Value of Flow Diagrams in Reports of Randomized Controlled Trials

Matthias Egger, MD
Peter Juni, MD
Christopher Bartlett, PhD
for the CONSORT Group

The Consolidated Standards for Reporting of Trials (CONSORT) criteria were published in 1996 to assist authors in reporting randomized controlled trials (RCTs). The CONSORT statement consists of a checklist for items in the methods, results, and discussion sections of a trial report and a diagram to show the flow of participants through the various stages of the trial (Figure A). The CONSORT statement has been endorsed by an increasing number of journals,2-5 but it is unclear whether the use of CONSORT has improved the quality of reports of RCTs. We examined to what extent 5 leading medical journals adopted flow diagrams, analyzed the information contributed by these diagrams, and assessed completeness of reporting overall. The results guided our attempt to improve the design of the flow diagram (Figure B).

Methods
One of us (C.B.) searched each issue published in 1998 of the Annals of Internal Medicine (AIM), BMJ, JAMA, The Lancet, and The New England Journal of Medicine (NEJM) for published reports of RCTs. A study was defined as an RCT if the assignment of participants to interventions was described as randomized by words such as randomly, random, and randomization. Two of us (M.E., P.J.) independently extracted data on the characteristics of reports and examined whether a flow diagram was included. Articles were examined in random sequence with 1 assessor examining trials in the opposite order. We searched each journal’s Web site for diagrams that were published electronically but did not appear in print.

We examined whether 6 counts were provided in diagrams: (1) number of patients assessed for eligibility, (2) number found to be eligible, (3) number randomized (per group), (4) number who received allocated intervention (per group), (5) number who were lost to follow-up (per group), and (6) number included in the main analysis of primary outcomes (per group). In a separate step, we assessed whether these counts were provided either in a flow diagram or anywhere else in the article.

Each article was read independently by 2 of us in an incomplete randomized Latin square design, and inter-rater reliability6 for assessment of

information in flow diagrams and overall was determined. Disagreements were resolved by consensus. We compared overall completeness of reporting between journals using \( \chi^2 \) tests and used logistic regression to assess the association of flow diagrams with completeness of reporting.

**RESULTS**

The hand search of these journals identified 290 articles. We excluded 1 quasi-randomized trial, 10 cluster trials, and 9 reports for which the focus was not on randomized comparisons. Our study sample thus consisted of 270 reports of RCTs. Most trials were of parallel group design (256 [94.8%]) and evaluated pharmacological interventions (173 [64.1%]). A total of 139 reports (51.5%) included a flow diagram but this varied widely among journals (Table 1). Five (29.4%) of 17 short reports included a flow diagram. No additional diagrams were found on the journals’ Web sites. Interrater reliability \(^6\) for the assessment of information provided in flow diagrams was good (\( \kappa \) values ranging from 0.60 to 0.81), but showed more disagreement for assessment of overall completeness of reporting (\( \kappa \), 0.21-0.71).

Characteristics of flow diagrams were similar across journals. Sixty-five diagrams (46.8%) included the number of participants assessed for eligibility, and 94

---

**Figure.** Original and Proposed Revised Versions of CONSORT Flow Chart

**Table 1.** Information on Participant Flow Provided in 270 Reports of Randomized Controlled Trials Published in 5 Journals in 1998 \(^*\)

<table>
<thead>
<tr>
<th>Journals</th>
<th>Annals of Internal Medicine (n = 19)</th>
<th>BMJ (n = 42)</th>
<th>JAMA (n = 45)</th>
<th>The Lancet (n = 81)</th>
<th>The New England Journal of Medicine (n = 83)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. (%) of articles including flow diagram</td>
<td>4 (21.0)</td>
<td>16 (38.1)</td>
<td>36 (80.0)</td>
<td>76 (93.8)</td>
<td>7 (8.4)</td>
</tr>
<tr>
<td>No. (%) of articles reporting numbers of participants (in flow diagram or text)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assessed for eligibility (overall)</td>
<td>9 (47.4)</td>
<td>21 (50.0)</td>
<td>25 (55.6)</td>
<td>38 (46.9)</td>
<td>30 (36.1)</td>
</tr>
<tr>
<td>Found to be eligible (overall)</td>
<td>11 (57.9)</td>
<td>27 (64.3)</td>
<td>36 (80.0)</td>
<td>54 (66.7)</td>
<td>36 (43.4)</td>
</tr>
<tr>
<td>Randomized (per group)</td>
<td>15 (78.9)</td>
<td>39 (29.2)</td>
<td>43 (95.6)</td>
<td>76 (93.8)</td>
<td>67 (80.7)</td>
</tr>
<tr>
<td>Received allocated intervention (per group)</td>
<td>10 (52.6)</td>
<td>27 (64.3)</td>
<td>27 (60.0)</td>
<td>56 (69.1)</td>
<td>28 (33.7)</td>
</tr>
<tr>
<td>Lost to follow-up (per group)</td>
<td>8 (42.1)</td>
<td>26 (61.9)</td>
<td>34 (75.6)</td>
<td>55 (67.9)</td>
<td>27 (32.5)</td>
</tr>
<tr>
<td>Excluded from analysis (per group)</td>
<td>13 (68.4)</td>
<td>36 (85.7)</td>
<td>36 (80.0)</td>
<td>71 (87.7)</td>
<td>56 (67.5)</td>
</tr>
<tr>
<td>Included in main analysis (per group)</td>
<td>16 (84.2)</td>
<td>39 (29.2)</td>
<td>37 (82.2)</td>
<td>76 (93.8)</td>
<td>74 (89.2)</td>
</tr>
</tbody>
</table>

\(^*\) Differences were statistically significant for inclusion of flow diagrams \((P < .001)\) and reporting on the number found eligible \((P = .001)\) the number randomized \((P = .02)\), the number who received allocated intervention \((P < .001)\), the number lost to follow-up \((P < .001)\), and the number excluded from analysis \((P = .01)\). Probability from \( \chi^2 \) tests.

©2001 American Medical Association. All rights reserved.
(67.6%) reported the number found to be eligible. The number assigned to each study group was reported in 129 flow diagrams (92.8%), and 73 (52.5%) indicated whether interventions had been received as allocated. Most flow diagrams (114 [82.0%]) indicated how many patients in each group were lost to follow-up, but only 32 (23.0%) reported the number included in the analysis. Overall reporting was more complete in journals that publish flow diagrams more frequently (Table 1). However, there is clearly room for improvement in all journals, in particular with regard to the number receiving the allocated intervention and the number lost to follow-up. Independent of journal, reporting tended to be less complete if the trial was published as a short report. In logistic regression analysis, completeness of reporting was associated with the publication of a flow diagram (Table 2). In multivariable analyses, there was little evidence for confounding by journal, and differences among journals were largely explained by the publication or nonpublication of flow diagrams.

**COMMENT**

A diagram showing the flow of participants from enrollment to analysis is an important element of the CONSORT standards for the reporting of clinical trials. Flow diagrams provide an aid to trialists when writing trial results and assist readers in the critical appraisal of the internal and external validity of a trial. We found that the information on the progress of participants through the trial was more complete in journals that frequently publish flow diagrams and that the publication of these diagrams is associated with better reporting in individual articles. Our results are consistent with those of Moher and colleagues, who compared reports published in the same journals before and after the CONSORT statement became available. The proportion of articles that included a flow diagram varied widely across journals and was lower in short reports. Flow diagrams take up precious journal space, and editors may sometimes feel that this space is better used otherwise. Our findings provide strong support for the idea that RCTs always should be published as full articles including a flow diagram.

The shortcomings of the flow diagram template recommended in 1996 may be another reason some editors are reluctant to publish flow diagrams. Meinert pointed out that the terms used in the 1996 CONSORT statement lacked clarity and that the information presented in the flow diagram was incomplete. Our results indicate that there were problems with both the original design of the flow diagram and its implementation by authors. For example, most flow diagrams provided the number of individuals randomized, although this count was not explicitly requested. Conversely, only about half of flow diagrams included the number of participants who actually received treatments as allocated, an item included in the original template. The number of participants included in the main analysis was not an item in the recommended flow diagram, and this number was included in only a few diagrams (23.0%). This finding is of concern because the latter count is essential for appraising whether a trial has been analyzed by intention to treat. A recent study found that intention-to-treat approaches are often inadequately described and inadequately applied. The number of persons assessed for eligibility was also frequently missing. Although this number is not relevant for assessing the internal validity of a trial, it is useful to estimate whether trial participants were likely to be representative of all patients seen.11

The CONSORT criteria are an evolving tool designed to help improve the quality of reporting of RCTs. Based on the findings of our present study and earlier criticism of CONSORT, we submit that the flow diagram is useful but in need of revision. We revised the template in collaboration with the CONSORT Group (Figure). Three tiers of boxes cover the enrollment phase and 1 box each is set apart for allocation of interventions, follow-up, and analysis. At enrollment a distinction is made between the exclusion of persons not meeting the criteria specified in the protocol and persons excluded for other reasons. The revised flow diagram explicitly includes the number of participants allocated to each group, the number initially receiving the intervention as assigned, and, if applicable, the reasons why some participants did not receive allocated interventions. Information on follow-up includes the number of participants lost to follow-up and the number of patients who stopped interventions along with the reasons why this occurred. The bottom tier of boxes includes the number of patients included in the main analysis and, if applicable, the reasons why some patients were excluded. We trust that this template could be improved further and would appreciate comments from readers through the CONSORT Web site.13

**Table 2. Association of Presence of a Flow Diagram With Completeness of Reporting**

<table>
<thead>
<tr>
<th>Reporting on No. of Participants</th>
<th>Crude Adjusted for Journal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assessed for eligibility (overall)</td>
<td>1.79 (1.10-2.91) 1.90 (0.93-3.87)</td>
</tr>
<tr>
<td>Found to be eligible (overall)</td>
<td>3.28 (1.97-5.48) 3.19 (1.47-6.92)</td>
</tr>
<tr>
<td>Randomized (per group)</td>
<td>3.31 (1.42-7.72) 1.87 (0.53-6.52)</td>
</tr>
<tr>
<td>Received allocated intervention (per group)</td>
<td>2.79 (1.70-4.58) 1.75 (0.85-3.60)</td>
</tr>
<tr>
<td>Lost to follow-up (per group)</td>
<td>5.49 (2.34-9.28) 5.01 (2.33-11.1)</td>
</tr>
<tr>
<td>Excluded from analysis (per group)</td>
<td>3.27 (1.75-6.12) 3.10 (1.22-7.86)</td>
</tr>
<tr>
<td>Included in main analysis (per group)</td>
<td>2.45 (1.07-5.63) 4.96 (1.51-16.3)</td>
</tr>
</tbody>
</table>

*Results from univariate and bivariate logistic regression analyses. Odds ratios (ORs) >1 indicate that articles with flow diagrams are more likely to report the number of participants than articles without flow diagrams. CI indicates confidence interval.*
Author Contributions: Study concept and design: Egger, Juni.
Acquisition of data: Egger, Juni, Bartlett.
Analysis and interpretation of data: Egger, Juni.
Drafting of the manuscript: Egger.
Critical revision of the manuscript for important intellectual content: Egger, Juni, Bartlett.
Statistical expertise: Egger, Juni.
Administrative, technical, or material support: Egger, Bartlett.
Study supervision: Egger.

Funding/Support: The Ottawa meeting was made possible by funding from Abbott Laboratories, American College of Physicians, GlaxoWellcome, National Library of Medicine, The Lancet, Merck, and the Medical Research Council of Canada. Drs Juni and Egger are supported by the Swiss National Science Foundation and the United Kingdom’s Medical Research Council Health Services Research Collaboration.

Acknowledgment: We thank all members of the following groups: Standards for Reporting Trials (SORT), Asilomar Working Group on Recommendations for Reporting of Clinical Trials in the Biomedical Literature, and the CONSORT Group for their support and encouragement in helping to bring the CONSORT statement to its present form. We are grateful to Frank Davidoff, MD, Peter Gøtzsche, MD, Sylvan Green, PhD, Barbara Hawkins, PhD, Richard Horton, MD, Terry Klassen, MD, Tom Lang, MA, Curt Meinert, PhD, David Moher, MSc, and Tim Peters, PhD, for helpful comments on an earlier version of the manuscript. The work reported herein was prompted by discussions during the second meeting of the CONSORT Group in Ottawa, Ontario, in May 1999. The members of the CONSORT Group who participated in the Ottawa meeting were Doug Altman, PhD, Jesse Berlin, PhD, Frank Davidoff, MD, Matthias Egger, MD, Diana Elbourne, PhD, Peter Gøtzsche, MD, Sylvan Green, PhD, Barbara Hawkins, PhD, Richard Horton, MD, Terry Klassen, MD, Tom Lang, MA, Curt Meinert, PhD, David Moher, MSc, Drummond Rennie, MD, Ken Schulz, PhD, and Bruce Squires, MD.

REFERENCES

Scientific activity is the only one which is obviously and undoubtedly cumulative and progressive.
—George Sarton (1884-1956)