Effects of Technical Editing in Biomedical Journals
A Systematic Review

Elizabeth Wager, MA
Philippa Middleton, MPH

Most articles in biomedical journals undergo some form of editing between acceptance and publication. The intensity and type of editing vary between journals but the process often includes applying house style to references, abbreviations, and numbers, checking articles for consistency, clarity, and completeness, and correcting grammatical errors. For this review, we examined any processes applied to articles between acceptance and publication that were designed to improve accuracy or clarity or impose a predefined style.

METHODS

We performed a systematic review using Cochrane methods, searching MEDLINE, EMBASE, Current Contents, and 10 other databases from earliest entries to February 2000. We used inclusive search terms such as writing, editing, accuracy, and readability. Relevant journals (eg, Learned Publishing, Journal of Information Sciences, and JAMA special issues on peer review) were hand searched. We contacted researchers working in the field to request publications that we had missed. The full search strategy is described in the protocol.1

The full review included any comparative studies of the effects of processes designed to improve the quality of accepted research articles in biomedical journals.1 Those on editorial decision making and peer review formed the basis of a separate review.2

Although we sought evidence about biomedical journals, we did not restrict our search to studies published in such journals. Studies of readability or comprehension had to involve journal readers (ie, health care professionals). We included articles in any language but restricted our searches on evidence about improving writing style to English. This article presents the findings from the subset of articles that focused on editorial processes between acceptance and publication. Details of studies on the effects of other processes, such as providing instructions to contributors and imposing a structured abstract format, can be found in the Cochrane review.3

RESULTS

Because few studies used comparable methods, we performed a descriptive review of 11 articles (TABLE). The study at the Dutch Medical Journal3 demonstrated significant improvements in papers between acceptance and publication, as measured by Gunning Fog and Flesch reading ease scores, may improve other aspects of their quality, can increase the accuracy of references and quotations, and raises the quality of abstracts. Supplied authors with abstract preparation instructions had no discernible effect.

Conclusions Considering the time and resources devoted to technical editing, remarkably little is known about its effects or the effects of imposing different house styles. Studies performed at 3 journals employing relatively large numbers of professional technical editors suggest that their editorial processes are associated with increases in readability and quality of articles, but these findings may not be generalizable to other journals.

JAMA. 2002;287:2821-2824

Author Affiliations: Sideview, Buckinghamshire, England (Ms Wager); Australian Safety and Efficacy Register of New Intervention Procedures-Surgical, Royal Australasian College of Surgeons, Adelaide, Australia (Ms Middleton)

Corresponding Author and Reprints: Elizabeth Wager, MA, Sideview, 19 Station Rd, Princes Risborough, Bucks, HP27 9DE England (e-mail: liz@sideview.demon.co.uk).
Table. Summary of Studies on Technical Editing

<table>
<thead>
<tr>
<th>Source, y</th>
<th>Methods</th>
<th>Outcome Measures</th>
<th>Results</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Quality of Articles Between Acceptance and Publication</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pierie et al, 1996</td>
<td>Compared accepted vs published versions of 50 articles at Dutch Medical Journal</td>
<td>5-Point, 17-item scoring system applied by readers</td>
<td>Significant improvements in 11 out of 16 items</td>
<td>Quality of papers improved by technical editing</td>
</tr>
<tr>
<td>Goodman et al, 1994</td>
<td>Compared submitted vs published versions of 109 articles at Annals of Internal Medicine</td>
<td>5-Point, 34-item scoring system applied by expert assessors</td>
<td>Percentage of items scoring &gt;3 rose from 75% to 85%; significant increases in discussion of study limitations (47%-65%), acknowledging generalizations (58%-79%), appropriate conclusions (71%-85%), and use of confidence intervals (65%-81%)</td>
<td>Quality of papers improved between submission and publication</td>
</tr>
<tr>
<td><strong>Readability Studies</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biddle and Aker, 1996</td>
<td>Compared submitted vs published versions of 33 articles from American Association of Nurse Anesthetists Journal</td>
<td>Flesch reading ease, Gunning Fog scores</td>
<td>Mean Flesch score 24.6 before, 32.5 after; mean Gunning Fog score 19.4 before, 14.9 after</td>
<td>Readability improved between submission and publication</td>
</tr>
<tr>
<td>Roberts et al, 1994</td>
<td>Compared submitted vs published versions of 101 articles from Annals of Internal Medicine</td>
<td>Flesch reading ease, Gunning Fog scores</td>
<td>Mean Flesch score 28.2 before, 29.1 after; mean Gunning Fog score 17.2 before, 16.9 after</td>
<td>Readability improved between submission and publication</td>
</tr>
<tr>
<td><strong>Effects of Editing on Accuracy of References</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Holbma and Overbeke, 1992</td>
<td>Compared reference accuracy in submitted vs published letters in BMJ (248 references submitted, 67 published; 61 quotes submitted, 28 published)</td>
<td>Reference accuracy</td>
<td>Incorrect citations fell from 70% to 31%; between submission and publication</td>
<td>In-house checking improved citation accuracy</td>
</tr>
<tr>
<td>Lowry, 1985</td>
<td>Compared reference and quotation accuracy in submitted vs published letters in BMJ</td>
<td>Reference accuracy, quotation accuracy</td>
<td>Percentage of correct references rose from 69% to 92%; percentage of accurate quotations rose from 50% to 67%</td>
<td>Copyediting improved reference and quotation accuracy</td>
</tr>
<tr>
<td>George and Robbins, 1994</td>
<td>Compared journals checking reference accuracy (British Journal of Dermatology, Journal of the American Academy of Dermatology) with those that do not (Archives of Dermatology, Journal of Investigative Dermatology) (60 references from each)</td>
<td>Reference accuracy</td>
<td>35% Inaccurate references in journals that check vs 47.5% in those that do not (P = .07)</td>
<td>In-house checking may improve reference accuracy</td>
</tr>
<tr>
<td><strong>Effects of Editing on Abstract Quality</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pitkin et al, 1999</td>
<td>Compared abstract quality in 6 journals</td>
<td>Percentage of abstracts assessed as defective by editors</td>
<td>Percentage of defective abstracts varied between journals (32%-68%)</td>
<td>Journal interventions may affect abstract quality</td>
</tr>
<tr>
<td>Pitkin et al, 2000</td>
<td>Compared abstract quality before and after improvement initiative at JAMA (50 abstracts before vs 50 after)</td>
<td>Percentage of defective abstracts as assessed by editorial staff</td>
<td>Percentage of defective abstracts decreased from 52% before to 20% after the intervention</td>
<td>Intensive in-house editing raised abstract quality</td>
</tr>
<tr>
<td>Winker, 1999</td>
<td>Compared abstract quality before and after improvement initiative at JAMA (21 abstracts before vs 21 after)</td>
<td>Percentage of defective abstracts</td>
<td>Percentage of defective abstracts decreased from &gt;50% before to 0% after the intervention</td>
<td>Intensive in-house editing raised abstract quality</td>
</tr>
<tr>
<td><strong>Effects of Instructions to Authors</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pitkin and Branagan, 1998</td>
<td>Randomized controlled trial of providing instructions on abstracts to authors of papers accepted by Obstetrics and Gynecology (89 received instructions, 114 did not)</td>
<td>Percentage of defective abstracts as assessed by editorial staff</td>
<td>No discernible difference in abstract accuracy; 26% inaccurate in control group vs 26% in the instructed group</td>
<td>Providing instructions on abstracts did not improve their quality</td>
</tr>
</tbody>
</table>
united to distinguish the effects of peer review from technical editing.

Two studies compared readability scores and demonstrated improvements between submission and publication. Although the improvements were statistically significant, the absolute increases were small, and scores indicate that, even after editing, the articles remained difficult to read. Guidelines for using the Flesch readability scale suggest that scores of 50 to 60 are desirable for standard documents, but articles in studies scored around 30. The Gunning Fog test, which gives lower scores for clearer writing, indicated that papers remained difficult to read, with scores of 15 to 17, on a par with legal contracts or corporate reports (as opposed to quality newspapers, which score about 9). The articles did not provide details about which interventions were thought to affect readability.

Three studies considered the effects of technical editing on the accuracy of reference citations and quotations. Those that compared submitted and published versions showed clear improvements in accuracy. A comparison of journals with different checking policies was less conclusive.

A study of abstract completeness and consistency indicated variation between journals but could not identify the processes associated with better results, while 2 studies relating to the introduction of more intensive abstract editing at JAMA showed clear improvements in accuracy after the intervention. Another study by Pitkin and Branagan examined the effects of sending abstract preparation instructions to authors after their article had been accepted but did not detect any improvement in the quality of abstracts from authors who had received the intervention.

### Quality of Study Methods
The quality of study methods varied; 6 compared different versions of articles, 2 used before and after designs, 2 made comparisons between journals, and 1 used a randomized design. Some of the studies comparing different versions of articles (eg, accepted vs published) attempted to mask the articles' status, but often unsuccessfully (eg, in the Dutch Medical Journal study, 72% of assessors correctly identified the version of the article). Comparisons between journals are hard to interpret because findings may be influenced by extraneous factors. Comparisons over time may be similarly confounded.

Another aspect of methodological quality is the use of appropriate rating scales and assessors. Two studies used Flesch and Gunning readability scores, which have not been validated on scholarly literature and are based not on adult readers, but on the ways in which children learn to read. Both produce scores based on sentence and word length but do not take into account word familiarity or sentence complexity. Therefore, they are at best surrogate markers of the comprehensibility of a research report. A study in which students assessed structured and unstructured abstracts found that the structured version was considered significantly easier to read, despite the fact that Flesch scores were similar for both versions. Further work is needed to test the ability of these scores to measure the comprehensibility of biomedical articles.

Only one study recruited journal readers to act as assessors. In other cases, the investigators themselves, experts, or professional editors rated the quality of papers. Although this step may have reduced interrater variability, the everyday use of journals may not have been reflected, and there is evidence that readers do not always agree with experts on the quality of papers. Several studies used specially devised scales to rate the quality of papers or abstracts. In most cases, no details of interrater reliability or face validity were given. However, the study by Goodman et al reported poor reliability (intraclass correlation coefficient of 0.12) for its scoring system.

### Definitions of Technical Editing and Generalizability
Most studies that examined changes between acceptance and publication did not specify the processes that took place. However, the Dutch Medical Journal investigators note that during editing, the information in the article is checked scientifically and linguistically, corrected and clarified if necessary, numbers are checked when possible, and the references are made to conform to the so-called Vancouver system.

Studies of reference accuracy indicate how the intensity of editing varies between journals. For example, editors at the Dutch Medical Journal and British Journal of Dermatology check references against MEDLINE, while those at the Journal of the American Academy of Dermatology check references only from their own journal. Lowry, who studied letters in the BMJ in 1985, explains that although the journal does not check all references . . . the subeditors correct any obvious errors. References are put into the house style, which allows many mistakes to be spotted, especially where the fault is an incomplete reference, which is inevitably corrected.

Many of the studies of technical editing relate to general medical journals that employ professional technical editors. They are likely to have received more training and to have more time than academic editors who edit their journal in addition to holding a full-time job. Thus, Goodman et al. comment that the relatively large editorial staff at Annals is not typical of any but the largest medical journals, and the generalization to others with different . . . editing processes cannot easily be made.

Three of the 4 studies examining overall quality or readability failed to distinguish between changes occurring between submission and acceptance, which might be considered part...
REFERENCES

CONCLUSIONS
Our review suggests a pattern of improvement in the accuracy and readability of research articles, which occurs between acceptance and publication. However, few studies have attempted to determine which processes contribute most to this improvement or how variations in these processes might affect the quality of published papers. Without elucidation of the processes involved, these findings may not be generalizable. Despite the time and resources spent on technical editing, articles remain difficult to read, although intensive editing and checking may lead to improvements in the accuracy of references and abstracts. Few studies have consulted journal readers about their needs or views or worked with authors to improve manuscript quality.

Author Contributions: Study concept and design: Wager, Middleton. Acquisition of data: Wager, Middleton. Analysis and interpretation of data: Wager, Middleton. Drafting of the manuscript: Wager. Critical revision of the manuscript for important intellectual content: Wager, Middleton.

Disclosure: Much of the work of this review was done when Philippa Middleton was working at the UK Cochrane Centre and Elizabeth Wager was employed at GlaxoSmithKline, but the project did not receive any specific funding.

Acknowledgment: We thank other members of the Cochrane Peer Review and Technical Editing Systematic review (PIRATES) team (Tom Jefferson, MD, Phil Alderson, MBCHB, and Frank Davidoff, MD) for their help in developing the protocol and performing this review.

EFFECTS OF TECHNICAL EDITING
of peer review, and those between acceptance and publication, which meet our definition of technical editing. Only the Dutch study specifically examined and demonstrated improvements occurring between acceptance and publication. A more recent study at Annals of Internal Medicine found that only 3% of substantive changes introduced in manuscripts resulted from technical editing, although the origin of 47% of changes was unknown, so this figure may have been substantially higher. It is also possible that other textual changes (that were not considered in the Annals study) contributed to manuscript clarity and readability (Frank Davidoff, oral communication, September 2001).

Need for Further Research
Nearly all journals impose a house style that includes elements of typographic design, such as typeface, and scientific conventions, such as number format. The only aspect of journal style that has attracted research is the structuring of abstracts. Although studies comparing readability scores are inconclusive, structured abstracts are preferred by readers and are more comprehensive but are also longer than unstructured ones. Apart from this, we found no research about the effects of different styles on legibility or readability, and we conclude that the imposition of such styles is not evidence based, unless journals have undertaken unpublished research.