Can the Accuracy of Abstracts Be Improved by Providing Specific Instructions?  
A Randomized Controlled Trial

Roy M. Pitkin, MD; Mary Ann Branagan

Context.—The most-read section of a research article is the abstract, and therefore it is especially important that the abstract be accurate.

Objective.—To test the hypothesis that providing authors with specific instructions about abstract accuracy will result in improved accuracy.

Design.—Randomized controlled trial of an educational intervention specifying 3 types of common defects in abstracts of articles that had been reviewed and were being returned to the authors with an invitation to revise.

Mean Outcome Measure.—Proportion of abstracts containing 1 or more of the following defects: inconsistency in data between abstract and body of manuscript (text, tables, and figures), data or other information given in abstract but not in body, and/or conclusions not justified by information in the abstract.

Results.—Of 250 manuscripts randomized, 13 were never revised and 34 were lost to follow-up, leaving a final comparison between 89 in the intervention group and 114 in the control group. Abstracts were defective in 25 (28%) and 30 (26%) cases, respectively (P = .78). Among 55 defective abstracts, 28 (51%) had inconsistencies, 16 (29%) contained data not present in the body, 8 (15%) had both types of defects, and 3 (5%) contained unjustified conclusions.

Conclusions.—Defects in abstracts, particularly inconsistencies between abstract and body and the presentation of data in abstract but not in body, occur frequently. Specific instructions to authors who are revising their manuscripts are ineffective in lowering this rate. Journals should include in their editing processes specific and detailed attention to abstracts.
data were tabulated by a different staff member, the managing editor. Experience indicated that at least a quarter of abstracts are defective in 1 or more of the ways described. Based on an assumed reduction from 25% to 10% with the intervention, and assuming a of .05 and of .80, we projected a sample size of 100 in each arm. Because of anticipated losses, we enrolled 250 manuscripts. The proportion of defective abstracts in the instructed and uninstructed groups were analyzed by \( \chi^2 \) test.

RESULTS

Of 250 manuscripts enrolled, 119 were assigned to receive the instruction sheet and 131 to the uninstructed or control group. Thirteen manuscripts were withdrawn (ie, a revision was not returned) and 34 were otherwise lost to follow-up analysis, leaving for final analysis 89 in the intervention group and 114 controls (Figure). One or another of the types of defects was identified in 25 instructed abstracts (28%; 95% confidence interval [CI], 19%-37%) and in 30 uninstructed abstracts (26%; 95% CI, 18%-34%), insignificant differences (\( P = .78 \)). With respect to specific type of defect found, 28 of the 55 defective abstracts (51%; 95% CI, 38%-64%) had inconsistencies with the body of the manuscript, 16 (29%; 95% CI, 17%-41%) contained data or other information not found in the body, 8 (15%; 95% CI, 10%-20%) had both types of defects, and only 3 (5%; 95% CI, 3%-7%) contained inappropriate or unjustified conclusions. There were no differences apparent between intervention and control groups with respect to type of defect found.

The proportion of manuscripts withdrawn or otherwise lost to analysis was large, and the distribution between intervention groups was disproportional. Of the 13 withdrawals, 4 had been assigned to the instructed groups and 9 to the uninstructed. Of 34 otherwise un-

available, 26 had been assigned to the intervention group and 8 to the control group. We recalculated the results with the assumption that none of the 30 withdrawn or unavailable manuscripts assigned to the intervention would have been returned with defective abstracts, and that 9 (56%) of the 16 unavailable or withdrawn manuscripts assigned to control would have had defective abstracts. Under these highly unlikely conditions, the number of defective abstracts would be 25 (21%; 95% CI, 14%-28%) of 119 in the instructed group and 39 (30%; 95% CI, 22%-38%) of 131 in the uninstructed. This difference is still not statistically significant (\( \chi^2, 2.12; P = .15 \)).

COMMENT

The abstract of a research article, it could reasonably be argued, is the most important part of the article. It is by far more likely to be read than any other section of the report. The ubiquitous availability and widespread use of automated literature search mechanisms, which provide an (often truncated) abstract, have done nothing but increase this likelihood.

Given the importance of the abstract, there has been surprisingly little research into its accuracy. Narine and colleagues analyzed the quality of abstracts of original research papers published in 1989 in the Canadian Medical Association Journal, using a special instrument developed specifically for the study. A number of deficiencies were identified, but none related specifically to consistency of data in the abstract and the body or to the basis of data in the abstract. Roberts and associates examined the effects of peer review and editing on the readability of articles from the Annals of Internal Medicine; evidence was found of a modest increase in readability in both text and abstract, but accuracy was not addressed. Goodman and associates compared the effects of peer review and editing on manuscript quality in the same journal; only 1 of 34 items involved the abstract, and it was a very general assessment.

The major development involving abstracts during the last decade has been the introduction of structured abstracts. Although there have been objections to the structured format, based on length and aesthetic concerns, it has been adopted widely in one form or another, and there seems to be general acceptance that it is more informative than the unstructured variety. Some evidence exists that both quality and understanding may be improved by structured abstracts. However, there is little reason to suspect that requiring a structure will lessen the types of discrepancies and omissions we assessed.

The present study was designed to test the effectiveness of an educational intervention on types of defects in abstracts of papers undergoing revision. We found that the proportion of abstracts with one or another of 3 defects was 26% to 28% and—disappointingly—not affected by the simple intervention tested. The most common defect, present alone or in combination in 17% of all manuscripts and 65% of defective ones, was inconsistency between abstract and body.

To determine if the types of defects were identified are regularly recognized by all journals and dealt with during the copyediting process, much as are errors in grammar and syntax, we did a small study involving analysis of published abstracts in 4 journals: American Journal of Obstetrics and Gynecology (July 1996 issue), Pediatrics (October 1996 issue), JAMA (July 26, August 2, and August 9, 1995, issues), and The New England Journal of Medicine (August 29, September 5, and September 12, 1996, issues). The results of this analysis of published articles are summarized in the Table. Surprisingly, the proportion of published abstracts that we found defective was as high or higher than what we observed in those that had been reviewed and revised but were not in final copyedited form. This was a small and not scientifically derived sample, but the findings suggest that journals do not, as part of their regular copyediting procedures, scrutinize abstracts for these types of problems.

We conclude that inconsistencies in data between abstract and body and reporting of data and other information solely in the abstract are relatively common and that a simple educational intervention directed to the author is ineffective in reducing that frequency. Until effective interventions are devised, journals and publishers should incorporate
Reporting of Randomized Clinical Trial Descriptors and Use of Structured Abstracts

Robert W. Scherer, PhD; Barbara Crawley, MS

...into their copyediting procedures the practice of detailed and specific verification of all data and other information in the abstract.

We thank Leon F. Burmeister, PhD, for advising us on statistical and design matters and providing the set of random numbers.

References