Reporting of harms data in RCTs: a systematic review of empirical assessments against the CONSORT harms extension

Alex Hodkinson, Jamie J Kirkham, Catrin Tudur-Smith, Carrol Gamble

ABSTRACT

Objective: To determine the standard of reporting of harms-related data, in randomised controlled trials (RCTs) according to the Consolidated Standards of Reporting Trials (CONSORT) statement extension for harms.

Design: Systematic review.

Data sources: The Cochrane library, Ovid MEDLINE, Scopus and ISI Web of Knowledge were searched for relevant literature.

Eligibility criteria for selecting studies: We included publications of studies that used the CONSORT harms extension to assess the reporting of harms in RCTs.

Results: We identified 7 studies which included between 10 and 205 RCTs. The clinical areas of the 7 studies were: hypertension (1), urology (1), epilepsy (1), complimentary medicine (2) and two not restricted to a clinical topic. Quality of the 7 studies was assessed by a risk of bias tool and was found to be variable. Adherence to the CONSORT harms criteria reported in the 7 studies was inadequate and variable across the items in the checklist. Adverse events are poorly defined, with 6 studies failing to exceed 50% adherence to the items in the checklist.

Conclusions: Readers of RCT publications need to be able to balance the trade-offs between benefits and harms of interventions. This systematic review suggests that this is compromised due to poor reporting of harms which is evident across a range of clinical areas. Improvements in quality could be achieved by wider adoption of the CONSORT harms criteria by journals reporting RCTs.

INTRODUCTION

Every healthcare intervention is associated with a risk of harmful or adverse events, that must be balanced against the potential favourable outcomes.1

The Consolidated Standards of Reporting Trials (CONSORT) statement aims to improve the quality of published reports of randomised controlled trials (RCTs) and has been widely endorsed by healthcare journals, leading to improvements in quality when used by manuscript authors and peer reviewers.2–4 However, some reports suggest that assessment and reporting of harms in clinical trials may be suboptimal.5–7

The standard CONSORT statement8 is primarily aimed at reporting the intended, usually beneficial effects of intervention(s) with only one item (item 19) devoted to unintended adverse events (harms) in the original 2001 checklist. Owing to accumulating evidence that reporting on harms-related data in RCTs was of poor quality with an imbalanced ratio of benefit–harms reporting, a CONSORT statement extension was developed in 2004 to improve harms reporting (CONSORT-harms) and to help address perceived shortcomings in measurement, analysis and reporting of harms data.9 The subsequent update of the standard CONSORT statement, published in 2010,10 now specifically refers to the additional CONSORT-harms extension but it is still unclear whether authors and journals routinely adopt the use of this extension. The aim...
of this paper is to systematically review the evidence from previously conducted empirical studies that have assessed the adequacy of harms reporting in RCTs using the CONSORT-harms extension as a benchmark.

METHODS
A protocol for the systematic review was developed by AH, CTS, CG and JJK.

Study inclusion criteria
We included published and unpublished research that evaluated the quality of harms reporting in RCTs against the CONSORT-harms recommendations. No restriction was placed on the clinical area or type of intervention studied. Excluded studies were those that assessed harms reporting using assessment criteria other than CONSORT-harms and studies that assessed harms reporting using study designs for which the CONSORT guideline was not intended (eg, observational studies).

Identification of studies
AH, CTS and CG developed the search strategy with support from an information specialist which was then implemented by AH in the following databases: the Cochrane Methodology register, Database of Abstracts of Reviews of Effects (DARE), Ovid MEDLINE, Scopus and ISI Web of Knowledge. Conference abstracts were searched for in the Web of Knowledge Conference Proceedings Citation Indexes (CPCI-S or CPCI-SSH) and the Zetoc database. An unpublished Masters dissertation involving one of the authors (JJK) was also obtained. Date filters were not used during the search strategy; our interest lies only within reviews published after 2004, with the cut-off date June 2012.

Titles and abstracts of reports identified by the search were screened by AH and full articles obtained for all potentially eligible studies. Each full article was assessed independently by two reviewers (AH and CTS) to determine eligibility.

Quality assessment
Two reviewers (AH and JJK) independently assessed the methodological quality of each study using the Cochrane Risk of Bias (RoB) tool as a guideline to cover the following aspects. Criteria were graded as low risk, high risk or unclear as indicated.

1. Were the trials included in the study a representative sample, for example, unselected journals, and reasonable time scale?
   - Low risk of bias: studies included trials from a primary search of all the available literature.
   - High risk of bias: studies were highly selective of the trials included, for example, high-impact journals or specialised journals only.
   - Unclear risk of bias: not stated how studies were selected.

2. During the data extraction of CONSORT-harms criteria, were reviewers blinded to study authors, institution, journal name and sponsors?
   - Low risk of bias: reviewers were blinded.
   - High risk of bias: reviewers were not blinded.
   - Unclear risk of bias: not stated.

3. Is there evidence of selective outcome reporting in the study (ie, were all CONSORT-harms recommendations considered and if not were suitable reasons provided)?
   - Low risk of bias: studies that considered all CONSORT-harms criteria or reasons for excluding specific criteria were transparent and justified.
   - High risk of bias: studies did not consider all CONSORT-harms criteria.
   - Unclear risk of bias: unclear whether all CONSORT-harms criteria were considered.

4. Did more than one reviewer assess the CONSORT-harms criteria for each primary RCT, with a description of how agreement was achieved?
   - Low risk of bias: data extraction was completed independently by two people or reasonable attempts were made to maximise data extraction reliability.
   - High risk of bias: data extraction not completed independently by two people.
   - Unclear risk of bias: not stated.

Data collection and extraction
Two reviewers (AH and JJK) independently extracted the data and any discrepancies were resolved through a consensus discussion with a third reviewer (CTS). Data extraction included:
- Study characteristics: inclusion criteria including clinical area, types of interventions, databases or journals searched within the study and any search date restrictions.
- Sample size (defined by the number of RCT reports assessed for reporting quality).
- Reporting quality: inclusion of any of the 10 recommendations from the 2004 CONSORT-harms checklist (table 1 and supplemental data: see CONSORT plots).

Lead authors were contacted through email with any queries relating to the quality assessment or data extraction.

Data analysis and presentation
For each study, the percentage of included RCTs that satisfied each CONSORT-harms recommendation is presented with 95% CIs. Some studies had presented data for individual items described within each of the 10 criteria rather than overall data. These are presented as such in tables with footnotes to provide further explanation. Forest plots were used to graphically depict the levels of adherence to the CONSORT harms recommendations so that readers can easily discern the extent of compliance and heterogeneity between studies with the I² statistic (included as supplementary material online). We refrained from statistically combining results from the different studies due to the differences in their study characteristics. In accordance
with the Cochrane Handbook, $I^2$ statistics were interpreted as (0–40%, might not be important; 30–60%, may represent moderate heterogeneity; 50–90% may represent substantial heterogeneity; 75–100%, considerable heterogeneity).\textsuperscript{13}

**RESULTS**

The search strategy identified 5083 potentially eligible study cohorts from which seven studies assessing the quality of reporting across almost 800 RCTs were included (figure 1).
Five studies,\textsuperscript{14-18} (with the study\textsuperscript{19} recently published) contained trials focusing on specific clinical areas with two\textsuperscript{20-21} covering multiple clinical areas (table 2). Four studies\textsuperscript{14 17 20 21} included trials using drug interventions, one comparing acupuncture\textsuperscript{18} and another alternative complementary medicines,\textsuperscript{16} the interventions were unclear in one study.\textsuperscript{15} MEDLINE was used by four of the studies\textsuperscript{17 18 20 21} to identify the relevant literature; three\textsuperscript{14 16 17} used the Cochrane database of RCTs and three\textsuperscript{15-17} searched specialised-journal databases. The date restrictions used in the search strategy of each study ranged from a 1 year period up to a 9 years span. The

<table>
<thead>
<tr>
<th>Review characteristic</th>
<th>Bagul and Kirkham\textsuperscript{14}</th>
<th>Breau et al\textsuperscript{15}</th>
<th>Turner et al\textsuperscript{16}</th>
<th>Shukralla et al\textsuperscript{17}</th>
<th>Capili et al\textsuperscript{18}</th>
<th>Pitrou et al\textsuperscript{20}</th>
<th>Haidich et al\textsuperscript{21}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinical area</td>
<td>Hypertension</td>
<td>Urology</td>
<td>Acupuncture therapies and other complementary alternative medicines (CAM)</td>
<td>Acupuncture, massage therapies and herbal medicines.</td>
<td>Epilepsy</td>
<td>Acupuncture</td>
<td>Mixed</td>
</tr>
<tr>
<td>Type of intervention(s)</td>
<td>Drug interventions for hypertension</td>
<td>Unclear</td>
<td>Drug interventions for epilepsy</td>
<td>Drug interventions (excluding studies that evaluated acupressure, laser acupuncture, and auricular (acupuncture)</td>
<td></td>
<td></td>
<td>Drug interventions</td>
</tr>
<tr>
<td>Journals/ databases searched</td>
<td>Cochrane Collaboration’s CAM Field Specialised Regist Trials</td>
<td>Cochrane Libr; Epilepsy Group Trial Registry</td>
<td>MEDLINE; Cochrane Libr; Epilepsy Group Trial Registry</td>
<td>MEDLINE; Allied Complement Med; Cumulative Index Nurs Allied Health Lit; Evid Based Med Rev</td>
<td>MEDLINE via PubMed (NEJM, Lancet, JAMA;BMJ; Ann Intern Med; PLosS Med)</td>
<td>MEDLINE (Ann Intern Med, BMJ, JAMA, Lancet, NEJM)</td>
<td></td>
</tr>
<tr>
<td>Number of RCTs included in the review</td>
<td>41</td>
<td>152</td>
<td>205</td>
<td>152</td>
<td>10</td>
<td>133</td>
<td>102</td>
</tr>
<tr>
<td>Inclusion criteria</td>
<td>Randomised controlled hypertension trials comparing two parallel arms reported in standard CONSORT endorsing journals</td>
<td>RCTs of therapeutic interventions published in the three selected journals for the 2 years.</td>
<td>All Cochrane Complementary Medicine RCTs pertaining to 15 CAM intervention categories.</td>
<td>RCTs comparing AEDs (antiepileptic drugs); RCT patient population with epilepsy; RCTs published in English.</td>
<td>Studies published in the English language, acupuncture for pain reduction, a method for evaluating level of pain, and randomised allocation to treatment group.</td>
<td>Articles were included if the study was identified as an RCT with two parallel arms (in selected journals).</td>
<td>Published RCTs assessing drugs in the selected journals for the 2 years.</td>
</tr>
</tbody>
</table>
studies were published after 2008, 4 years after the release of the harms extension with three,15 17 21 of them including trials that had been published before the publication of CONSORT-harms. Five studies14–17 20 excluded trials published in a non-English language.

**Risk of bias**

Lead authors were contacted by email with any queries relating to the quality of their study, or CONSORT criteria; however, two authors failed to respond.15 18 The risk of bias for the seven included studies, assessed across four domains, is summarised in table 3. Six studies14 15 17 18 20 21 were classified as high risk for bias for at least one domain with two of these studies14 20 classified as high risk for three domains. Four studies15 14 20 21 did not include trials of a representative sample targeting specific journals rather than a database search. Blinding of assessors was only implemented in two studies15 16 with one unclear.18 Most studies used all the CONSORT-harms criteria with the exception of the subgroup analysis item; one study16 however, discarded the use of recommendation eight, since it was captured elsewhere within the data extraction, and recommendation 10, which was considered too vague to assess with any objectivity. Reporting of the assessment within three15 18 20 of the seven identified studies was unclear and authors were contacted. The authors did not respond for two studies15 17 and in another study20 a response was received but some details remained unclear. Six studies15–18 20 21 had used two independent data extractors while one study14 had not and was classified as high risk of bias for this domain.

**CONSORT harms recommendations**

Results extracted for the CONSORT-harms criteria (table 4) demonstrate variability in the level of adherence to items. Heterogeneity is highlighted by the individual Forest plots where inflated I² values of over 85% are represented for all recommendations, denoting considerable heterogeneity.

Of the six studies that assess inclusion of harms in the title and abstract of their included RCTs, three,16 20 21 reported compliance in over 70% of RCTs, but three14–16 reported compliance in less than 50% of RCTs. The introduction section of the included RCTs reflect an imbalance in the reporting benefit–harms, with one study16 reporting that less than 5% of RCTs had mentioned harms in the introduction, and one study17 reporting more than 70% of its included RCTs has satisfied this criteria.

The definition of adverse events in reports is unsatisfactory with most studies14–16 17 20 indicating that fewer than 20% of RCTs satisfy these criteria adequately. The collection of harms-related information is described by more than 80% of RCTs in two studies,20 21 but this high level is not consistent across the other five studies with one study14 suggesting that as few as 10% of RCTs had provided an adequate description. The analysis and coding of adverse events is poorly described, with less than 50% of RCTs satisfying this criteria across six studies,13–18 20 21 with one of these studies13 indicating that none of the RCTs had provided an adequate description. The reporting of participant withdrawals due to harms was inconsistent with two studies15 16 suggesting infrequent reporting (less than 40% of RCTs

---

**Table 3 Risk of bias assessment**

<table>
<thead>
<tr>
<th>Risk of bias criteria</th>
<th>Bagul and Kirkham14</th>
<th>Breau et al15</th>
<th>Turner et al16</th>
<th>Shukralla et al17</th>
<th>Capilli et al18</th>
<th>Pitrou et al20</th>
<th>Haidich et al21</th>
</tr>
</thead>
<tbody>
<tr>
<td>Representativeness of sample of trials (low if trials were searched across unselected journals and across a reasonable time period)</td>
<td>High</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>Blinding of reviewers during CONSORT-harms data extraction (low if reviewers blinded to study authors, institution, journal name and sponsors)</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td>Unclear</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>Selective outcome reporting (low if all CONSORT-harms criteria assessed)</td>
<td>Low*</td>
<td>Low*</td>
<td>Low*†</td>
<td>Low*</td>
<td>High</td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>Reliability of data extraction (low if more than one reviewer assessed the CONSORT harms criteria for each review that was undertaken, with a description of how agreement was achieved)</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
</tr>
</tbody>
</table>

*Recommendation 9 was not included in these studies as subgroup analysis was either not reported in any of the included studies or considered to be irrelevant for the therapeutic area being investigated.

†Authors response: ‘Recommendation 8 has been captured elsewhere in data extraction, to report this item would be to duplicate information presented’.

*Recommendation 10 was considered too vague to assess with any objectivity so we decided to leave this item, especially given that some of our primary outcomes were already reasonably subjective.*
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>CONSORT Recommendation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% of trials (95% CI) that adhered to each recommendation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1) Title &amp; Abstract</td>
<td>20 (9, 35)</td>
<td>12 (6, 20)</td>
<td>21 (16, 27)</td>
<td>88 (81, 92)</td>
<td>NR</td>
<td>71 (63, 79)</td>
<td>76 (67, 84)</td>
</tr>
<tr>
<td>1i)</td>
<td></td>
<td>12 (6, 20)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1ii)</td>
<td></td>
<td>64 (53, 74)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2) Introduction</td>
<td>34 (20, 51)</td>
<td>54 (43, 65)</td>
<td>4 (2, 8)</td>
<td>74 (67, 81)</td>
<td>NR</td>
<td>NR</td>
<td>48 (38, 58)</td>
</tr>
<tr>
<td>(3) Definition of adverse events</td>
<td>0 (0, 9)</td>
<td>15 (8, 24)</td>
<td>6 (3, 11)</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>(4) Collection of harms data</td>
<td>10 (3, 23)</td>
<td>4i) 22 (14, 32)</td>
<td>17 (12, 22)</td>
<td>20 (3, 56)</td>
<td>89 (82, 94)</td>
<td>81 (74, 89)</td>
<td></td>
</tr>
<tr>
<td>4i)</td>
<td></td>
<td>4ii) 6 (2, 13)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4iii)</td>
<td></td>
<td>0 (0, 4)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(5) Analysis of harms</td>
<td>0 (0, 9)</td>
<td>76 (66, 84)</td>
<td>6 (3, 10)</td>
<td>20 (3, 56)</td>
<td>12 (7, 19)</td>
<td>44 (34, 54)</td>
<td></td>
</tr>
<tr>
<td>(6) Withdrawals</td>
<td>51 (35, 67)</td>
<td>35 (25, 45)</td>
<td>30 (24, 37)</td>
<td>70 (35, 93)</td>
<td>53 (44, 61)</td>
<td>59 (50, 69)</td>
<td></td>
</tr>
<tr>
<td>(7) Number of patients analysed</td>
<td>17 (7, 32)</td>
<td>35 (25, 45)</td>
<td>18 (13, 24)</td>
<td>NR</td>
<td>84 (77, 90)</td>
<td>74 (64, 82)</td>
<td></td>
</tr>
<tr>
<td>(8) Results for each adverse event</td>
<td>39 (24, 56)</td>
<td>8i) 0 (0, 4)</td>
<td>–</td>
<td>NR</td>
<td>73 (65, 80)</td>
<td>89 (82, 95)</td>
<td></td>
</tr>
<tr>
<td>8ii)</td>
<td></td>
<td>28 (19, 38)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(9) Subgroup Analysis</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>NR</td>
<td>NR</td>
<td>53 (43, 63)</td>
<td></td>
</tr>
<tr>
<td>(10) Balanced discussion</td>
<td>5 (1, 17)</td>
<td>10i) 61 (50, 71)</td>
<td>–</td>
<td>10a) 68 (60, 76)</td>
<td>NR</td>
<td>NR</td>
<td>83 (76, 91)</td>
</tr>
</tbody>
</table>

NR: Not reported in manuscript, and no response from authors when contacted.

- Author detailed reasons for not reporting the recommendation.
- 1) (i) Harm, safety or similar term used in title; (ii) Harm addressed in abstract.
- 4) (i) When harm information was collected; (ii) Methods to attribute harm to intervention; (iii) Stopping rules.
- 8) (i) Effect sizes for harms; (ii) Stratified serious and minor harms.
- 10) (i) Interpret harm outcome; (ii) Discuss generalizability; (iii) Discuss current evidence.
- 3) (a) Definition of AE; (b) All or selected sample; (c) Treatment Emergent AE; (d) Validated instrument; (e) Validated dictionary.
- 4) (a) Mode of AE collection; (b) Timing of AE; (c) Details of attribution.
- 5) (a) Details of presentation and analysis; (b) Handling of recurrent AE.
- 6) (a) Early or late withdrawals; (b) Serious AEs or death.
- 7) (a) Provide denominators for AEs; (b) Provide definitions used for analysis set.
- 8) (a) Same analysis set used for efficacy and safety; (b) Results presented separately; (c) Severity and grading of AEs; (d) Provide both number of AEs and number of patients with AEs.
- 10) (a) Discusses prior AE data; (b) Discussion is balanced; (c) Discusses limitations.
had mentioned withdrawals), three studies\(^1\)\(^3\)\(^2\)\(^1\)\(^7\)\(^2\)\(^1\)\(^5\)\(^6\) suggesting occasional reporting (50–60% of RCTs had mentioned withdrawals) and two studies\(^1\)\(^7\) suggesting that reporting of withdrawals was quite common (approximately 70% of RCTs had mentioned withdrawals).

When providing the denominators within trial reports, the results were also varied across studies, with three\(^1\)\(^7\)\(^2\)\(^1\)\(^5\)\(^6\)\(^2\)\(^1\)\(^2\)\(^1\) studies identifying more than 70% of trials that satisfied this criterion, but two studies\(^1\)\(^3\)\(^5\)\(^6\)\(^2\)\(^1\)\(^2\)\(^1\)\(^5\)\(^6\) identifying less than 20% adherence. The risk and severity grading of adverse events, is detailed in more than 70% of trial across two studies,\(^1\)\(^2\)\(^1\)\(^5\)\(^6\)\(^2\)\(^1\)\(^2\)\(^1\)\(^5\)\(^6\) but the reporting is inadequate in three studies.\(^1\)\(^3\)\(^5\)\(^6\)\(^2\)\(^1\)\(^2\)\(^1\)\(^5\)\(^6\) An assessment of reporting of harms within subgroup analysis was only carried out within one study.\(^2\)\(^1\)

Four studies\(^4\)\(^5\)\(^6\)\(^2\)\(^1\)\(^2\)\(^1\)\(^5\)\(^6\) assessed their included RCTs for a balanced report on the benefits and harms within their discussion: one study\(^5\)\(^6\) identified a very low percentage (<10%), two studies\(^4\)\(^5\)\(^6\)\(^2\)\(^1\)\(^2\)\(^1\)\(^5\)\(^6\) identified a moderate percentage (approximately 60%), and one study\(^2\)\(^1\) identified a high percentage (over 80%) of trials that met this criterion.

**DISCUSSION**

**Summary of findings**

This is the first study to systematically review empirical studies assessing the quality of reporting according to the CONSORT-harms guideline.\(^3\)\(^7\)\(^2\)\(^1\)\(^5\)\(^6\) Data were extracted from seven studies that had each assessed the quality of reporting across almost 800 RCTs from a range of clinical specialities. Eight years have now passed since the release of the harms extension, allowing adequate time for the guideline implementation. This review highlights that the reporting of harms in RCTs is inconsistent, and at times very poor. Heterogeneity is easily discerned between studies for each recommendation. Further adherence to the CONSORT-harms is needed.

The standard CONSORT is well established in health research with building evidence to support the use of the guideline.\(^5\)\(^6\) Currently the standard CONSORT is endorsed by over 50% of the core medical journals in the abridged Index Medicus on PubMed.\(^2\)\(^2\) In a review\(^2\)\(^3\)\(^2\)\(^1\) of 116 health research journals, 41 provided online instructions to authors. Almost half (19/41 (46%)) mentioned the standard CONSORT guideline but none referred to the CONSORT extension for harms.

**Strengths and weaknesses of the study**

In this study we have focused on assessing reporting according to the CONSORT-harms criteria only. The included studies contained trials reported prior to the publication of the CONSORT-harms guideline. However, we have not assessed changes in reporting over time. Nevertheless, our results support those from previous studies\(^3\)\(^4\) that used various guidelines published before the release of the CONSORT-harms extension. This study should be regarded as a reflection of reporting standards in general rather than an assessment of adherence to the CONSORT-harms extension.

This review was strengthened by its assessment of quality of the included studies across four key domains. With the guidance of the Cochrane review\(^1\)\(^2\)\(^1\)\(^3\) we have designed a RoB tool to perform a generalisable assessment of the included studies. In this assessment only the one study,\(^1\)\(^5\) demonstrated low RoB across all four of the assessment criteria. No restriction was placed on the inclusion criteria of the included studies such that the time span and clinical areas of their included studies varied. While this is a strength in terms of generalisability of results, it may also be considered as a level of heterogeneity that cannot be explored due to the limited number of studies.

**Conclusions and implications**

Complete and accurate reporting is essential to guide decisions on advances in medical interventions. The responsibility to ensure greater balance between reporting of both benefits and harms lies with authors of research and journals publishing that research. We recognised that journals have limited space for the reporting of all outcomes which can lead to selective outcomes reporting.\(^2\)\(^4\)\(^2\)\(^5\) We recommend the use of supplementary online tables to help summarise key results on harms.

Further dissemination strategies should be used to ensure that trial journal editors and trial investigators are aware of the importance of adequate reporting of harms-related data in RCTs. As it stands, it is unclear as to whether the problem of the poor reporting of harms data in trial publications is a result of the lack of awareness of the CONSORT for harms statement, or journals and peer reviewers not implementing this guideline. The most effective strategy would follow that of the CONSORT statement with the extension for harms comprehensively incorporated in journal requirements along with clear instructions to peer reviewers for guidelines of acceptance.

**Acknowledgements**

The authors are grateful to Su Golder and Fiona Beyer York University for expanding on the item words used within the databases and Alison Beaumont University of Liverpool, for recommending databases to search conference abstracts.

**Contributors**

AH and CTS carried out all screening of literature; AH and JJK extracted data; AH, JJK, CTS and CG interpreted results and drafted the manuscript.

**Funding**

This work was supported by the award of a Capacity-Building Studentship to AH from the Medical Research Council (MRC) (grant number G000397 – 1/1) North West Hub for Trials Methodology Research, UK (grant number G08000792).

**Competing interests**

None.

**Patient consent**

Obtained.

**Provenance and peer review**

Not commissioned; externally peer reviewed.

**Data sharing statement**

Data extraction form and protocol available on request from ahoddy@liverpool.ac.uk.

**Open Access**

This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 3.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/3.0/.
REFERENCES