Methodological quality and reporting of ethical requirements in phase III cancer trials

J J Tuech, P Pessaux, G Moutel, V Thoma, S Schraub, C Herve

Background: The approval of a research ethics committee (REC) and obtaining informed consent from patients (ICP) could be considered the main issues in the ethics of research with human beings. The aim of this study was to assess both methodological quality and ethical quality, and also to assess the relationship between these two qualities in randomised phase III cancer trials.

METHODS

Methodological quality (Jadad score) and ethical quality (Berdeu score) were assessed for all randomised controlled trials (RCTs) published in 10 international journals between 1999 and 2001 (n = 231). The mean Jadad score was 9.86 ± 1.117. The methodological quality was poor in 75 RCTs (Jadad score <9). The mean Berdeu score was 0.42 ± 0.133. The mean ethical quality score for poor methodological quality RCTs (n = 75) was 0.39 ± 0.133; it was 0.43 ± 0.133 for good (n = 156) methodological quality RCTs (p = 0.07). There was improvement in ethical quality according to the year of commencement of the trials (p < 0.001). There was no correlation between methodological quality and the number of participating patients (R2 = 0.003, p = 0.78), between ethical quality and the number of participating patients (R2 = 0.003, p = 0.76), or between ethical quality and methodological quality (R2 = 0.012, p = 0.1). ICP and REC approval were not obtained for 21 and 77 trials respectively.

Conclusion: The association between methodological quality and the reporting of ethical requirements probably reflects the respect shown for patients during the whole research process. These results suggest that closer attention to the conduct of clinical research, as well as the reporting of its ethical aspects, is needed.
7. Was the method used to assess adverse effects described?
8. Were the methods of statistical analysis described?

Items are scored as follows:

- Give either a score of 1 point for each “yes” or 0 points for each “no”. There are no in-between marks.
- Give 1 additional point if, for question 1, the method to generate the sequence of randomisation was described and was appropriate (table of random numbers, computer generated, etc.) and/or if, for question 2, the method of double blinding was described and was appropriate (identical placebo, active placebo, dummy, etc.).
- Deduct 1 point if, for question 1, the method to generate the sequence of randomisation was described and was inappropriate (patients were allocated alternately, or according to date of birth, hospital number, etc.) and/or if, for question 2, the study was described as double blind but the method of blinded was inappropriate (for example, comparison of tablet versus injection with no double dummy).

Ethical quality was evaluated using the Berdeu 10-item scale. Each item scores one point; the score was obtained by dividing the sum of the individual scores by the maximum possible score, expressed as a decimal number ranging from 0 to 1.

Statistical analysis
Comparison of the qualitative variables was carried out by means of the $\chi^2$ test and with Fisher’s exact test. Comparison of the quantitative values was carried out using Student’s t-test. Multiple comparisons of the quantitative values were by analysis of variance. The year of commencement of the trial was redetermined as follows: before 1990, from 1990 to 1995, and after 1995. Correlations between two variables were carried out by a linear regression expressed by its straight regression line and its coefficient of correlation $R$ squared. A difference was considered as significant when the p value was less than 0.05 (p values are given as two-tailed values).

Ethical aspects
This study was not approved by a REC, nor did we request Ethical aspects

0.05 (p values are given as two-tailed values).

line and its coefficient of correlation $R$ squared. A difference out by a linear regression expressed by its straight regression after 1995. Correlations between two variables were carried out using Student's test and with Fisher's exact test. Comparison of the quantitative values was carried out using Student's t-test. Multiple comparisons of the quantitative values were by analysis of variance. The year of commencement of the trial was redefined as follows: before 1990, from 1990 to 1995, and after 1995. Correlations between two variables were carried out by a linear regression expressed by its straight regression line and its coefficient of correlation $R$ squared. A difference was considered as significant when the p value was less than 0.05 (p values are given as two-tailed values).

RESULTS
Between January 1999 and December 2001, 259 phase III cancer trials were published in the journals studied. Twenty-eight trials had already been published; the remaining 231 trials are the subject of this study.

The methodological quality score (Jadad scale) ranged from 6 to 13, with a mean of 9.86 ± 1.117. Only 24 articles, in which the study was described as double blind and could therefore score an extra 2 points, were eligible for the maximum possible score of 13, and just 7 achieved this score. The remainder were: 12 points, $n = 8$; 11 points, $n = 4$; and 10 points, $n = 5$. The maximum possible score for the remaining 207 trials was 11. For 75 trials (32.5%) the methodological quality was insufficient (score ≤ 9; table 1). Fifty-five obtained the maximum possible score: 13 points, $n = 7$ for double blind trials; and 11 points, $n = 48$ for non-double blind trials.

The score for ethical quality ranged from 0.11 to 0.77 (mean 0.42 ± 0.133). The item by item frequency of endorsement (Berdeu scale) is summarised in table 2. This analysis showed that criteria 1, 2, 4, and 7 were respected in more than 60% of cases.

A total of 154 trial reports (66.7%) stated that a REC had approved the research and 210 (90.9%) reported that ICP had been requested from the participants.

There was no significant difference in methodological quality ($p = 0.38$) according to the year of commencement of the trial.

There was improvement in ethical quality according to the year of commencement (overall comparisons, $p < 0.001$), particularly when comparing trials that started before 1990 with those starting in the period 1990–1995 ($p = 0.0006$), or after 1995 ($p < 0.0001$). There was no statistical difference on comparing trials that started during 1990–1995 with those that started after 1995 ($p = 0.051$).

There was no statistical difference in methodological quality ($p = 0.15$) or ethical quality ($p = 0.34$) according to the journal in which reports were published.

No correlations were shown between numbers of participating patients and either methodological quality ($R^2 = 0.003$, $p = 0.78$) or ethical score ($R^2 = 0.003$, $p = 0.76$).

There was no correlation between ethical score and methodological score ($R^2 = 0.012$, $p = 0.1$). The mean ethical quality for the reports ($n = 75$) that had insufficient methodological quality was 0.39 ± 0.133 (0.11–0.66); this was 0.43 ± 0.133 (0.11–0.77) for the reports ($n = 156$) of high methodological quality ($p = 0.07$).

Informed consent was not obtained in 7 (9.3%) trials with insufficient methodological quality ($n = 75$) or for 14 (8.9%) trials with high methodological quality ($n = 156$) ($p < 0.001$). The maximum possible score for the reports ($n = 156$) (p = 0.9).

REC approval was not reported for 10 (13.3%) trials with insufficient methodological quality ($n = 75$) or for 67 trials (42.9%) with high methodological quality ($n = 156$) ($p < 0.0001$).

DISCUSSION
This study was limited to phase III oncological trials. We selected randomised controlled trials (RCTs) from 10 journals. This restrictive choice was led by the recognised quality of the four medical journals selected (leading journals that publish research reports in all fields and have a broad readership) and because the six oncology journals comprised a good sample of oncology around the world (12 international oncology societies). The purpose of this choice was to create a homogeneous group of publications and conditions that allowed standardised analysis. This arbitrary choice may have introduced a bias causing overestimation of the quality of the RCTs analysed. Randomisation is recognised as the best available technique for approximating the equality of patient groups being compared. The validity of a clinical trial, however, depends on much more than the proper conduct of the randomisation process. Interpretation and application of the results depends on an adequate description of the patients accepted as well as of those not accepted into the trial, and also on the experimental and supplementary treatment regimens, withdrawals, blinding where appropriate, testing of how well the rules have been followed, and the use of proper statistical analysis. Inadequate reporting makes the interpretation of RCTs difficult if not impossible.

Meta-analyses of RCTs are being published with increasing frequency, resulting in great interest in assessing the quality of the trials included in these meta-analyses. Numerous scales and checklists have been suggested to evaluate the quality of RCTs. However, the 11-item scale devised by Jadad et al is the only one known to have been formulated using standard scale development techniques. Although this scale was developed and validated to assess the quality of reports of pain relief, because of its efficiency it has been used extensively in other clinical areas. Methodological
quality is the first ethical requirement in clinical trials. Three methods are used to assess quality: individual markers, checklists, and scales. Scales have the theoretical advantage over the other methods in that they provide quantitative estimates of quality that can be replicated easily and incorporated formally into the peer review process and into systematic reviews. The main disadvantage of quality scales is that there is a dearth of evidence supporting either the inclusion or exclusion of items or the numerical scores attached to each of those items. Another disadvantage is that assessments of quality depend on the information available in the reports. Space constraints in most printed journals, the referral of readers to previous publications as sources for detailed description of the methods used, and the publication of trials in abstract form could all lead to the assumption that a trial was methodologically deficient, even when it has been designed, conducted, and analysed appropriately. However, in interpreting trial results, the reader has only the published paper on which to rely.

Chalmers et al.14 analysed the phenomenon of incomplete reporting. The authors of 41 of 59 trials responded to a questionnaire, whereupon 58% of the missing items were found to have been carried out. Many reviews have documented deficiencies in reports of clinical trials. For example, only 12 (27%) of 45 reports published in three medical journals in 1985 defined a primary endpoint, and only 43% of 37 trials with negative findings published in 1990 reported a sample size calculation.4 Reporting is not only frequently incomplete but also sometimes inaccurate. Among 119 reports stating that all participants were included in the analysis in the groups to which they were originally assigned (intention-to-treat analysis), 15 (13%) either excluded patients or did not analyse all randomised patients as allocated.20 Unfortunately, reporting of the methods used for the allocation of participants to interventions is also occasionally inadequate. For example, at least 5% of 206 reports of supposed RCTs in obstetrics and gynaecology journals described studies that were not truly randomised.21

Results from poorly designed and reported trials can mislead decision making in health care at all levels, from treatment decisions for individual patients to the formulation of national public health policies. In this study, 67.5% (156/231) RCT were of good methodological quality; this proportion could be overestimated by the choice of journals in which the RCT were selected. However, only 55 RCTs (23.8%) obtained the maximum score that all RCTs should attain.

Another disadvantage of quality scales is that assessments of quality are done late in the process, at the time of publication, which has little impact on improving the quality of RCTs. A group of scientists and editors developed the CONSORT (CONsolidated Standards Of Reporting Trials) statement to improve the quality of reporting of RCTs.22-25 This consists of a checklist and flow diagram to assist authors. The CONSORT statement is not meant to be used as a quality assessment instrument; its objective is to facilitate the critical appraisal and interpretation of RCTs by providing guidance to authors on how to improve the reporting of their trials. Since its publication, CONSORT has been supported by an increasing number of journals and several editorial groups,22-25 including the International Committee of Medical Journal Editors.26 A study has been performed to determine whether the use of the CONSORT statement is associated with improvement in the quality of reporting RCTs.27 A comparative before and after evaluation was carried out, in which reports of RCTs published in 1994 (pre-CONSORT) were compared with those from the same journals published in 1998 (post-CONSORT). Reports from the British Medical Journal, the Journal of the American Medical Association, and The Lancet (journals that adopted CONSORT) were compared with those published in the New England Journal of Medicine (a journal that did not adopt CONSORT) and the results analysed. Compared with 1994, the methodological quality based on the Jadad scale increased in all four journals in 1998; this increase was statistically significant for endorsing journals. These findings encourage our belief that evidence based approaches such as the CONSORT statement will improve the quality of reporting RCTs, which will ultimately result in less bias and more appropriate information for consumers. In the present study we highlighted that there was no improvement in methodological quality with time; however only five of the journals included have endorsed the CONSORT statement (The Lancet, the British Medical Journal, the Journal of the American Medical Association, Annals of Oncology, and the European Journal of Cancer).

### Table 1

<table>
<thead>
<tr>
<th>Score</th>
<th>No. trials</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>22</td>
</tr>
<tr>
<td>4</td>
<td>48</td>
</tr>
<tr>
<td>5</td>
<td>89</td>
</tr>
<tr>
<td>6</td>
<td>52</td>
</tr>
<tr>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>8</td>
<td>7</td>
</tr>
</tbody>
</table>

### Table 2

<table>
<thead>
<tr>
<th>Item</th>
<th>n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Obtaining informed consent from patients*</td>
<td>210 (90.9)</td>
</tr>
<tr>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>21 (9.1)</td>
</tr>
<tr>
<td>2. Approval by a REC*</td>
<td>154 (66.7)</td>
</tr>
<tr>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>77 (33.3)</td>
</tr>
<tr>
<td>3. Risk-benefit ratio valuation</td>
<td>5 (2.2)</td>
</tr>
<tr>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>226 (97.8)</td>
</tr>
<tr>
<td>4. Respect for the principle of a priori equivalence</td>
<td>231 (100)</td>
</tr>
<tr>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>0 (0)</td>
</tr>
<tr>
<td>5. Refusal of consent</td>
<td>9 (3.9%)</td>
</tr>
<tr>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>222 (96.1)</td>
</tr>
<tr>
<td>6. Placebo ethical justification†</td>
<td>1</td>
</tr>
<tr>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>23</td>
</tr>
<tr>
<td>7. Fairness of participant selection (inclusion/exclusion criteria)</td>
<td>230 (&gt;99)</td>
</tr>
<tr>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1 (&lt;1)</td>
</tr>
<tr>
<td>8. Planned interim analysis</td>
<td>67 (29.0)</td>
</tr>
<tr>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>164 (71.0)</td>
</tr>
<tr>
<td>9. Prospectively defined stopping rules</td>
<td>37 (16.0)</td>
</tr>
<tr>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>194 (84.0)</td>
</tr>
<tr>
<td>10. Independent monitoring committee</td>
<td>33 (14.3)</td>
</tr>
<tr>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>198 (85.7)</td>
</tr>
</tbody>
</table>

*13 clinical trials (5.6%) reported neither that informed consent had been requested from the participants nor that a REC had approved the research.
†Only 24 articles were described as double blind.

---

1. Chalmers et al.14
2. Three methods are used to assess quality: individual markers, checklists, and scales.
3. Scales have the theoretical advantage over the other methods in that they provide quantitative estimates of quality that can be replicated easily and incorporated formally into the peer review process and into systematic reviews.
4. Reporting is not only frequently incomplete but also sometimes inaccurate.
5. Among 119 reports stating that all participants were included in the analysis in the groups to which they were originally assigned (intention-to-treat analysis), 15 (13%) either excluded patients or did not analyse all randomised patients as allocated.
6. Unfortunately, reporting of the methods used for the allocation of participants to interventions is also occasionally inadequate.
7. For example, at least 5% of 206 reports of supposed RCTs in obstetrics and gynaecology journals described studies that were not truly randomised.
8. Results from poorly designed and reported trials can mislead decision making in health care at all levels, from treatment decisions for individual patients to the formulation of national public health policies.
9. In this study, 67.5% (156/231) of RCT were of good methodological quality; this proportion could be overestimated by the choice of journals in which the RCTs were selected. However, only 55 RCTs (23.8%) obtained the maximum score that all RCTs should attain.
10. Another disadvantage of quality scales is that assessments of quality are done late in the process, at the time of publication, which has little impact on improving the quality of RCTs.
11. A group of scientists and editors developed the CONSORT (CONsolidated Standards Of Reporting Trials) statement to improve the quality of reporting of RCTs.
12. This consists of a checklist and flow diagram to assist authors.
13. The CONSORT statement is not meant to be used as a quality assessment instrument; its objective is to facilitate the critical appraisal and interpretation of RCTs by providing guidance to authors on how to improve the reporting of their trials.
14. Since its publication, CONSORT has been supported by an increasing number of journals and several editorial groups, including the International Committee of Medical Journal Editors.
15. A study has been performed to determine whether the use of the CONSORT statement is associated with improvement in the quality of reporting RCTs.
16. A comparative before and after evaluation was carried out, in which reports of RCTs published in 1994 (pre-CONSORT) were compared with those from the same journals published in 1998 (post-CONSORT).
17. Reports from the British Medical Journal, the Journal of the American Medical Association, and The Lancet (journals that adopted CONSORT) were compared with those published in the New England Journal of Medicine (a journal that did not adopt CONSORT) and the results analysed.
18. Compared with 1994, the methodological quality based on the Jadad scale increased in all four journals in 1998; this increase was statistically significant for endorsing journals.
19. These findings encourage our belief that evidence based approaches such as the CONSORT statement will improve the quality of reporting RCTs, which will ultimately result in less bias and more appropriate information for consumers.
20. In the present study we highlighted that there was no improvement in methodological quality with time; however only five of the journals included have endorsed the CONSORT statement (The Lancet, the British Medical Journal, the Journal of the American Medical Association, Annals of Oncology, and the European Journal of Cancer).
The total ethical score (Berdeu scale) calculated for the 231 RCTs studied was rather disappointing at 0.42 ± 0.133, being lower than that obtained by Berdeu et al\textsuperscript{28} with a sample of 24 RCTs (0.44 ± 0.118), but higher than with clinical trials involving elderly patients (0.334 ± 0.118).\textsuperscript{29} Of the reports on the 231 clinical trials, 8 (3.5%) did not state that informed consent had been requested from the participants, 64 (27.7%) did not note that an institutional review board had approved the research, and 13 (5.6%) reported neither REC approval nor ICP.

Ruiz-Canela et al\textsuperscript{30} assessed the frequency of reporting ICP and REC approval in trials published between 1993 and 1995.\textsuperscript{31} For 767 RCTs, 543 (70.8%) stated that a REC had approved the research and 612 (79.8%) that ICP had been requested. Both types of information were included in 64% of the RCTs. When information on ICP or REC approval was missing these authors mailed a questionnaire to the corresponding authors. The response rate was 73.7%; 22.4% stated that they had not sought approval and 20.6% that they had not requested ICP. These results are surprising, showing that not more than 60% (but more than 40%) of the authors did not respect ethical criteria; this percentage is reduced to 21.7% when taking into account authors who responded to the questionnaire. However, failing to report ethical requirements in original articles may already imply a lower concern for ethical issues in trials.\textsuperscript{32} It is even more surprising to note that these articles were accepted for publication in prestigious journals when the authors respected neither the patients involved in the trial nor fundamental ethical principles.

We have highlighted that there was an improvement in ethical quality with time and that this was not related to the particular journals in which the RCTs were published. There were no significant differences in the proportion of reports that included information on institutional review board approval and ICP in relation to the number of patients included. This result is reassuring because it demonstrates that RCTs with a small number of participants have the same level of ethical quality as those with more participants. On the other hand there was no correlation between ethical quality and methodological quality. RCTs with higher methodological quality standards did not give more attention to ethical aspects than those with a poor methodological quality. Unfortunately, readers cannot consider studies of good methodological quality as models and examples of good research. In the current study, 42.9% of the RCTs with good methodological quality did not state that a REC had approved the research.

Improvement in the ethical quality of RCTs is mandatory. Appropriate effort must be made throughout the life of a trial, from its conception to the publication and communication of results.\textsuperscript{33} The sensitisation of researchers to the ethical aspects of research is a step by step process. We believe that this sensitisation should occur throughout both initial and continuing training of all persons concerned with research, from researcher to reader.

Ethics committees have to protect the interests of research participants and must have a policy of encouraging research, but very few such committees monitor what happens to the research they approve. Acceptance is growing that RECs have responsibilities that continue after their approval of research.\textsuperscript{34, 35} One of these responsibilities should be regular auditing of the results of the research that RECs have approved.

Medical journals have a key role in any movement towards raising ethical standards in medical research. Editors have an important part to play by requiring guarantees regarding the obtaining of permission from RECs and the ethical monitoring of medical research. Some journals have included in their instructions to authors that a paragraph must be devoted to the ethical aspects of trials. These must be described together with details of the ethical monitoring carried out during the research. These recommendations should be widely disseminated.

Similar to the CONSORT statement, ethical standards for the reporting of trials must be developed. This policy would have three advantages: to publish only studies respecting the patients involved in the research; to assist investigators throughout, from conception of the research project to the publication of results; and to facilitate appraisal by peer reviewers and editors.

CONCLUSION

The phase III randomised cancer trials that were analysed were shown to have satisfactory methods in 67.5% of cases. Their ethical quality, however, was disappointing, with approximately 30% of authors not respecting fundamental ethical principles. Improvements in the teaching and sensitising of researchers are mandatory, as well as the reinforcement of trials monitoring by RECs. The editors of medical journals and RECs are the watchdogs of patients’ rights; they are also implicated in this walk for ethical quality and their level of vigilance must be increased. It is essential to develop an assessment of the ethical value of research featured in journals in the same way as methodological value is assessed. Scientific publications have a duty to contribute to the spreading and application of ethical principles. In the same way that today a study can be rejected for publication because of methodological inadequacies, in the future it could more likely be refused for ethical shortcomings. In RCTs, failing to obtain or forgetting to report the obtaining of informed consent from participants, or approval of the protocol by a REC, suggest that the authors considered these steps to be unimportant details, if not obstacles. This is clearly a misconception because the aim of human research is to serve the participants, not to use them.

The association between methodological quality and the reporting of ethical requirements could be seen to reflect the respect shown for patients during the whole research process.

Authors’ affiliations
J J Tuech, G Moutel, S Schraub, C Herve, Laboratoire d’Ethique Medicale et de Sante Publique, Faculte de Medicine Necker, Universite Paris Rene Descartes, France
J J Tuech, Department of Digestive Surgery, University Hospital, France
P Pessaux, Departement de Statistiques Bio-medicales, CHU Angers, France
V Thoma, S Schraub, CRILCC Paul Strauss, Strasbourg, France

REFERENCES

www.jmedethics.com
Methodological quality and reporting of ethical requirements in phase III cancer trials


29 Rennie D, Yank V. Disclosure to the reader of institutional review board approval and informed consent. JAMA 1997;277:922–3.

