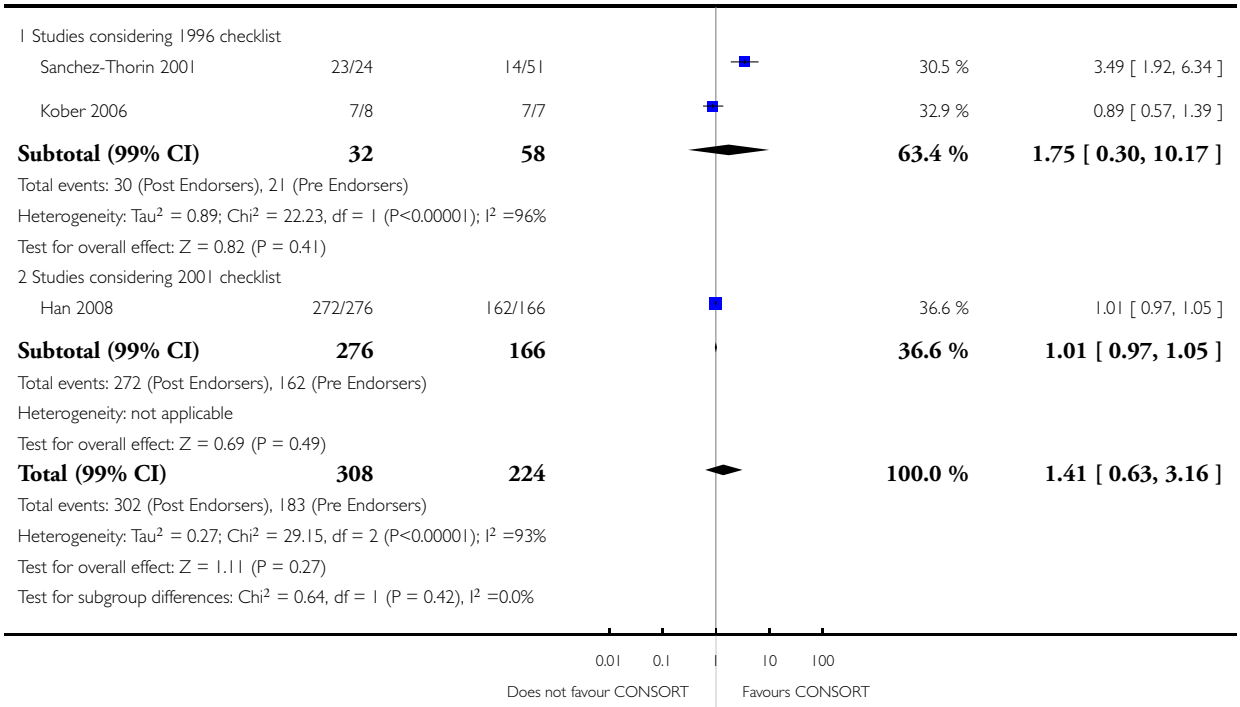
Analysis 2.1. Comparison 2 CONSORT-endorsing journals before and after CONSORT endorsement, Outcome 1 Title and abstract.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 2 CONSORT-endorsing journals before and after CONSORT endorsement

Outcome: 1 Title and abstract

Study or subgroup	Post Endorsers n/N	Pre Endorsers n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
1 Studies considering 1996 checklist					
Sanchez-Thorin 2001	23/24	14/51		30.5 %	3.49 [1.92, 6.34]
Kober 2006	7/8	7/7		32.9 %	0.89 [0.57, 1.39]
Subtotal (99% CI)	**32**	**58**		**63.4 %**	**1.75 [0.30, 10.17]**
Total events: 30 (Post Endorsers), 21 (Pre Endorsers)					
Heterogeneity: $Tau^2 = 0.89$; $Chi^2 = 22.23$, df = 1 ($P<0.00001$); $I^2 = 96\%$					
Test for overall effect: Z = 0.82 (P = 0.41)					
2 Studies considering 2001 checklist					
Han 2008	272/276	162/166		36.6 %	1.01 [0.97, 1.05]
Subtotal (99% CI)	**276**	**166**		**36.6 %**	**1.01 [0.97, 1.05]**
Total events: 272 (Post Endorsers), 162 (Pre Endorsers)					
Heterogeneity: not applicable					
Test for overall effect: Z = 0.69 (P = 0.49)					
Total (99% CI)	**308**	**224**		**100.0 %**	**1.41 [0.63, 3.16]**
Total events: 302 (Post Endorsers), 183 (Pre Endorsers)					
Heterogeneity: $Tau^2 = 0.27$; $Chi^2 = 29.15$, df = 2 ($P<0.00001$); $I^2 = 93\%$					
Test for overall effect: Z = 1.11 (P = 0.27)					
Test for subgroup differences: $Chi^2 = 0.64$, df = 1 (P = 0.42), $I^2 = 0.0\%$					



Analysis 2.2. Comparison 2 CONSORT-endorsing journals before and after CONSORT endorsement, Outcome 2 Introduction.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 2 CONSORT-endorsing journals before and after CONSORT endorsement

Outcome: 2 Introduction

Study or subgroup	Post Endorsers n/N	Pre Endorsers n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
Han 2008	276/276	160/166		99.4 %	1.04 [1.00, 1.08]
Kober 2006	7/8	6/7		0.6 %	1.02 [0.60, 1.73]
Total (99% CI)	**284**	**173**		**100.0 %**	**1.04 [1.00, 1.08]**

Total events: 283 (Post Endorsers), 166 (Pre Endorsers)

Heterogeneity: $Tau^2 = 0.0$; $Chi^2 = 0.01$, df = 1 (P = 0.93); $I^2 = 0.0\%$

Test for overall effect: Z = 2.40 (P = 0.016)

Test for subgroup differences: Not applicable

Analysis 2.3. Comparison 2 CONSORT-endorsing journals before and after CONSORT endorsement, Outcome 3 Participants.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 2 CONSORT-endorsing journals before and after CONSORT endorsement

Outcome: 3 Participants

Study or subgroup	Post Endorsers n/N	Pre Endorsers n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
Sanchez-Thorin 2001 (1)	22/24	48/51		24.8 %	0.97 [0.81, 1.17]
Faunce 2003	2/2	5/5		2.2 %	1.00 [0.48, 2.09]
Han 2008	272/276	163/166		68.9 %	1.00 [0.97, 1.04]
Alvarez 2009	21/53	27/45		4.1 %	0.66 [0.39, 1.13]
Total (99% CI)	**355**	**267**		**100.0 %**	**0.98 [0.88, 1.09]**

Total events: 317 (Post Endorsers), 243 (Pre Endorsers)

Heterogeneity: $Tau^2 = 0.00$; $Chi^2 = 4.16$, df = 3 (P = 0.24); $I^2 = 28\%$

Test for overall effect: Z = 0.48 (P = 0.63)

Test for subgroup differences: Not applicable

(1) Alvarez- this is multicentre versus single centre may not be applicable

Analysis 2.4. Comparison 2 CONSORT-endorsing journals before and after CONSORT endorsement, Outcome 4 Interventions.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 2 CONSORT-endorsing journals before and after CONSORT endorsement

Outcome: 4 Interventions

Study or subgroup	Post Endorsers n/N	Pre Endorsers n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
Sanchez-Thorin 2001	23/24	50/51		14.6 %	0.98 [0.87, 1.10]
Kober 2006	6/8	7/7		0.6 %	0.77 [0.43, 1.38]
Han 2008	270/276	158/166		83.6 %	1.03 [0.98, 1.08]
Alvarez 2009	33/53	26/45		1.2 %	1.08 [0.70, 1.65]
Total (99% CI)	**361**	**269**		**100.0 %**	**1.02 [0.97, 1.07]**

Total events: 332 (Post Endorsers), 241 (Pre Endorsers)

Heterogeneity: $Tau^2 = 0.0$; $Chi^2 = 2.62$, df = 3 (P = 0.45); $I^2 = 0.0\%$

Test for overall effect: Z = 1.05 (P = 0.29)

Test for subgroup differences: Not applicable

Analysis 2.5. Comparison 2 CONSORT-endorsing journals before and after CONSORT endorsement, Outcome 5 Objectives.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 2 CONSORT-endorsing journals before and after CONSORT endorsement

Outcome: 5 Objectives

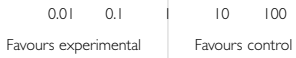
Study or subgroup	Post Endorsers n/N	Pre Endorsers n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
Han 2008	276/276	154/166		58.8 %	1.08 [1.02, 1.14]
Sanchez-Thorin 2001	23/24	50/51		41.2 %	0.98 [0.87, 1.10]
Total (99% CI)	**300**	**217**		**100.0 %**	**1.04 [0.91, 1.17]**

Total events: 299 (Post Endorsers), 204 (Pre Endorsers)

Heterogeneity: $Tau^2 = 0.00$; $Chi^2 = 3.62$, df = 1 (P = 0.06); $I^2 = 72\%$

Test for overall effect: Z = 0.73 (P = 0.47)

Test for subgroup differences: Not applicable



Analysis 2.6. Comparison 2 CONSORT-endorsing journals before and after CONSORT endorsement, Outcome 6 Outcomes.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 2 CONSORT-endorsing journals before and after CONSORT endorsement

Outcome: 6 Outcomes

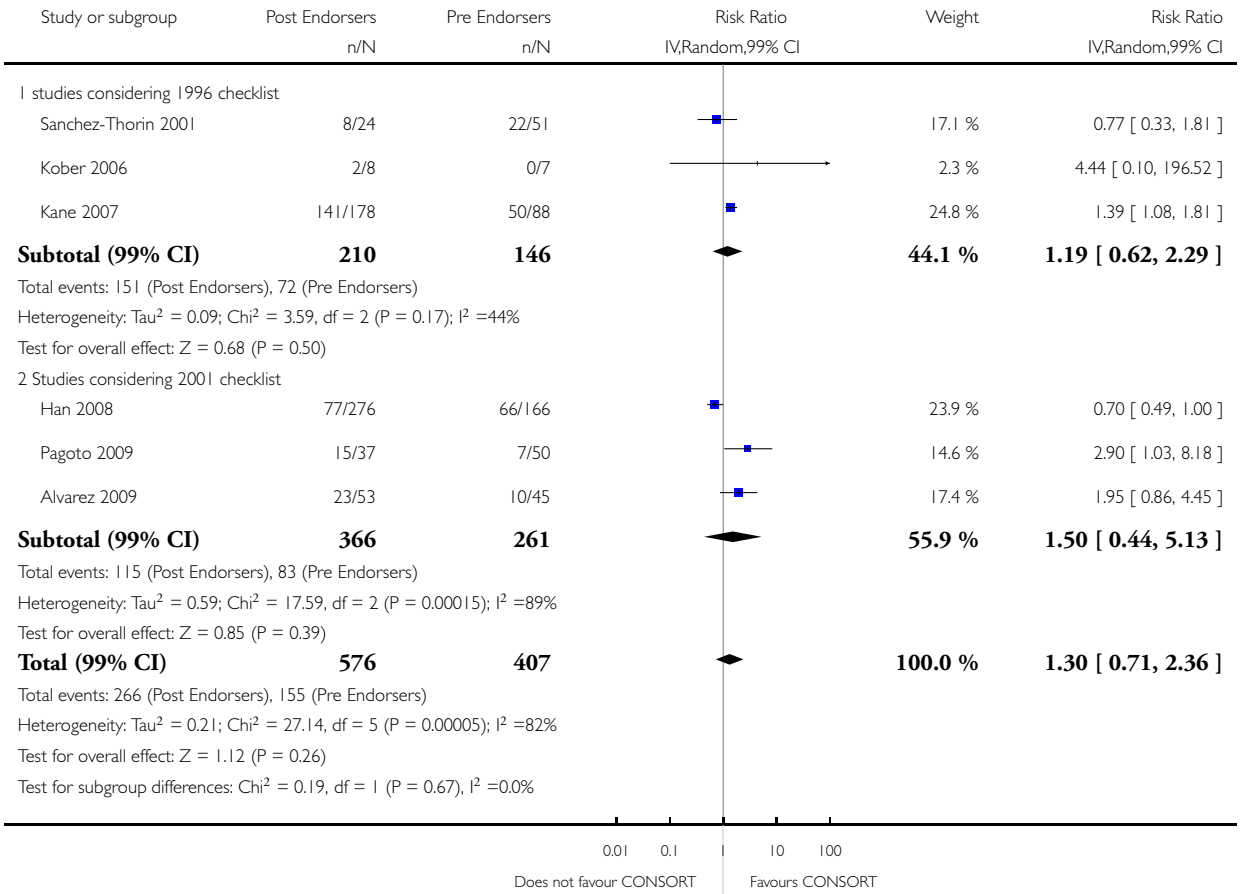
Study or subgroup	Post Endorsers n/N	Pre Endorsers n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
1 Studies considering 1996 checklist					
Sanchez-Thorin 2001	14/24	5/51		12.2 %	5.95 [1.83, 19.38]
Kober 2006	6/8	5/6		18.7 %	0.90 [0.44, 1.82]
Subtotal (99% CI)	**32**	**57**		**30.9 %**	**2.23 [0.20, 25.38]**
Total events: 20 (Post Endorsers), 10 (Pre Endorsers)					
Heterogeneity: Tau^2 = 1.64; Chi^2 = 12.51, df = 1 (P = 0.00040); I^2 =92%					
Test for overall effect: Z = 0.85 (P = 0.39)					
2 Studies considering 2001 checklist					
Han 2008	206/276	56/166		24.6 %	2.21 [1.65, 2.97]
Alvarez 2009	33/53	19/45		21.4 %	1.47 [0.87, 2.50]
Pagoto 2009	26/37	29/50		23.1 %	1.21 [0.80, 1.83]
Subtotal (99% CI)	**366**	**261**		**69.1 %**	**1.61 [0.95, 2.72]**
Total events: 265 (Post Endorsers), 104 (Pre Endorsers)					
Heterogeneity: Tau^2 = 0.10; Chi^2 = 10.13, df = 2 (P = 0.01); I^2 =80%					
Test for overall effect: Z = 2.32 (P = 0.020)					
Total (99% CI)	**398**	**318**		**100.0 %**	**1.68 [0.96, 2.96]**
Total events: 285 (Post Endorsers), 114 (Pre Endorsers)					
Heterogeneity: Tau^2 = 0.18; Chi^2 = 23.08, df = 4 (P = 0.00012); I^2 =83%					
Test for overall effect: Z = 2.38 (P = 0.017)					
Test for subgroup differences: Chi^2 = 0.12, df = 1 (P = 0.73), I^2 =0.0%					

Analysis 2.7. Comparison 2 CONSORT-endorsing journals before and after CONSORT endorsement, Outcome 7 Sample size.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 2 CONSORT-endorsing journals before and after CONSORT endorsement

Outcome: 7 Sample size



Analysis 2.8. Comparison 2 CONSORT-endorsing journals before and after CONSORT endorsement, Outcome 8 Sequence generation.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 2 CONSORT-endorsing journals before and after CONSORT endorsement

Outcome: 8 Sequence generation

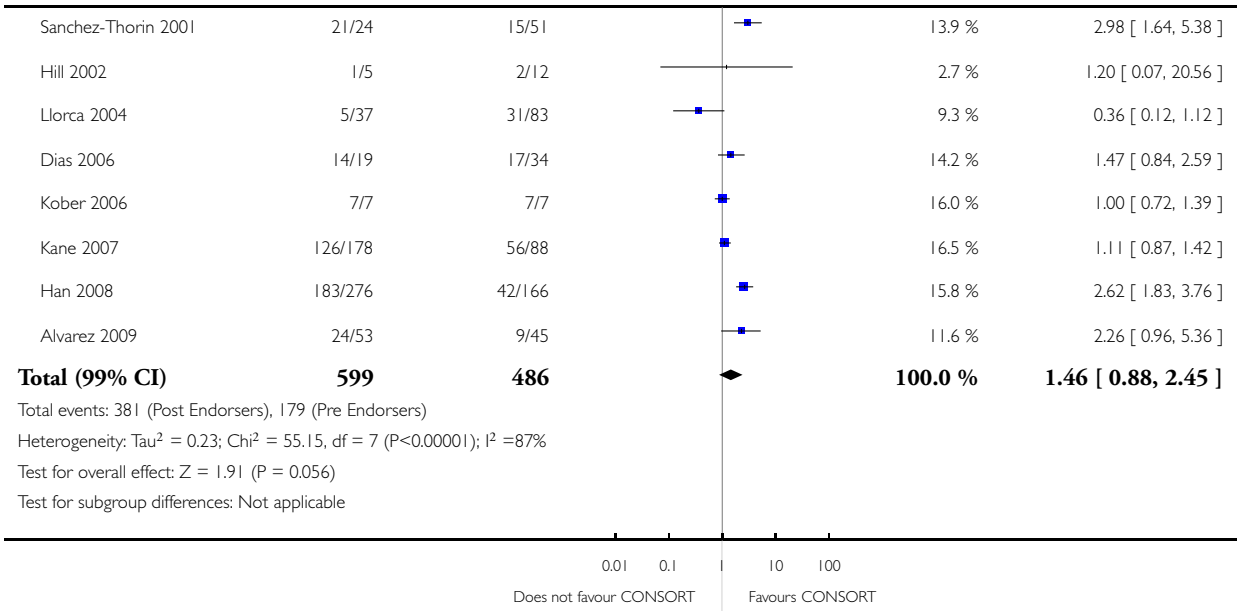
Study or subgroup	Post Endorsers n/N	Pre Endorsers n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
Sanchez-Thorin 2001	21/24	15/51		13.9 %	2.98 [1.64, 5.38]
Hill 2002	1/5	2/12		2.7 %	1.20 [0.07, 20.56]
Llorca 2004	5/37	31/83		9.3 %	0.36 [0.12, 1.12]
Dias 2006	14/19	17/34		14.2 %	1.47 [0.84, 2.59]
Kober 2006	7/7	7/7		16.0 %	1.00 [0.72, 1.39]
Kane 2007	126/178	56/88		16.5 %	1.11 [0.87, 1.42]
Han 2008	183/276	42/166		15.8 %	2.62 [1.83, 3.76]
Alvarez 2009	24/53	9/45		11.6 %	2.26 [0.96, 5.36]
Total (99% CI)	**599**	**486**		**100.0 %**	**1.46 [0.88, 2.45]**

Total events: 381 (Post Endorsers), 179 (Pre Endorsers)

Heterogeneity: $Tau^2 = 0.23$; $Chi^2 = 55.15$, df = 7 ($P<0.00001$); $I^2 = 87\%$

Test for overall effect: Z = 1.91 (P = 0.056)

Test for subgroup differences: Not applicable



Analysis 2.9. Comparison 2 CONSORT-endorsing journals before and after CONSORT endorsement, Outcome 9 Allocation concealment.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 2 CONSORT-endorsing journals before and after CONSORT endorsement

Outcome: 9 Allocation concealment

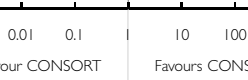
Study or subgroup	Post Endorsers n/N	Pre Endorsers n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
Sanchez-Thorin 2001	6/24	17/51		16.8 %	0.75 [0.26, 2.13]
Moher 2001	47/77	28/71		21.9 %	1.55 [0.99, 2.42]
Hill 2002	1/5	4/12		7.0 %	0.60 [0.05, 7.55]
Llorca 2004	5/37	31/83		16.0 %	0.36 [0.12, 1.12]
Dias 2006	9/19	8/34		17.1 %	2.01 [0.73, 5.53]
Han 2008	123/276	24/166		21.4 %	3.08 [1.84, 5.16]
Total (99% CI)	**438**	**417**		**100.0 %**	**1.23 [0.55, 2.74]**

Total events: 191 (Post Endorsers), 112 (Pre Endorsers)

Heterogeneity: $Tau^2 = 0.41$; $Chi^2 = 26.87$, df = 5 (P = 0.00006); $I^2 = 81\%$

Test for overall effect: Z = 0.67 (P = 0.50)

Test for subgroup differences: Not applicable



Analysis 2.10. Comparison 2 CONSORT-endorsing journals before and after CONSORT endorsement, Outcome 10 Implementation.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 2 CONSORT-endorsing journals before and after CONSORT endorsement

Outcome: 10 Implementation

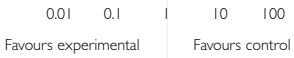
Study or subgroup	Post Endorsers n/N	Pre Endorsers n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
Han 2008	84/276	10/166		51.3 %	5.05 [2.22, 11.51]
Sanchez-Thorin 2001	5/24	15/51		48.7 %	0.71 [0.22, 2.28]
Total (99% CI)	**300**	**217**		**100.0 %**	**1.94 [0.15, 24.36]**

Total events: 89 (Post Endorsers), 25 (Pre Endorsers)

Heterogeneity: $Tau^2 = 1.78$; $Chi^2 = 12.55$, df = 1 (P = 0.00040); $I^2 = 92\%$

Test for overall effect: Z = 0.68 (P = 0.50)

Test for subgroup differences: Not applicable

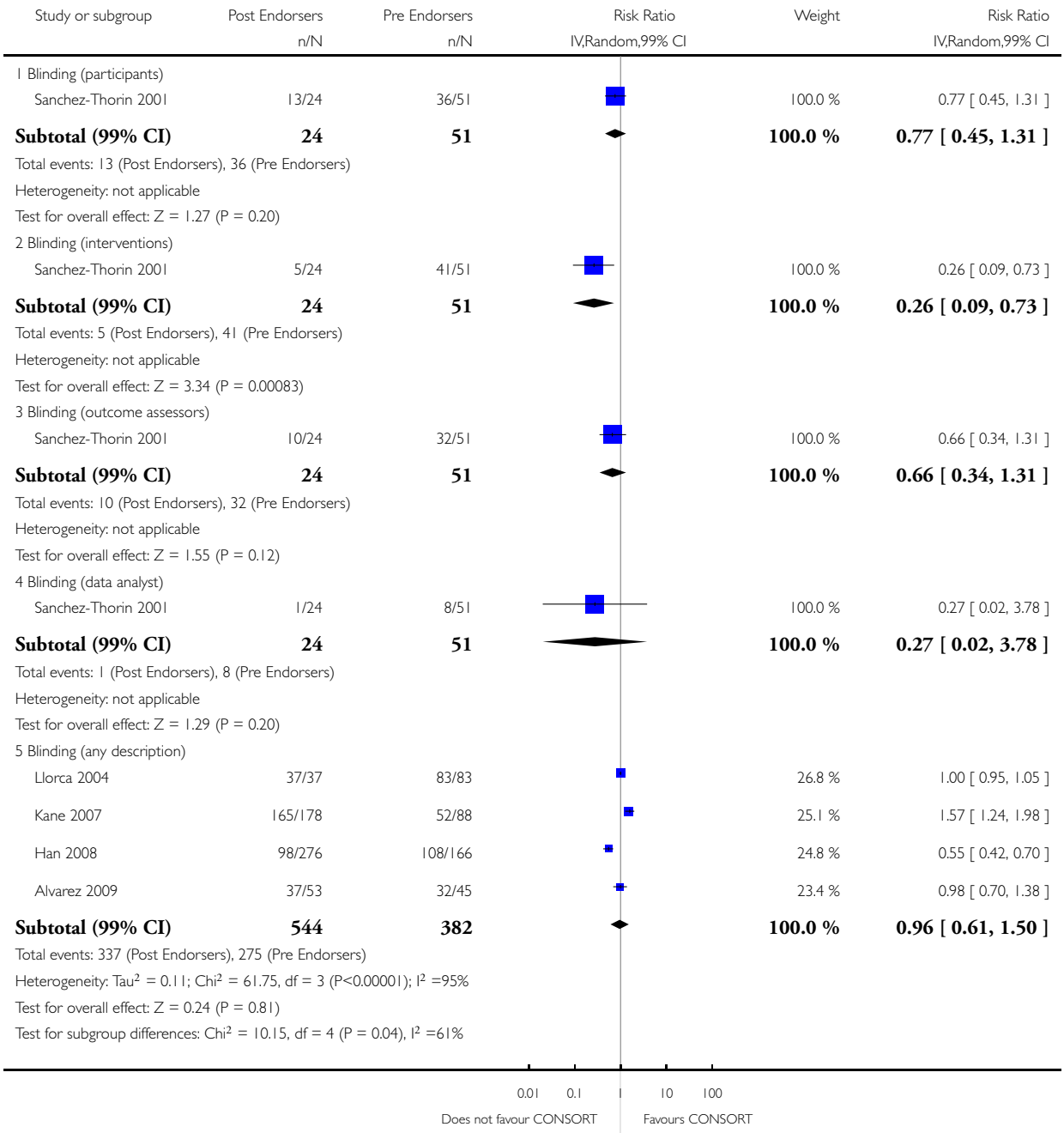


Analysis 2.11. Comparison 2 CONSORT-endorsing journals before and after CONSORT endorsement, Outcome 11 Blinding.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 2 CONSORT-endorsing journals before and after CONSORT endorsement

Outcome: 11 Blinding



Analysis 2.12. Comparison 2 CONSORT-endorsing journals before and after CONSORT endorsement, Outcome 12 Statistical methods.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 2 CONSORT-endorsing journals before and after CONSORT endorsement

Outcome: 12 Statistical methods

Study or subgroup	Post Endorsers n/N	Pre Endorsers n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
Sanchez-Thorin 2001	22/24	48/51		22.5 %	0.97 [0.81, 1.17]
Hill 2002	3/5	3/12		3.5 %	2.40 [0.49, 11.83]
Llorca 2004	27/37	77/83		20.8 %	0.79 [0.60, 1.03]
Kober 2006	5/8	5/7		8.0 %	0.88 [0.34, 2.23]
Han 2008	272/276	154/166		24.1 %	1.06 [1.00, 1.13]
Han 2008	98/276	108/166		21.1 %	0.55 [0.42, 0.70]
Total (99% CI)	**626**	**485**		**100.0 %**	**0.86 [0.62, 1.19]**

Total events: 427 (Post Endorsers), 395 (Pre Endorsers)

Heterogeneity: $Tau^2 = 0.07$; $Chi^2 = 51.54$, df = 5 ($P<0.00001$); $I^2 = 90\%$

Test for overall effect: Z = 1.18 (P = 0.24)

Test for subgroup differences: Not applicable

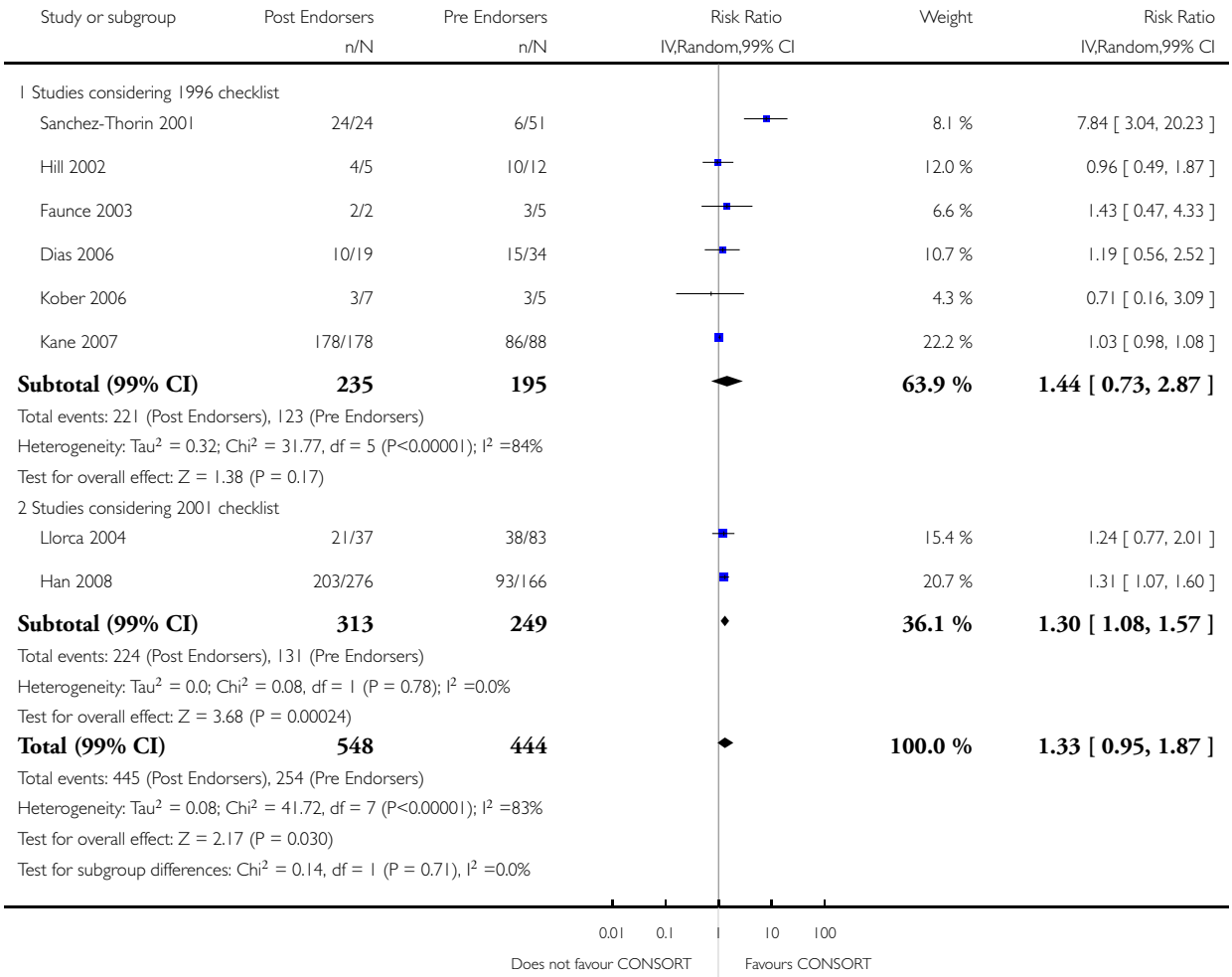
Analysis 2.13. Comparison 2 CONSORT-endorsing journals before and after CONSORT endorsement, Outcome 13 Participant flow.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 2 CONSORT-endorsing journals before and after CONSORT endorsement

Outcome: 13 Participant flow

Study or subgroup	Post Endorsers n/N	Pre Endorsers n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
1 Studies considering 1996 checklist					
Sanchez-Thorin 2001	24/24	6/51		8.1 %	7.84 [3.04, 20.23]
Hill 2002	4/5	10/12		12.0 %	0.96 [0.49, 1.87]
Faunce 2003	2/2	3/5		6.6 %	1.43 [0.47, 4.33]
Dias 2006	10/19	15/34		10.7 %	1.19 [0.56, 2.52]
Kober 2006	3/7	3/5		4.3 %	0.71 [0.16, 3.09]
Kane 2007	178/178	86/88		22.2 %	1.03 [0.98, 1.08]
Subtotal (99% CI)	**235**	**195**		**63.9 %**	**1.44 [0.73, 2.87]**
Total events: 221 (Post Endorsers), 123 (Pre Endorsers)					
Heterogeneity: Tau^2 = 0.32; Chi^2 = 31.77, df = 5 (P<0.00001); I^2 =84%					
Test for overall effect: Z = 1.38 (P = 0.17)					
2 Studies considering 2001 checklist					
Llorca 2004	21/37	38/83		15.4 %	1.24 [0.77, 2.01]
Han 2008	203/276	93/166		20.7 %	1.31 [1.07, 1.60]
Subtotal (99% CI)	**313**	**249**		**36.1 %**	**1.30 [1.08, 1.57]**
Total events: 224 (Post Endorsers), 131 (Pre Endorsers)					
Heterogeneity: Tau^2 = 0.0; Chi^2 = 0.08, df = 1 (P = 0.78); I^2 =0.0%					
Test for overall effect: Z = 3.68 (P = 0.00024)					
Total (99% CI)	**548**	**444**		**100.0 %**	**1.33 [0.95, 1.87]**
Total events: 445 (Post Endorsers), 254 (Pre Endorsers)					
Heterogeneity: Tau^2 = 0.08; Chi^2 = 41.72, df = 7 (P<0.00001); I^2 =83%					
Test for overall effect: Z = 2.17 (P = 0.030)					
Test for subgroup differences: Chi^2 = 0.14, df = 1 (P = 0.71), I^2 =0.0%					



Analysis 2.14. Comparison 2 CONSORT-endorsing journals before and after CONSORT endorsement, Outcome 14 Recruitment.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 2 CONSORT-endorsing journals before and after CONSORT endorsement

Outcome: 14 Recruitment

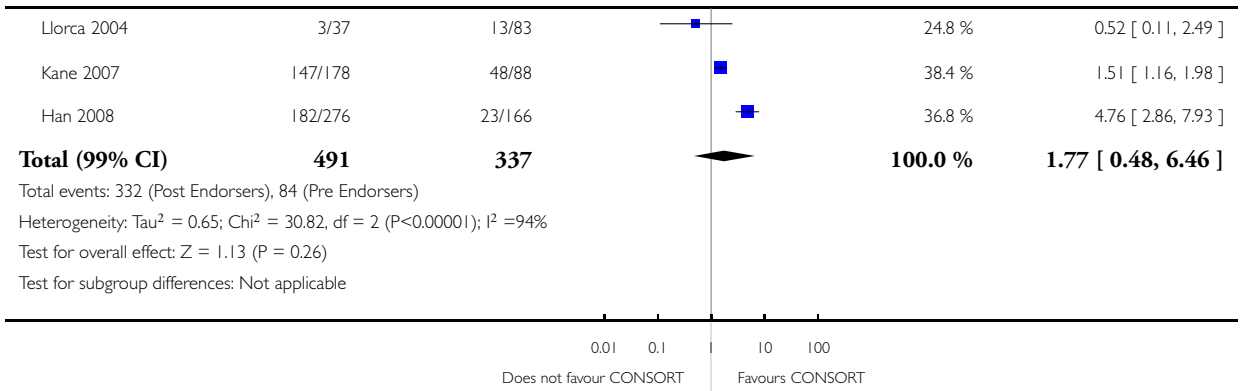
Study or subgroup	Post Endorsers n/N	Pre Endorsers n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
Llorca 2004	3/37	13/83		24.8 %	0.52 [0.11, 2.49]
Kane 2007	147/178	48/88		38.4 %	1.51 [1.16, 1.98]
Han 2008	182/276	23/166		36.8 %	4.76 [2.86, 7.93]
Total (99% CI)	**491**	**337**		**100.0 %**	**1.77 [0.48, 6.46]**

Total events: 332 (Post Endorsers), 84 (Pre Endorsers)

Heterogeneity: $Tau^2 = 0.65$; $Chi^2 = 30.82$, df = 2 ($P<0.00001$); $I^2 = 94\%$

Test for overall effect: Z = 1.13 (P = 0.26)

Test for subgroup differences: Not applicable



Analysis 2.15. Comparison 2 CONSORT-endorsing journals before and after CONSORT endorsement, Outcome 15 Baseline data.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 2 CONSORT-endorsing journals before and after CONSORT endorsement

Outcome: 15 Baseline data

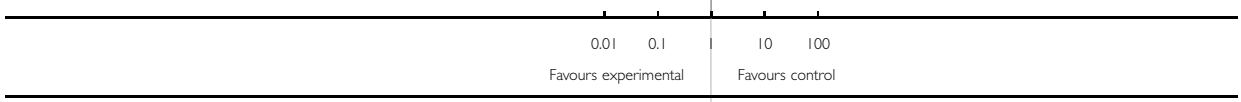
Study or subgroup	Post Endorsers n/N	Pre Endorsers n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
Han 2008	264/276	111/166		84.7 %	1.43 [1.24, 1.65]
Pagoto 2009	31/37	31/50		15.3 %	1.35 [0.96, 1.90]
Total (99% CI)	**313**	**216**		**100.0 %**	**1.42 [1.24, 1.62]**

Total events: 295 (Post Endorsers), 142 (Pre Endorsers)

Heterogeneity: $Tau^2 = 0.0$; $Chi^2 = 0.16$, df = 1 (P = 0.69); $I^2 = 0.0\%$

Test for overall effect: Z = 6.76 (P < 0.00001)

Test for subgroup differences: Not applicable



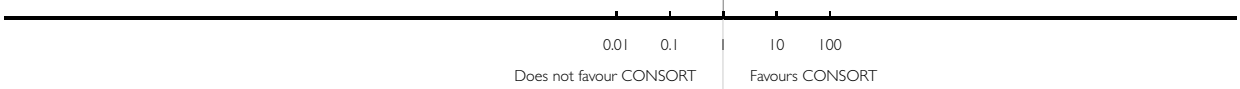
Analysis 2.16. Comparison 2 CONSORT-endorsing journals before and after CONSORT endorsement, Outcome 16 Numbers analysed.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 2 CONSORT-endorsing journals before and after CONSORT endorsement

Outcome: 16 Numbers analysed

Study or subgroup	Post Endorsers n/N	Pre Endorsers n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
1 Studies considering 1996 checklist					
Sanchez-Thorin 2001	24/24	41/51		28.6 %	1.23 [1.01, 1.49]
Kober 2006	1/8	1/7		1.2 %	0.88 [0.03, 25.96]
Kane 2007	116/178	29/88		22.3 %	1.98 [1.30, 3.00]
Subtotal (99% CI)	**210**	**146**		**52.1 %**	**1.50 [0.86, 2.62]**
Total events: 141 (Post Endorsers), 71 (Pre Endorsers)					
Heterogeneity: Tau^2 = 0.08; Chi^2 = 7.24, df = 2 (P = 0.03); I^2 =72%					
Test for overall effect: Z = 1.86 (P = 0.063)					
2 Studies considering 2001 checklist					
Llorca 2004	2/37	3/83		2.4 %	1.50 [0.15, 14.85]
Han 2008	259/276	81/166		28.3 %	1.92 [1.56, 2.37]
Pagoto 2009	26/37	16/50		17.2 %	2.20 [1.21, 3.99]
Subtotal (99% CI)	**350**	**299**		**47.9 %**	**1.95 [1.60, 2.37]**
Total events: 287 (Post Endorsers), 100 (Pre Endorsers)					
Heterogeneity: Tau^2 = 0.0; Chi^2 = 0.38, df = 2 (P = 0.83); I^2 =0.0%					
Test for overall effect: Z = 8.75 (P < 0.00001)					
Total (99% CI)	**560**	**445**		**100.0 %**	**1.72 [1.18, 2.49]**
Total events: 428 (Post Endorsers), 171 (Pre Endorsers)					
Heterogeneity: Tau^2 = 0.07; Chi^2 = 21.22, df = 5 (P = 0.00074); I^2 =76%					
Test for overall effect: Z = 3.74 (P = 0.00018)					
Test for subgroup differences: Chi^2 = 1.30, df = 1 (P = 0.26), I^2 =23%					



Analysis 2.17. Comparison 2 CONSORT-endorsing journals before and after CONSORT endorsement, Outcome 17 Outcomes and estimation.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 2 CONSORT-endorsing journals before and after CONSORT endorsement

Outcome: 17 Outcomes and estimation

Study or subgroup	Post Endorsers n/N	Pre Endorsers n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
Sanchez-Thorin 2001	24/24	49/51		43.6 %	1.03 [0.92, 1.15]
Kober 2006	5/8	3/7		14.5 %	1.46 [0.39, 5.50]
Han 2008	233/276	80/166		41.9 %	1.75 [1.41, 2.18]
Total (99% CI)	**308**	**224**		**100.0 %**	**1.35 [0.73, 2.51]**

Total events: 262 (Post Endorsers), 132 (Pre Endorsers)
Heterogeneity: $Tau^2 = 0.13$; $Chi^2 = 31.85$, df = 2 (P<0.00001); $I^2 = 94\%$
Test for overall effect: Z = 1.26 (P = 0.21)
Test for subgroup differences: Not applicable

Analysis 2.18. Comparison 2 CONSORT-endorsing journals before and after CONSORT endorsement, Outcome 18 Ancillary analyses.

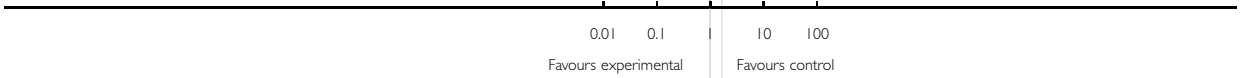
Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 2 CONSORT-endorsing journals before and after CONSORT endorsement

Outcome: 18 Ancillary analyses

Study or subgroup	Post Endorsers n/N	Pre Endorsers n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
Han 2008	253/276	44/166		100.0 %	3.46 [2.47, 4.84]
Total (99% CI)	**276**	**166**		**100.0 %**	**3.46 [2.47, 4.84]**

Total events: 253 (Post Endorsers), 44 (Pre Endorsers)
Heterogeneity: not applicable
Test for overall effect: Z = 9.51 (P < 0.00001)
Test for subgroup differences: Not applicable



Analysis 2.19. Comparison 2 CONSORT-endorsing journals before and after CONSORT endorsement, Outcome 19 Adverse events.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 2 CONSORT-endorsing journals before and after CONSORT endorsement

Outcome: 19 Adverse events

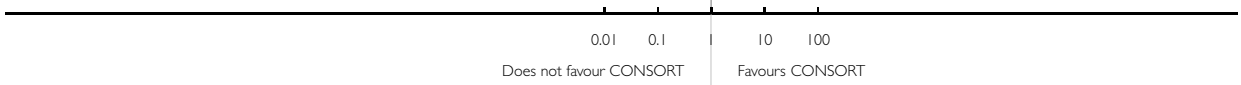
Study or subgroup	Post Endorsers n/N	Pre Endorsers n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
Dias 2006	3/19	3/34		1.2 %	1.79 [0.25, 12.83]
Kober 2006	6/8	3/4		5.8 %	1.00 [0.40, 2.49]
Han 2008	193/276	82/166		93.0 %	1.42 [1.13, 1.78]
Total (99% CI)	**303**	**204**		**100.0 %**	**1.39 [1.12, 1.73]**

Total events: 202 (Post Endorsers), 88 (Pre Endorsers)

Heterogeneity: $Tau^2 = 0.0$; $Chi^2 = 1.02$, df = 2 (P = 0.60); $I^2 = 0.0\%$

Test for overall effect: Z = 3.90 (P = 0.000097)

Test for subgroup differences: Not applicable



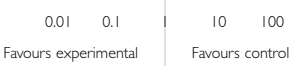
Analysis 2.20. Comparison 2 CONSORT-endorsing journals before and after CONSORT endorsement, Outcome 20 Interpretation.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 2 CONSORT-endorsing journals before and after CONSORT endorsement

Outcome: 20 Interpretation

Study or subgroup	Post Endorsers n/N	Pre Endorsers n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
Han 2008	276/276	164/166		100.0 %	1.01 [0.99, 1.04]
Total (99% CI)	**276**	**166**		**100.0 %**	**1.01 [0.99, 1.04]**
Total events: 276 (Post Endorsers), 164 (Pre Endorsers)					
Heterogeneity: not applicable					
Test for overall effect: Z = 1.34 (P = 0.18)					
Test for subgroup differences: Not applicable					



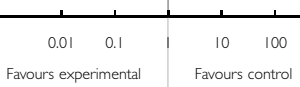
Analysis 2.21. Comparison 2 CONSORT-endorsing journals before and after CONSORT endorsement, Outcome 21 Generalisability.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 2 CONSORT-endorsing journals before and after CONSORT endorsement

Outcome: 21 Generalisability

Study or subgroup	Post Endorsers n/N	Pre Endorsers n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
Han 2008	270/276	92/166		100.0 %	1.77 [1.47, 2.11]
Total (99% CI)	**276**	**166**		**100.0 %**	**1.77 [1.47, 2.11]**
Total events: 270 (Post Endorsers), 92 (Pre Endorsers)					
Heterogeneity: not applicable					
Test for overall effect: Z = 8.10 (P < 0.00001)					
Test for subgroup differences: Not applicable					



Analysis 2.22. Comparison 2 CONSORT-endorsing journals before and after CONSORT endorsement, Outcome 22 Overall evidence.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 2 CONSORT-endorsing journals before and after CONSORT endorsement

Outcome: 22 Overall evidence

Study or subgroup	Post Endorsers n/N	Pre Endorsers n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
Sanchez-Thorin 2001	24/24	43/51		48.1 %	1.17 [0.98, 1.39]
Han 2008	276/276	114/166		51.9 %	1.46 [1.27, 1.67]
Total (99% CI)	**300**	**217**		**100.0 %**	**1.31 [0.99, 1.73]**

Total events: 300 (Post Endorsers), 157 (Pre Endorsers)

Heterogeneity: $Tau^2 = 0.02$; $Chi^2 = 6.45$, df = 1 (P = 0.01); $I^2 = 84\%$

Test for overall effect: Z = 2.50 (P = 0.013)

Test for subgroup differences: Not applicable

Analysis 2.23. Comparison 2 CONSORT-endorsing journals before and after CONSORT endorsement, Outcome 23 Total sum score.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 2 CONSORT-endorsing journals before and after CONSORT endorsement

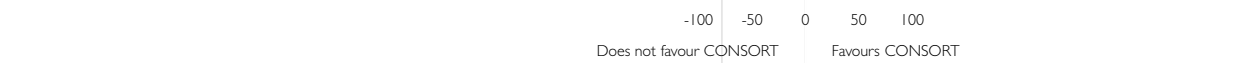
Outcome: 23 Total sum score

Study or subgroup	Post Endorsers N	Mean(SD)	Pre Endorsers N	Mean(SD)	Std. Mean Difference IV,Random,99% CI	Weight	Std. Mean Difference IV,Random,99% CI
Moher 2001	77	27.1 (4.81)	71	23.4 (5.1)		100.0 %	0.74 [0.30, 1.18]
Total (99% CI)	**77**		**71**			**100.0 %**	**0.74 [0.30, 1.18]**

Heterogeneity: not applicable

Test for overall effect: Z = 4.37 (P = 0.000013)

Test for subgroup differences: Not applicable



Analysis 3.1. Comparison 3 Sample of RCTs before and after CONSORT publication, Outcome 1 Title and abstract.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 3 Sample of RCTs before and after CONSORT publication

Outcome: 1 Title and abstract

Study or subgroup	Post-CONSORT n/N	Pre-CONSORT n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
1 Studies considering 1996 checklist					
Anttila 2006	7/7	5/7		5.9 %	1.36 [0.71, 2.63]
Thabane 2007	8/13	17/50		4.7 %	1.81 [0.85, 3.87]
Ziogas 2009	168/187	61/74		21.4 %	1.09 [0.94, 1.27]
Partsinevelou 2009	201/237	21/27		16.0 %	1.09 [0.83, 1.43]
Subtotal (99% CI)	**444**	**158**		**48.0 %**	**1.13 [0.96, 1.33]**
Total events: 384 (Post-CONSORT), 104 (Pre-CONSORT)					
Heterogeneity: $Tau^2 = 0.00$; $Chi^2 = 3.52$, df = 3 (P = 0.32); $I^2 = 15\%$					
Test for overall effect: Z = 1.93 (P = 0.053)					
2 Studies considering 2001 checklist					
Wang 2007	2774/4492	1353/2930		24.5 %	1.34 [1.26, 1.42]
Par s 2008	17/27	8/15		5.0 %	1.18 [0.57, 2.45]
Ladd 2010	83/89	63/70		22.5 %	1.04 [0.91, 1.18]
Subtotal (99% CI)	**4608**	**3015**		**52.0 %**	**1.18 [0.88, 1.59]**
Total events: 2874 (Post-CONSORT), 1424 (Pre-CONSORT)					
Heterogeneity: $Tau^2 = 0.03$; $Chi^2 = 22.25$, df = 2 (P = 0.00001); $I^2 = 91\%$					
Test for overall effect: Z = 1.45 (P = 0.15)					
Total (99% CI)	**5052**	**3173**		**100.0 %**	**1.18 [0.98, 1.42]**
Total events: 3258 (Post-CONSORT), 1528 (Pre-CONSORT)					
Heterogeneity: $Tau^2 = 0.02$; $Chi^2 = 31.95$, df = 6 (P = 0.00002); $I^2 = 81\%$					
Test for overall effect: Z = 2.35 (P = 0.019)					
Test for subgroup differences: $Chi^2 = 0.12$, df = 1 (P = 0.73), $I^2 = 0.0\%$					

Analysis 3.2. Comparison 3 Sample of RCTs before and after CONSORT publication, Outcome 2 Introduction.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 3 Sample of RCTs before and after CONSORT publication

Outcome: 2 Introduction

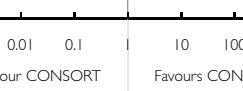
Study or subgroup	Post-CONSORT n/N	Pre-CONSORT n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
Anttila 2006	7/7	6/7		6.5 %	1.15 [0.70, 1.90]
Thabane 2007	13/13	50/50		14.3 %	1.00 [0.87, 1.15]
Peckitt 2007	43/45	21/26		11.4 %	1.18 [0.91, 1.53]
Wang 2007	2697/4492	1350/2930		15.5 %	1.30 [1.23, 1.38]
Par s 2008	27/27	14/15		12.4 %	1.08 [0.87, 1.35]
Partsinevelou 2009	210/237	24/27		13.2 %	1.00 [0.83, 1.20]
Ziogas 2009	142/187	45/74		11.3 %	1.25 [0.96, 1.62]
Ladd 2010	83/87	68/69		15.4 %	0.97 [0.90, 1.04]
Total (99% CI)	**5095**	**3198**		**100.0 %**	**1.10 [0.94, 1.30]**

Total events: 3222 (Post-CONSORT), 1578 (Pre-CONSORT)

Heterogeneity: $Tau^2 = 0.03$; $Chi^2 = 78.01$, df = 7 ($P<0.00001$); $I^2 = 91\%$

Test for overall effect: Z = 1.55 (P = 0.12)

Test for subgroup differences: Not applicable



Analysis 3.3. Comparison 3 Sample of RCTs before and after CONSORT publication, Outcome 3 Participants.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 3 Sample of RCTs before and after CONSORT publication

Outcome: 3 Participants

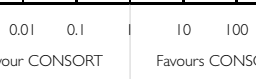
Study or subgroup	Post-CONSORT n/N	Pre-CONSORT n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
Anttila 2006	5/7	5/7		2.5 %	1.00 [0.42, 2.39]
Thabane 2007	13/13	37/50		15.7 %	1.31 [1.02, 1.69]
Wang 2007	2844/4492	1770/2930		29.3 %	1.05 [1.00, 1.10]
Bausch 2009	88/105	150/239		21.4 %	1.34 [1.13, 1.58]
Ziogas 2009	181/187	73/74		29.2 %	0.98 [0.93, 1.03]
Partsinevelou 2009	171/237	5/27		1.8 %	3.90 [1.37, 11.08]
Total (99% CI)	**5041**	**3327**		**100.0 %**	**1.15 [0.99, 1.33]**

Total events: 3302 (Post-CONSORT), 2040 (Pre-CONSORT)

Heterogeneity: $Tau^2 = 0.01$; $Chi^2 = 39.73$, df = 5 ($P<0.00001$); $I^2 = 87\%$

Test for overall effect: Z = 2.44 (P = 0.015)

Test for subgroup differences: Not applicable



Analysis 3.4. Comparison 3 Sample of RCTs before and after CONSORT publication, Outcome 4 Interventions.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 3 Sample of RCTs before and after CONSORT publication

Outcome: 4 Interventions

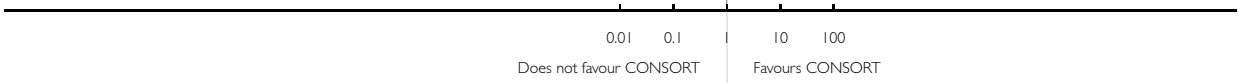
Study or subgroup	Post-CONSORT n/N	Pre-CONSORT n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
Anttila 2006	6/7	5/7		0.2 %	1.20 [0.58, 2.50]
Thabane 2007	13/13	49/50		4.2 %	0.99 [0.86, 1.15]
Wang 2007	2904/4492	1826/2930		34.4 %	1.04 [0.99, 1.09]
Par s 2008	27/27	15/15		5.0 %	1.00 [0.88, 1.14]
Partsinevelou 2009	234/237	27/27		17.6 %	1.00 [0.94, 1.07]
Ziogas 2009	179/187	73/74		27.7 %	0.97 [0.92, 1.02]
Ladd 2010	84/88	67/70		11.0 %	1.00 [0.91, 1.09]
Total (99% CI)	**5051**	**3173**		**100.0 %**	**1.00 [0.97, 1.04]**

Total events: 3447 (Post-CONSORT), 2062 (Pre-CONSORT)

Heterogeneity: $Tau^2 = 0.00$; $Chi^2 = 6.46$, df = 6 (P = 0.37); $I^2 = 7\%$

Test for overall effect: Z = 0.39 (P = 0.70)

Test for subgroup differences: Not applicable



Analysis 3.5. Comparison 3 Sample of RCTs before and after CONSORT publication, Outcome 5 Objectives.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 3 Sample of RCTs before and after CONSORT publication

Outcome: 5 Objectives

Study or subgroup	Post-CONSORT n/N	Pre-CONSORT n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
Anttila 2006	6/7	5/7		0.5 %	1.20 [0.58, 2.50]
Peckitt 2007	44/47	15/20		2.0 %	1.25 [0.88, 1.77]
Wang 2007	3575/4492	2287/2930		50.4 %	1.02 [0.99, 1.05]
Partsinevelou 2009	231/237	23/27		5.2 %	1.14 [0.93, 1.41]
Ziogas 2009	183/187	73/74		41.9 %	0.99 [0.95, 1.04]
Total (99% CI)	**4970**	**3058**		**100.0 %**	**1.02 [0.97, 1.07]**

Total events: 4039 (Post-CONSORT), 2403 (Pre-CONSORT)

Heterogeneity: $Tau^2 = 0.00$; $Chi^2 = 6.75$, df = 4 (P = 0.15); $I^2 = 41\%$

Test for overall effect: Z = 0.97 (P = 0.33)

Test for subgroup differences: Not applicable

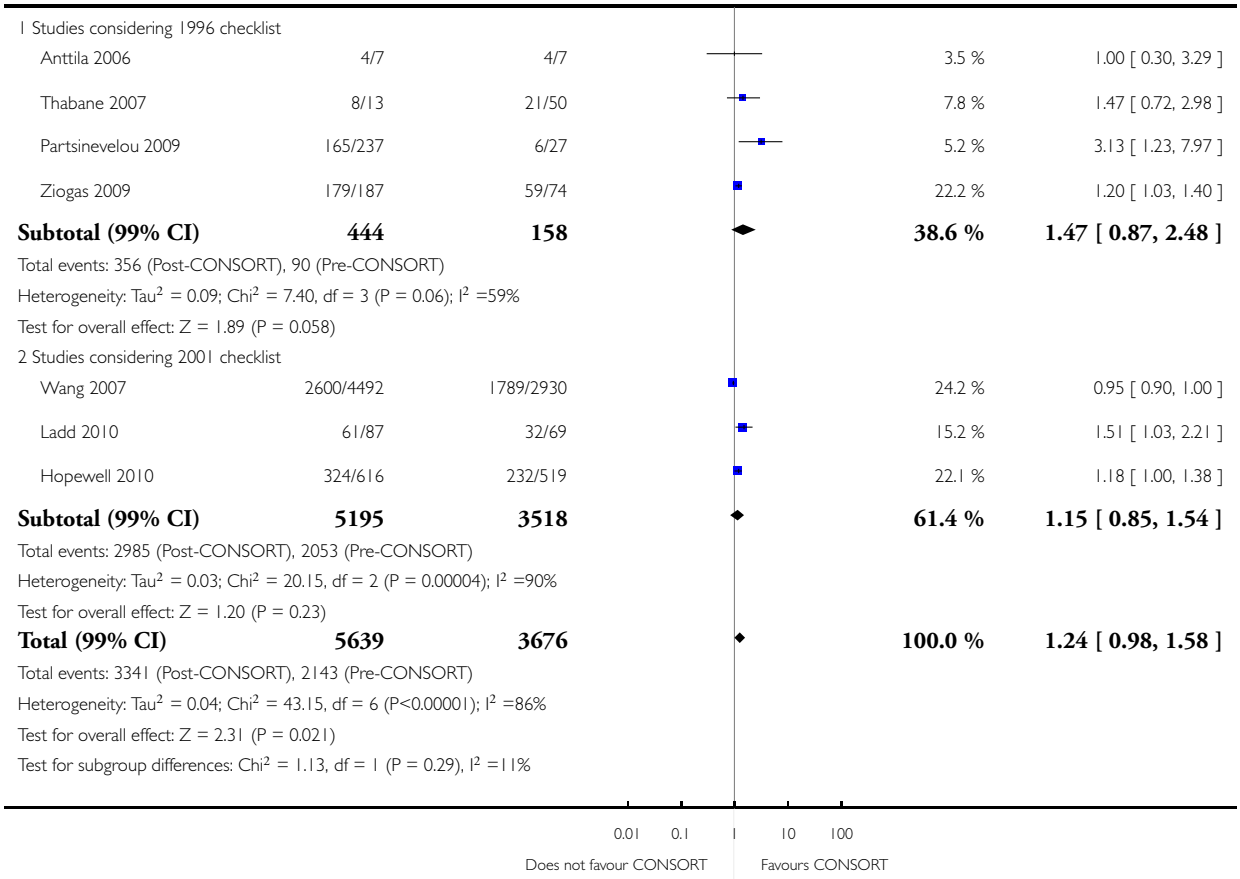
Analysis 3.6. Comparison 3 Sample of RCTs before and after CONSORT publication, Outcome 6 Outcomes.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 3 Sample of RCTs before and after CONSORT publication

Outcome: 6 Outcomes

Study or subgroup	Post-CONSORT n/N	Pre-CONSORT n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
1 Studies considering 1996 checklist					
Anttila 2006	4/7	4/7		3.5 %	1.00 [0.30, 3.29]
Thabane 2007	8/13	21/50		7.8 %	1.47 [0.72, 2.98]
Partsinevelou 2009	165/237	6/27		5.2 %	3.13 [1.23, 7.97]
Ziogas 2009	179/187	59/74		22.2 %	1.20 [1.03, 1.40]
Subtotal (99% CI)	**444**	**158**		**38.6 %**	**1.47 [0.87, 2.48]**
Total events: 356 (Post-CONSORT), 90 (Pre-CONSORT)					
Heterogeneity: Tau^2 = 0.09; Chi^2 = 7.40, df = 3 (P = 0.06); I^2 =59%					
Test for overall effect: Z = 1.89 (P = 0.058)					
2 Studies considering 2001 checklist					
Wang 2007	2600/4492	1789/2930		24.2 %	0.95 [0.90, 1.00]
Ladd 2010	61/87	32/69		15.2 %	1.51 [1.03, 2.21]
Hopewell 2010	324/616	232/519		22.1 %	1.18 [1.00, 1.38]
Subtotal (99% CI)	**5195**	**3518**		**61.4 %**	**1.15 [0.85, 1.54]**
Total events: 2985 (Post-CONSORT), 2053 (Pre-CONSORT)					
Heterogeneity: Tau^2 = 0.03; Chi^2 = 20.15, df = 2 (P = 0.00004); I^2 =90%					
Test for overall effect: Z = 1.20 (P = 0.23)					
Total (99% CI)	**5639**	**3676**		**100.0 %**	**1.24 [0.98, 1.58]**
Total events: 3341 (Post-CONSORT), 2143 (Pre-CONSORT)					
Heterogeneity: Tau^2 = 0.04; Chi^2 = 43.15, df = 6 (P<0.00001); I^2 =86%					
Test for overall effect: Z = 2.31 (P = 0.021)					
Test for subgroup differences: Chi^2 = 1.13, df = 1 (P = 0.29), I^2 =11%					



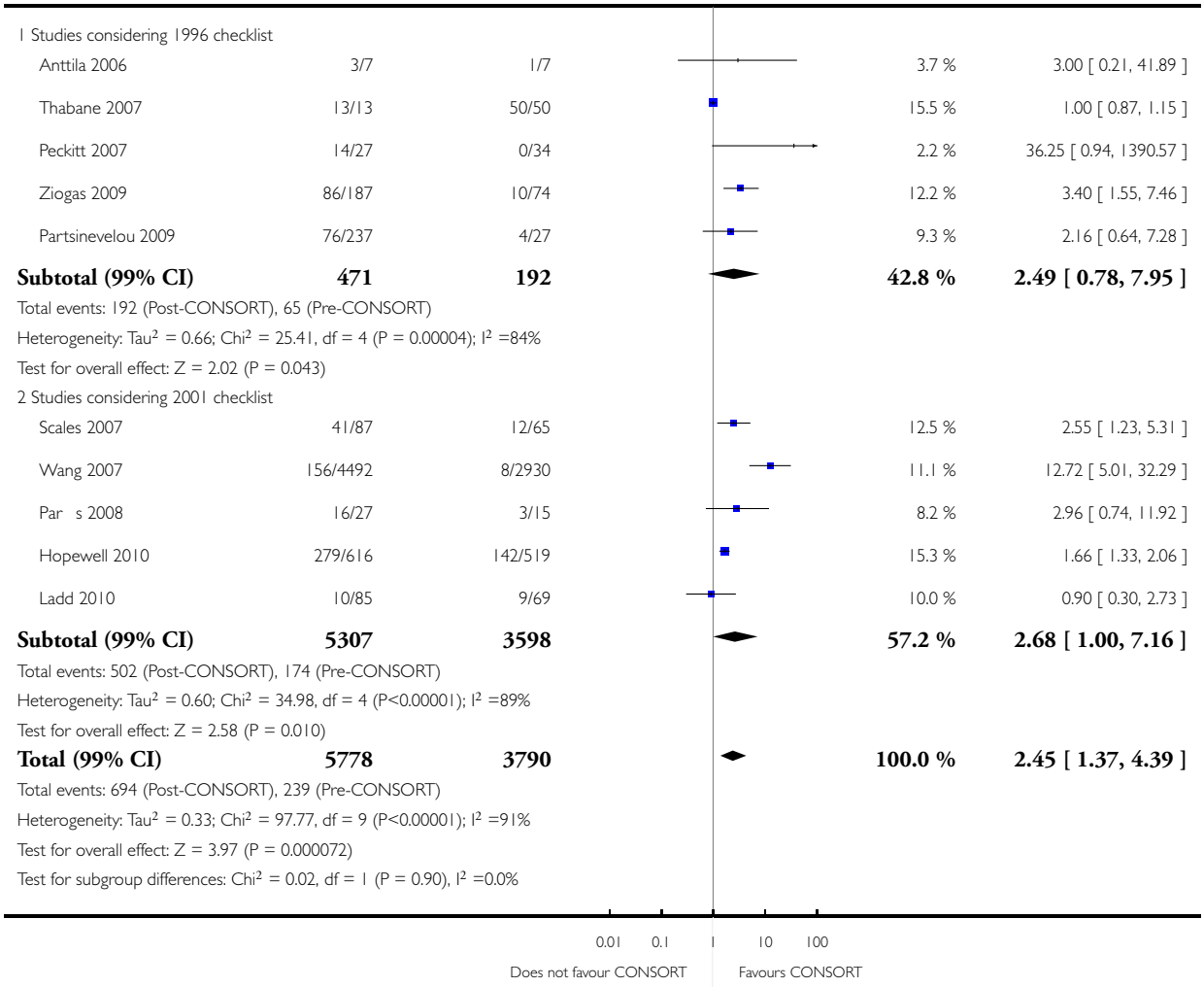
Analysis 3.7. Comparison 3 Sample of RCTs before and after CONSORT publication, Outcome 7 Sample size.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 3 Sample of RCTs before and after CONSORT publication

Outcome: 7 Sample size

Study or subgroup	Post-CONSORT n/N	Pre-CONSORT n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
1 Studies considering 1996 checklist					
Anttila 2006	3/7	1/7		3.7 %	3.00 [0.21, 41.89]
Thabane 2007	13/13	50/50		15.5 %	1.00 [0.87, 1.15]
Peckitt 2007	14/27	0/34		2.2 %	36.25 [0.94, 1390.57]
Ziogas 2009	86/187	10/74		12.2 %	3.40 [1.55, 7.46]
Partsinevelou 2009	76/237	4/27		9.3 %	2.16 [0.64, 7.28]
Subtotal (99% CI)	**471**	**192**		**42.8 %**	**2.49 [0.78, 7.95]**
Total events: 192 (Post-CONSORT), 65 (Pre-CONSORT)					
Heterogeneity: Tau^2 = 0.66; Chi^2 = 25.41, df = 4 (P = 0.00004); I^2 =84%					
Test for overall effect: Z = 2.02 (P = 0.043)					
2 Studies considering 2001 checklist					
Scales 2007	41/87	12/65		12.5 %	2.55 [1.23, 5.31]
Wang 2007	156/4492	8/2930		11.1 %	12.72 [5.01, 32.29]
Par s 2008	16/27	3/15		8.2 %	2.96 [0.74, 11.92]
Hopewell 2010	279/616	142/519		15.3 %	1.66 [1.33, 2.06]
Ladd 2010	10/85	9/69		10.0 %	0.90 [0.30, 2.73]
Subtotal (99% CI)	**5307**	**3598**		**57.2 %**	**2.68 [1.00, 7.16]**
Total events: 502 (Post-CONSORT), 174 (Pre-CONSORT)					
Heterogeneity: Tau^2 = 0.60; Chi^2 = 34.98, df = 4 (P<0.00001); I^2 =89%					
Test for overall effect: Z = 2.58 (P = 0.010)					
Total (99% CI)	**5778**	**3790**		**100.0 %**	**2.45 [1.37, 4.39]**
Total events: 694 (Post-CONSORT), 239 (Pre-CONSORT)					
Heterogeneity: Tau^2 = 0.33; Chi^2 = 97.77, df = 9 (P<0.00001); I^2 =91%					
Test for overall effect: Z = 3.97 (P = 0.000072)					
Test for subgroup differences: Chi^2 = 0.02, df = 1 (P = 0.90), I^2 =0.0%					



Analysis 3.8. Comparison 3 Sample of RCTs before and after CONSORT publication, Outcome 8 Sequence generation.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 3 Sample of RCTs before and after CONSORT publication

Outcome: 8 Sequence generation

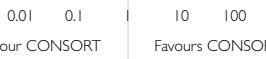
Study or subgroup	Post-CONSORT n/N	Pre-CONSORT n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
Anttila 2006	3/7	6/7		5.9 %	0.50 [0.15, 1.65]
Wang 2007	484/4492	97/2930		13.0 %	3.25 [2.46, 4.31]
Scales 2007	30/87	12/65		8.9 %	1.87 [0.86, 4.04]
Thabane 2007	2/13	12/50		3.4 %	0.64 [0.11, 3.86]
Par s 2008	20/27	9/15		10.2 %	1.23 [0.67, 2.29]
Prady 2008	19/39	7/51		7.1 %	3.55 [1.31, 9.63]
Ziogas 2009	80/187	22/74		11.2 %	1.44 [0.86, 2.39]
Bausch 2009	38/105	55/239		11.6 %	1.57 [1.00, 2.47]
Partsinevelou 2009	140/237	8/27		8.8 %	1.99 [0.92, 4.33]
Hopewell 2010	209/616	109/519		13.1 %	1.62 [1.24, 2.10]
Ladd 2010	16/81	8/66		6.9 %	1.63 [0.58, 4.57]
Total (99% CI)	**5891**	**4043**		**100.0 %**	**1.67 [1.14, 2.45]**

Total events: 1041 (Post-CONSORT), 345 (Pre-CONSORT)

Heterogeneity: $Tau^2 = 0.16$; $Chi^2 = 46.89$, df = 10 (P<0.00001); I^2 =79%

Test for overall effect: Z = 3.48 (P = 0.00050)

Test for subgroup differences: Not applicable



Analysis 3.9. Comparison 3 Sample of RCTs before and after CONSORT publication, Outcome 9 Allocation concealment.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 3 Sample of RCTs before and after CONSORT publication

Outcome: 9 Allocation concealment

Study or subgroup	Post-CONSORT n/N	Pre-CONSORT n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
Dickinson 2002	0/6	1/31		0.4 %	1.52 [0.03, 88.93]
Wang 2007	14/4492	7/2930		4.7 %	1.30 [0.40, 4.29]
Thabane 2007	1/13	5/50		1.0 %	0.77 [0.05, 11.51]
Scales 2007	19/87	5/65		4.5 %	2.84 [0.83, 9.65]
Prady 2008	19/39	7/51		6.5 %	3.55 [1.31, 9.63]
Par s 2008	18/27	5/15		6.4 %	2.00 [0.73, 5.46]
Selman 2008	21/35	10/39		9.9 %	2.34 [1.06, 5.14]
Bausch 2009	16/105	24/239		10.2 %	1.52 [0.70, 3.29]
Ziogas 2009	50/187	14/74		12.2 %	1.41 [0.71, 2.83]
Ladd 2010	14/86	11/66		7.1 %	0.98 [0.38, 2.52]
Hopewell 2010	156/616	94/519		37.2 %	1.40 [1.04, 1.89]
Total (99% CI)	**5693**	**4079**		**100.0 %**	**1.61 [1.23, 2.10]**

Total events: 328 (Post-CONSORT), 183 (Pre-CONSORT)

Heterogeneity: $Tau^2 = 0.02$; $Chi^2 = 11.50$, df = 10 (P = 0.32); $I^2 = 13\%$

Test for overall effect: Z = 4.60 (P < 0.00001)

Test for subgroup differences: Not applicable

Analysis 3.10. Comparison 3 Sample of RCTs before and after CONSORT publication, Outcome 10 Implementation.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 3 Sample of RCTs before and after CONSORT publication

Outcome: 10 Implementation

Study or subgroup	Post-CONSORT n/N	Pre-CONSORT n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
Anttila 2006	2/7	0/7		8.1 %	5.00 [0.11, 218.40]
Scales 2007	10/87	3/65		32.8 %	2.49 [0.48, 12.87]
Thabane 2007	0/13	3/50		7.9 %	0.52 [0.01, 23.64]
Ziogas 2009	13/187	7/74		51.2 %	0.73 [0.23, 2.33]
Total (99% CI)	**294**	**196**		**100.0 %**	**1.25 [0.41, 3.79]**

Total events: 25 (Post-CONSORT), 13 (Pre-CONSORT)

Heterogeneity: $Tau^2 = 0.16$; $Chi^2 = 3.76$, df = 3 (P = 0.29); $I^2 = 20\%$

Test for overall effect: Z = 0.51 (P = 0.61)

Test for subgroup differences: Not applicable

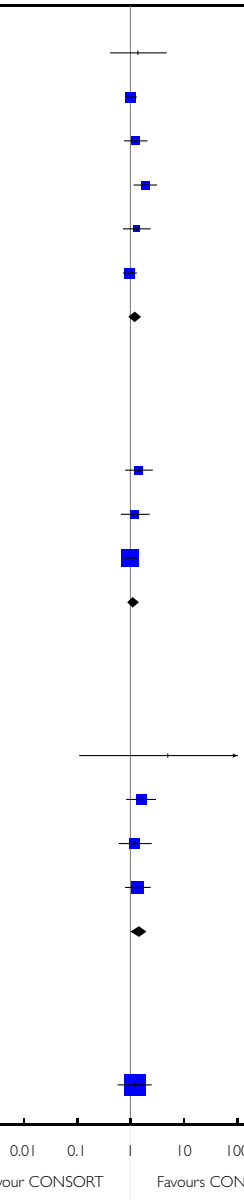
Analysis 3.11. Comparison 3 Sample of RCTs before and after CONSORT publication, Outcome 11 Blinding.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 3 Sample of RCTs before and after CONSORT publication

Outcome: 11 Blinding

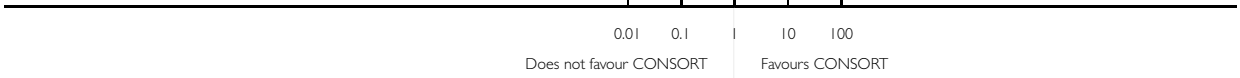
Study or subgroup	Post-CONSORT n/N	Pre-CONSORT n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
1 Blinding (participants)					
Dickinson 2002	3/6	11/31		4.2 %	1.41 [0.41, 4.79]
Wang 2007	232/4492	147/2930		26.5 %	1.03 [0.79, 1.34]
Scales 2007	42/87	25/65		16.0 %	1.26 [0.77, 2.06]
Thabane 2007	11/13	22/50		15.4 %	1.92 [1.15, 3.21]
Prady 2008	20/39	20/51		12.6 %	1.31 [0.72, 2.39]
Bausch 2009	55/105	127/239		25.3 %	0.99 [0.74, 1.31]
Subtotal (99% CI)	**4742**	**3366**		**100.0 %**	**1.21 [0.93, 1.58]**
Total events: 363 (Post-CONSORT), 352 (Pre-CONSORT)					
Heterogeneity: Tau^2 = 0.03; Chi^2 = 10.41, df = 5 (P = 0.06); I^2 =52%					
Test for overall effect: Z = 1.83 (P = 0.067)					
2 Blinding (intervenor)					
Scales 2007	35/87	18/65		15.3 %	1.45 [0.78, 2.69]
Prady 2008	19/39	20/51		15.3 %	1.24 [0.67, 2.30]
Bausch 2009	55/105	122/239		69.3 %	1.03 [0.77, 1.37]
Subtotal (99% CI)	**231**	**355**		**100.0 %**	**1.11 [0.88, 1.42]**
Total events: 109 (Post-CONSORT), 160 (Pre-CONSORT)					
Heterogeneity: Tau^2 = 0.0; Chi^2 = 1.97, df = 2 (P = 0.37); I^2 =0.0%					
Test for overall effect: Z = 1.16 (P = 0.25)					
3 Blinding (outcome assessor)					
Anttila 2006	2/7	0/7		0.9 %	5.00 [0.11, 218.40]
Scales 2007	34/87	16/65		30.3 %	1.59 [0.82, 3.06]
Prady 2008	16/39	17/51		25.9 %	1.23 [0.60, 2.50]
Bausch 2009	27/105	44/239		42.8 %	1.40 [0.80, 2.43]
Subtotal (99% CI)	**238**	**362**		**100.0 %**	**1.42 [0.99, 2.04]**
Total events: 79 (Post-CONSORT), 77 (Pre-CONSORT)					
Heterogeneity: Tau^2 = 0.0; Chi^2 = 1.20, df = 3 (P = 0.75); I^2 =0.0%					
Test for overall effect: Z = 2.51 (P = 0.012)					
4 Blinding (data analyst)					
Anttila 2006	6/7	5/7		100.0 %	1.20 [0.58, 2.50]



(*Continued . . .*)

(. . . *Continued*)

Study or subgroup	Post-CONSORT n/N	Pre-CONSORT n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
Subtotal (99% CI)	**7**	**7**		**100.0 %**	**1.20 [0.58, 2.50]**
Total events: 6 (Post-CONSORT), 5 (Pre-CONSORT)					
Heterogeneity: not applicable					
Test for overall effect: Z = 0.64 (P = 0.52)					
5 Blinding (any description)					
Ziogas 2009	26/187	10/74		6.5 %	1.03 [0.42, 2.51]
Partsinevelou 2009	110/237	10/27		11.4 %	1.25 [0.64, 2.45]
Hopewell 2010	160/616	148/519		82.1 %	0.91 [0.71, 1.17]
Subtotal (99% CI)	**1040**	**620**		**100.0 %**	**0.95 [0.76, 1.19]**
Total events: 296 (Post-CONSORT), 168 (Pre-CONSORT)					
Heterogeneity: Tau^2 = 0.0; Chi^2 = 1.37, df = 2 (P = 0.50); I^2 =0.0%					
Test for overall effect: Z = 0.56 (P = 0.58)					
Test for subgroup differences: Chi^2 = 6.88, df = 4 (P = 0.14), I^2 =42%					



Analysis 3.12. Comparison 3 Sample of RCTs before and after CONSORT publication, Outcome 12 Statistical methods.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 3 Sample of RCTs before and after CONSORT publication

Outcome: 12 Statistical methods

Study or subgroup	Post-CONSORT n/N	Pre-CONSORT n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
Anttila 2006	6/7	5/7		2.0 %	1.20 [0.58, 2.50]
Thabane 2007	13/13	47/50		15.6 %	1.04 [0.88, 1.22]
Wang 2007	2210/4492	1161/2930		22.2 %	1.24 [1.16, 1.33]
Par s 2008	26/27	15/15		15.7 %	0.98 [0.83, 1.15]
Partsinevelou 2009	228/237	24/27		14.7 %	1.08 [0.91, 1.29]
Ziogas 2009	174/187	57/74		15.1 %	1.21 [1.02, 1.43]
Ladd 2010	81/87	55/70		14.7 %	1.18 [0.99, 1.42]
Total (99% CI)	**5050**	**3173**		**100.0 %**	**1.13 [1.01, 1.25]**

Total events: 2738 (Post-CONSORT), 1364 (Pre-CONSORT)

Heterogeneity: $Tau^2 = 0.01$; $Chi^2 = 17.95$, df = 6 (P = 0.01); $I^2 = 67\%$

Test for overall effect: Z = 2.83 (P = 0.0047)

Test for subgroup differences: Not applicable

Analysis 3.13. Comparison 3 Sample of RCTs before and after CONSORT publication, Outcome 13 Participant flow.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 3 Sample of RCTs before and after CONSORT publication

Outcome: 13 Participant flow

Study or subgroup	Post-CONSORT n/N	Pre-CONSORT n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
1 Studies considering 1996 checklist					
Anttila 2006	0/7	1/7		0.5 %	0.33 [0.01, 18.28]
Thabane 2007	13/13	47/50		26.5 %	1.04 [0.88, 1.22]
Ziogas 2009	172/187	61/74		26.9 %	1.12 [0.96, 1.30]
Partsinevelou 2009	159/237	6/27		7.6 %	3.02 [1.19, 7.69]
Subtotal (99% CI)	**444**	**158**		**61.4 %**	**1.16 [0.87, 1.53]**
Total events: 344 (Post-CONSORT), 115 (Pre-CONSORT)					
Heterogeneity: Tau^2 = 0.02; Chi^2 = 9.27, df = 3 (P = 0.03); I^2 =68%					
Test for overall effect: Z = 1.35 (P = 0.18)					
2 Studies considering 2001 checklist					
Wang 2007	63/4492	24/2930		13.0 %	1.71 [0.93, 3.17]
Scales 2007	17/87	2/65		2.3 %	6.35 [0.97, 41.56]
Par s 2008	22/27	2/15		2.8 %	6.11 [1.10, 33.86]
Ladd 2010	55/88	37/67		20.5 %	1.13 [0.79, 1.61]
Subtotal (99% CI)	**4694**	**3077**		**38.6 %**	**2.14 [0.90, 5.09]**
Total events: 157 (Post-CONSORT), 65 (Pre-CONSORT)					
Heterogeneity: Tau^2 = 0.28; Chi^2 = 12.29, df = 3 (P = 0.01); I^2 =76%					
Test for overall effect: Z = 2.27 (P = 0.023)					
Total (99% CI)	**5138**	**3235**		**100.0 %**	**1.36 [1.01, 1.83]**
Total events: 501 (Post-CONSORT), 180 (Pre-CONSORT)					
Heterogeneity: Tau^2 = 0.05; Chi^2 = 24.93, df = 7 (P = 0.00078); I^2 =72%					
Test for overall effect: Z = 2.62 (P = 0.0088)					
Test for subgroup differences: Chi^2 = 3.04, df = 1 (P = 0.08), I^2 =67%					

Analysis 3.14. Comparison 3 Sample of RCTs before and after CONSORT publication, Outcome 14 Recruitment.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 3 Sample of RCTs before and after CONSORT publication

Outcome: 14 Recruitment

Study or subgroup	Post-CONSORT n/N	Pre-CONSORT n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
Anttila 2006	3/7	5/7		1.2 %	0.60 [0.17, 2.16]
Thabane 2007	13/13	49/50		29.5 %	0.99 [0.86, 1.15]
Wang 2007	3158/4492	1909/2930		41.9 %	1.08 [1.03, 1.13]
Partsinevelou 2009	97/237	1/27		0.3 %	11.05 [0.88, 139.51]
Ziogas 2009	151/187	61/74		27.1 %	0.98 [0.83, 1.16]
Total (99% CI)	**4936**	**3088**		**100.0 %**	**1.03 [0.89, 1.18]**

Total events: 3422 (Post-CONSORT), 2025 (Pre-CONSORT)

Heterogeneity: $Tau^2 = 0.01$; $Chi^2 = 10.76$, df = 4 (P = 0.03); $I^2 = 63\%$

Test for overall effect: Z = 0.47 (P = 0.64)

Test for subgroup differences: Not applicable

Analysis 3.15. Comparison 3 Sample of RCTs before and after CONSORT publication, Outcome 15 Baseline data.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 3 Sample of RCTs before and after CONSORT publication

Outcome: 15 Baseline data

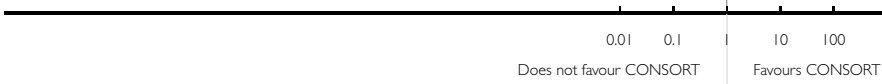
Study or subgroup	Post-CONSORT n/N	Pre-CONSORT n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
Anttila 2006	6/7	5/7		5.0 %	1.20 [0.58, 2.50]
Thabane 2007	4/13	12/50		1.9 %	1.28 [0.37, 4.49]
Wang 2007	4034/4492	2413/2930		42.6 %	1.09 [1.06, 1.12]
Prady 2008	35/39	32/51		18.2 %	1.43 [1.05, 1.95]
Ziogas 2009	131/187	47/74		22.2 %	1.10 [0.85, 1.43]
Partsinevelou 2009	198/237	14/27		10.0 %	1.61 [0.99, 2.61]
Total (99% CI)	**4975**	**3139**		**100.0 %**	**1.20 [1.01, 1.43]**

Total events: 4408 (Post-CONSORT), 2523 (Pre-CONSORT)

Heterogeneity: $Tau^2 = 0.01$; $Chi^2 = 9.51$, df = 5 (P = 0.09); $I^2 = 47\%$

Test for overall effect: Z = 2.73 (P = 0.0063)

Test for subgroup differences: Not applicable



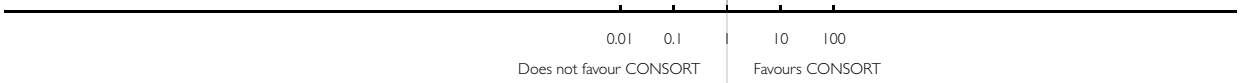
Analysis 3.16. Comparison 3 Sample of RCTs before and after CONSORT publication, Outcome 16 Numbers analysed.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 3 Sample of RCTs before and after CONSORT publication

Outcome: 16 Numbers analysed

Study or subgroup	Post-CONSORT n/N	Pre-CONSORT n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
1 Studies considering 1996 checklist					
Anttila 2006	5/7	5/7		14.8 %	1.00 [0.42, 2.39]
Partsinevelou 2009	33/237	0/27		2.0 %	7.88 [0.21, 298.34]
Ziogas 2009	67/187	7/74		13.6 %	3.79 [1.45, 9.89]
Subtotal (99% CI)	**431**	**108**		**30.4 %**	**2.32 [0.50, 10.87]**
Total events: 105 (Post-CONSORT), 12 (Pre-CONSORT)					
Heterogeneity: $Tau^2 = 0.70$; $Chi^2 = 8.05$, df = 2 (P = 0.02); $I^2 = 75\%$					
Test for overall effect: Z = 1.41 (P = 0.16)					
2 Studies considering 2001 checklist					
Scales 2007	25/87	22/65		18.1 %	0.85 [0.46, 1.58]
Prady 2008	13/39	8/51		12.9 %	2.13 [0.77, 5.89]
Par s 2008	3/14	3/10		6.5 %	0.71 [0.12, 4.38]
Bausch 2009	29/105	46/239		19.4 %	1.43 [0.84, 2.44]
Ladd 2010	21/88	7/70		12.6 %	2.39 [0.84, 6.79]
Subtotal (99% CI)	**333**	**435**		**69.6 %**	**1.37 [0.80, 2.36]**
Total events: 91 (Post-CONSORT), 86 (Pre-CONSORT)					
Heterogeneity: $Tau^2 = 0.10$; $Chi^2 = 7.82$, df = 4 (P = 0.10); $I^2 = 49\%$					
Test for overall effect: Z = 1.50 (P = 0.13)					
Total (99% CI)	**764**	**543**		**100.0 %**	**1.57 [0.91, 2.70]**
Total events: 196 (Post-CONSORT), 98 (Pre-CONSORT)					
Heterogeneity: $Tau^2 = 0.19$; $Chi^2 = 17.62$, df = 7 (P = 0.01); $I^2 = 60\%$					
Test for overall effect: Z = 2.13 (P = 0.033)					
Test for subgroup differences: $Chi^2 = 0.69$, df = 1 (P = 0.41), $I^2 = 0.0\%$					



Analysis 3.17. Comparison 3 Sample of RCTs before and after CONSORT publication, Outcome 17 Outcomes and estimation.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 3 Sample of RCTs before and after CONSORT publication

Outcome: 17 Outcomes and estimation

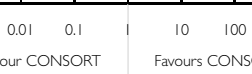
Study or subgroup	Post-CONSORT n/N	Pre-CONSORT n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
Anttila 2006	2/7	2/7		0.1 %	1.00 [0.11, 8.82]
Peckitt 2007	34/34	10/11		5.4 %	1.13 [0.85, 1.50]
Wang 2007	4352/4492	2756/2930		20.9 %	1.03 [1.02, 1.04]
Thabane 2007	13/13	48/50		11.3 %	1.01 [0.87, 1.18]
Par s 2008	26/27	13/15		5.6 %	1.11 [0.84, 1.47]
Ziogas 2009	179/187	69/74		16.4 %	1.03 [0.94, 1.12]
Partsinevelou 2009	235/237	27/27		18.1 %	1.01 [0.94, 1.08]
Bausch 2009	58/105	40/239		2.7 %	3.30 [2.14, 5.10]
Ladd 2010	89/89	68/69		19.4 %	1.02 [0.97, 1.07]
Total (99% CI)	**5191**	**3422**		**100.0 %**	**1.06 [0.98, 1.15]**

Total events: 4988 (Post-CONSORT), 3033 (Pre-CONSORT)

Heterogeneity: $Tau^2 = 0.00$; $Chi^2 = 49.92$, df = 8 (P<0.00001); $I^2 = 84\%$

Test for overall effect: Z = 2.06 (P = 0.040)

Test for subgroup differences: Not applicable



Analysis 3.18. Comparison 3 Sample of RCTs before and after CONSORT publication, Outcome 18 Ancillary analysis.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 3 Sample of RCTs before and after CONSORT publication

Outcome: 18 Ancillary analysis

Study or subgroup	Post-CONSORT n/N	Pre-CONSORT n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
Anttila 2006	1/7	2/7		6.3 %	0.50 [0.03, 8.54]
Wang 2007	130/4492	182/2930		25.8 %	0.47 [0.35, 0.62]
Partsinevelou 2009	99/237	4/27		16.8 %	2.82 [0.85, 9.41]
Ziogas 2009	131/187	36/74		25.5 %	1.44 [1.03, 2.01]
Yu 2010	113/422	84/355		25.6 %	1.13 [0.82, 1.56]
Total (99% CI)	**5345**	**3393**		**100.0 %**	**1.06 [0.47, 2.39]**

Total events: 474 (Post-CONSORT), 308 (Pre-CONSORT)

Heterogeneity: $Tau^2 = 0.37$; $Chi^2 = 57.21$, df = 4 (P<0.00001); $I^2 = 93\%$

Test for overall effect: Z = 0.19 (P = 0.85)

Test for subgroup differences: Not applicable

Analysis 3.19. Comparison 3 Sample of RCTs before and after CONSORT publication, Outcome 19 Adverse events.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 3 Sample of RCTs before and after CONSORT publication

Outcome: 19 Adverse events

Study or subgroup	Post-CONSORT n/N	Pre-CONSORT n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
Thabane 2007	9/13	23/50		5.6 %	1.51 [0.81, 2.79]
Wang 2007	869/4492	563/2930		37.4 %	1.01 [0.89, 1.14]
Par s 2008	12/17	10/11		8.8 %	0.78 [0.48, 1.25]
Partsinevelou 2009	123/237	9/27		4.2 %	1.56 [0.76, 3.20]
Ziogas 2009	121/187	42/74		17.6 %	1.14 [0.85, 1.53]
Ladd 2010	70/84	51/64		26.4 %	1.05 [0.85, 1.28]
Total (99% CI)	**5030**	**3156**		**100.0 %**	**1.06 [0.91, 1.24]**

Total events: 1204 (Post-CONSORT), 698 (Pre-CONSORT)

Heterogeneity: $Tau^2 = 0.01$; $Chi^2 = 8.07$, df = 5 (P = 0.15); $I^2 = 38\%$

Test for overall effect: Z = 0.94 (P = 0.35)

Test for subgroup differences: Not applicable

Analysis 3.20. Comparison 3 Sample of RCTs before and after CONSORT publication, Outcome 20 Interpretation.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 3 Sample of RCTs before and after CONSORT publication

Outcome: 20 Interpretation

Study or subgroup	Post-CONSORT n/N	Pre-CONSORT n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
Wang 2007	4228/4492	2776/2930		91.0 %	0.99 [0.98, 1.01]
Par s 2008	27/27	15/15		1.1 %	1.00 [0.88, 1.14]
Partsinevelou 2009	233/237	27/27		4.1 %	1.00 [0.93, 1.07]
Ziogas 2009	180/187	71/74		3.8 %	1.00 [0.93, 1.08]
Total (99% CI)	**4943**	**3046**		**100.0 %**	**0.99 [0.98, 1.01]**

Total events: 4668 (Post-CONSORT), 2889 (Pre-CONSORT)

Heterogeneity: $Tau^2 = 0.0$; $Chi^2 = 0.17$, df = 3 (P = 0.98); $I^2 = 0.0\%$

Test for overall effect: Z = 1.08 (P = 0.28)

Test for subgroup differences: Not applicable

Analysis 3.21. Comparison 3 Sample of RCTs before and after CONSORT publication, Outcome 21 Generalisability.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 3 Sample of RCTs before and after CONSORT publication

Outcome: 21 Generalisability

Study or subgroup	Post-CONSORT n/N	Pre-CONSORT n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
Thabane 2007	1/13	9/50		0.1 %	0.43 [0.03, 5.72]
Wang 2007	1218/4492	752/2930		54.5 %	1.06 [0.95, 1.17]
Ziogas 2009	172/187	62/74		28.1 %	1.10 [0.95, 1.27]
Partsinevelou 2009	219/237	24/27		17.4 %	1.04 [0.87, 1.25]
Total (99% CI)	**4929**	**3081**		**100.0 %**	**1.06 [0.99, 1.15]**

Total events: 1610 (Post-CONSORT), 847 (Pre-CONSORT)

Heterogeneity: $Tau^2 = 0.0$; $Chi^2 = 1.28$, df = 3 (P = 0.73); $I^2 = 0.0\%$

Test for overall effect: Z = 2.11 (P = 0.035)

Test for subgroup differences: Not applicable

Analysis 3.22. Comparison 3 Sample of RCTs before and after CONSORT publication, Outcome 22 Overall evidence.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 3 Sample of RCTs before and after CONSORT publication

Outcome: 22 Overall evidence

Study or subgroup	Post-CONSORT n/N	Pre-CONSORT n/N	Risk Ratio IV,Random,99% CI	Weight	Risk Ratio IV,Random,99% CI
Wang 2007	2841/4492	1791/2930		59.4 %	1.03 [0.99, 1.09]
Thabane 2007	0/13	5/50		0.1 %	0.33 [0.01, 13.73]
Partsinevelou 2009	215/237	22/27		16.2 %	1.11 [0.87, 1.42]
Ziogas 2009	168/187	56/74		24.3 %	1.19 [0.99, 1.42]
Total (99% CI)	**4929**	**3081**		**100.0 %**	**1.08 [0.97, 1.21]**

Total events: 3224 (Post-CONSORT), 1874 (Pre-CONSORT)

Heterogeneity: $Tau^2 = 0.00$; $Chi^2 = 4.65$, df = 3 (P = 0.20); $I^2 = 36\%$

Test for overall effect: Z = 1.80 (P = 0.072)

Test for subgroup differences: Not applicable

Analysis 3.23. Comparison 3 Sample of RCTs before and after CONSORT publication, Outcome 23 Total sum score.

Review: Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals

Comparison: 3 Sample of RCTs before and after CONSORT publication

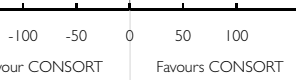
Outcome: 23 Total sum score

Study or subgroup	Post-CONSORT N	Mean(SD)	Pre-CONSORT N	Mean(SD)	Std. Mean Difference IV,Random,99% CI	Weight	Std. Mean Difference IV,Random,99% CI
Bian 2006	24	19.78 (3.85)	43	20.81 (4.21)		20.7 %	-0.25 [-0.91, 0.41]
Thoma 2006	7	12.14 (3.8)	11	8.36 (3.11)		14.2 %	1.06 [-0.28, 2.41]
Lai 2006 (1)	43	10 (2.22)	31	7 (3.7)		20.8 %	1.01 [0.37, 1.66]
Agha 2007	42	10.91 (2.49)	48	11.18 (2.43)		21.7 %	-0.11 [-0.65, 0.44]
Zhong 2010	227	14.6 (3.85)	52	10.88 (3.33)		22.7 %	0.99 [0.58, 1.40]
Total (99% CI)	**343**		**185**			**100.0 %**	**0.51 [-0.28, 1.30]**

Heterogeneity: $Tau^2 = 0.39$; $Chi^2 = 31.36$, df = 4 (P<0.00001); $I^2 = 87\%$

Test for overall effect: Z = 1.66 (P = 0.096)

Test for subgroup differences: Not applicable



(1) Score out of 15; Reported as median (IQR).

ADDITIONAL TABLES

Table 1. 2001 CONSORT checklist of items to include when reporting a randomised controlled trial

PAPER SECTION and topic	Item	
TITLE and ABSTRACT	1	How participants were allocated to interventions (e.g. 'random allocation', 'randomised', or 'randomly assigned')
INTRODUCTION Background	2	Scientific background and explanation of rationale
METHODS Participants	3	Eligibility criteria for participants and the settings and locations where the data were collected
Interventions	4	Precise details of the interventions intended for each group and how and when they were actually administered

Table 1. 2001 CONSORT checklist of items to include when reporting a randomised controlled trial (*Continued*)

Objectives	5	Specific objectives and hypotheses
Outcomes	6	Clearly defined primary and secondary outcome measures and, when applicable, any methods used to enhance the quality of measurements (e.g. multiple observations, training of assessors)
Sample size	7	How sample size was determined and, when applicable, explanation of any interim analyses and stopping rules
Randomisation Sequence generation	8	Method used to generate the random allocation sequence, including details of any restriction (e.g. blocking, stratification)
Randomisation Allocation concealment	9	Method used to implement the random allocation sequence (e.g. numbered containers or central telephone), clarifying whether the sequence was concealed until interventions were assigned
Randomisation Implementation	10	Who generated the allocation sequence, who enrolled participants, and who assigned participants to their groups
Blinding (masking)	11	Whether or not participants, those administering the interventions, and those assessing the outcomes were blinded to group assignment. When relevant, how the success of blinding was evaluated
Statistical methods	12	Statistical methods used to compare groups for primary outcome(s); methods for additional analyses, such as subgroup analyses and adjusted analyses
RESULTS Participant flow	13	Flow of participants through each stage (a diagram is strongly recommended). Specifically, for each group report the numbers of participants randomly assigned, receiving intended treatment, completing the study protocol, and analysed for the primary outcome. Describe protocol deviations from study as planned, together with reasons
Recruitment	14	Dates defining the periods of recruitment and follow-up
Baseline data	15	Baseline demographic and clinical characteristics of each group
Numbers analysed	16	Number of participants (denominator) in each group included in each analysis and whether the analysis was by 'intention-to-treat'. State the results in absolute numbers when feasible (e.g. 10/20, not 50%)
Outcomes and estimation	17	For each primary and secondary outcome, a summary of results for each group, and the estimated effect size and its precision (e.g. 95% confidence interval)
Ancillary analyses	18	Address multiplicity by reporting any other analyses performed, including subgroup analyses and adjusted analyses, indicating those pre-specified and those exploratory

Table 1. 2001 CONSORT checklist of items to include when reporting a randomised controlled trial (*Continued*)

Adverse events	19	All important adverse events or side effects in each intervention group
DISCUSSION Interpretation	20	Interpretation of the results, taking into account study hypotheses, sources of potential bias or imprecision, and the dangers associated with multiplicity of analyses and outcomes
Generalisability	21	Generalisability (external validity) of the trial findings
Overall evidence	22	General interpretation of the results in the context of current evidence

APPENDICES

Appendix 1. Search strategy

Ovid MEDLINE(R) <2005 to March Week 1 2010>

1 *randomized controlled trials/
2 *clinical trials/
3 Evidence-Based Medicine/
4 research design/
5 publishing/st
6 Practice Guidelines/
7 Guidelines/st
8 writing/st
9 or/1-8
10 quality control/
11 reproducibility of results/
12 "bias (epidemiology)"/
13 epidemiologic methods/
14 publication bias/
15 ethics, professional/
16 or/10-15
17 (consort or consolidat$ standard$).tw.
18 9 and 16
19 18 not randomized controlled trial.pt.
20 limit 19 to abstracts
21 19 not 20
22 17 or 21
23 limit 22 to (comment or editorial or guideline or letter)
24 22 not 23

EMBASE <1980 to 2010 Week 16>

1 (consort or consolidat$ standard$).tw.
2 *randomized controlled trials/
3 *clinical trials/
4 Evidence-Based Medicine/

5 research design/
6 Publishing/
7 Practice Guidelines/
8 Writing/
9 or/2-8
10 quality control/
11 reproducibility/
12 validation process/
13 epidemiology/
14 research ethics/
15 or/10-14
16 9 and 15
17 limit 16 to abstracts
18 16 not 17
19 1 or 18
20 limit 19 to "reviews (2 or more terms high specificity)"
21 limit 19 to (editorial or letter)
22 19 not (20 or 21)

ISI Web of Knowledge: 27 March 2010

TS=(consort AND (checklist* OR quality))

DocType=All document types; Language=All languages; Database=SCI-EXPANDED, SSCI;

Cochrane Methodology Register and Cochrane Database of Systematic Reviews, *The Cochrane Library, 2010, Issue 2 (Wiley interface)*

Cochrane [all fields]

PubMed 'Related Items' search (27 May 2010)

Using two PMIDS: PMID: 12161081 or PMID: 11308436.

Appendix 2. Validity assessment tool

Assessment of risk of bias (validity assessment) in included studies:

Question	Possible Responses
The RCTs included in the study represent a large cohort (i.e. an entire year), or were randomly chosen from a large cohort	Yes (low) No (high) Can't tell (medium)
The reviewer(s) who assessed CONSORT criteria was blinded to study authors, institutions, sponsorship, and/or journal name	Yes (low) No (high) Can't tell (medium)
Was consideration of potential clustering reported? (If potential for clustering does not exist, answer 'low' risk)	Yes (low) No (high) Can't tell (medium) Not applicable
There is no evidence of selective outcome reporting	Yes (low) No (high) Can't tell (medium)

(*Continued*)

More than one reviewer assessed CONSORT criteria	Yes (low) No (high) Can't tell (medium)
If applicable (i.e. more than one reviewer assessed CONSORT criteria), whether inter-reviewer agreement was greater than or equal to 90% agreement or a kappa statistic of 0.8	Yes (low) No (high) Can't tell (medium) Not applicable
If quality of included RCTs was assessed, the reviewer(s) conducted a blinded assessment	Yes (low) No (high) Can't tell (medium) Not applicable

WHAT'S NEW

Date	Event	Description
12 December 2012	Amended	Edited to re-format PDF.

CONTRIBUTIONS OF AUTHORS

DM, AP, LT, LS, and LW identified relevant evaluations to include in the review. JP identified endorsement status of RCTs and dates of endorsement for journals. LT and LS extracted data from the included evaluations. TK and SD provided own data and commented on various drafts. LT carried out the analysis. LT drafted the review with input from all authors. DM, DGA, and KFS provided conceptual and methodological supervision of the review.

DECLARATIONS OF INTEREST

Three review team members (DM, DGA, and KFS) comprise the executive of the CONSORT Group and have led the development of the CONSORT Statement since its inception in 1996. DM, KFS, and DGA are also members of the EQUATOR executive. One team member (LS) is CONSORT research staff, for which salary support is provided, in part, by the Medical Research Council, United Kingdom. Salary support for LT is provided under the Cochrane Bias Methods Group, funded by Canadian Institutes of Health Research.

SOURCES OF SUPPORT

Internal sources

- Cochrane Bias Methods Group, Canada.

Salary support for research assistance

External sources

- Medical Research Council (MRC), UK.

The CONSORT group is currently funded through a grant from the MRC. Grant no: MR/J004871/1

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

This was a complex review to complete. The review included evaluations with varied objectives, populations, and study methods. Although this was an updated review, and the protocol guidance extensive, based mainly on the quantity of included evaluations, some developments were not foreseen at the protocol stage and as such amendments were made as the evaluation progressed and documented here.

We encountered evaluations which did not report (or provide) data on CONSORT endorsement status for RCTs published in journals before and after publication of the CONSORT Statement (1996 or 2001). Due to the vast quantity of associated trials examined in included evaluations, resource limitations precluded us from obtaining the necessary data (from evaluation authors or, subsequently, included journals) in a timely and efficient manner. As such, we developed an additional comparison group (comparison 3) that allowed for these evaluations to be included subject in relation to their choice of evaluation design. Excluding comparison 3 would have omitted a substantial body of evidence and may have led to potentially misleading results. The robustness of comparison 3 is addressed in the Results and Discussion of this review.

In addition to comparison 3, we encountered data in included evaluations which could be used to form comparisons of potential 'control' groups, as described in the Objectives. Data for these comparisons were sparse and not included in this review; however they are available upon request.

During the review process, the search strategy for relevant literature as laid out in the protocol was broadened to remove the limitation of identifying literature published in only 'core clinical journals'.

During the review process, the secondary outcomes as specified in the protocol, were amended to reflect that only data on 'methodological quality' of included evaluations would be assessed. We did not collect data on overall quality since it was felt that there were no good or different measures for this outcome than those used to assess methodological quality. Validity assessment, while still carried out, was erroneously listed as a secondary outcome in the protocol and is not listed as one in this review. Validity assessment was nonetheless carried out, as described in Assessment of risk of bias in included studies.

In the protocol it was suggested that sensitivity analyses considering RCTs for which endorsement could not be strictly defined would be conducted. The protocol describes that RCTs would be excluded at the evaluation level (within evaluation). As a more efficient, and potentially more suitable, alternative we omitted the evaluations for which endorsement was not strictly compliant.

The protocol provides details of an analysis plan for assessing potential reporting biases across evaluations. Upon further consideration and consultation with statistical experts, given the type of data included in this review, standard means of assessing reporting bias were not suitable. Hence, we have not formally assessed reporting bias across the included evaluations.

Evaluations typically did not adjust for potential confounders in their analysis, and given the lack of information it was not possible for the research team to adjust for them. Due to methodological heterogeneity across evaluations, it was also not feasible to arbitrarily formulate an aggregate adjustment. As a result, we have included all results but used wider, more conservative, 99% confidence intervals, which is different from the standard 95% as detailed in the protocol.

NOTES

We were able to abstract additional data from four included evaluations which provided information on potential confounding. This included the improvement of reporting over time, or the difference in completeness of reporting in endorsing and non-endorsing journals. This information was very sparse and led to many empty forest plots; we did not feel this evidence contributed substantially to the results of the review so it is not reported here, but fully available upon request.

INDEX TERMS

Medical Subject Headings (MeSH)

Checklist [*standards]; Periodicals as Topic [*standards]; Publishing [*standards]; Randomized Controlled Trials as Topic [*standards]; Reference Standards