STUDY

Quality of Abstracts in 3 Clinical Dermatology Journals

Alain Dupuy, MD; Kiarash Khosrotehrani, MD; Céleste Lebbé, MD, PhD; Michel Rybojad, MD; Patrice Morel, MD

Background: Structured abstracts have been widely adopted in medical journals, with little demonstration of their superiority over unstructured abstracts.

Objectives: To compare abstract quality among 3 clinical dermatology journals and to compare the quality of structured and unstructured abstracts within those journals.

Design and Data Sources: Abstracts of a random sample of clinical studies (case reports, case series, and reviews excluded) published in 2000 in the *Archives of Dermatology*, *The British Journal of Dermatology*, and the *Journal of the American Academy of Dermatology* were evaluated. Each abstract was rated by 2 independent investigators, using a 30-item quality scale divided into 8 categories (objective, design, setting, subjects, intervention, measurement of variables, results, and conclusions). Items applicable to the study and present in the main text of the article were rated as being present or absent from the abstract. A global quality score (range, 0-1) for each abstract was established by calculating the proportion of criteria among the eligible criteria that was rated as being present. A score was also calculated for each category. Interrater agreement was assessed with a κ statistic. Mean±SD scores were compared among journals and between formats (structured vs unstructured) using analysis of variance.

Main Outcome Measures: Mean quality scores of abstracts by journal and by format.

Results: Interrater agreement was good (κ=0.71). Mean±SD quality scores of abstracts were significantly different among journals (*Archives of Dermatology*, 0.78±0.07; *The British Journal of Dermatology*, 0.67±0.17; and *Journal of the American Academy of Dermatology*, 0.64±0.15; $P=.045$) and between formats (structured, 0.71±0.11; and unstructured, 0.56±0.18; $P=.002$). The setting category had the lowest scores.

Conclusions: The quality of abstracts differed across the 3 tested journals. Unstructured abstracts were demonstrated to be of lower quality compared with structured abstracts and may account for the differences in quality scores among the journals. The structured format should be more widely adopted in dermatology journals.

Arch Dermatol. 2003;139:589-593

THE ABSTRACT is an important part of a biomedical publication, frequently read and easily accessed through computerized bibliographic databases. The abstract should help the reader decide whether reading the whole article is relevant to his or her subject. However, previous findings indicate that clinical decisions are made based on reading the abstract alone, without referring to the full text.[1]

Acknowledging the pivotal role of the abstract, recommendations were made in the late 1980s by a working group to promote a structured presentation of abstracts.[2,3] The structured format has been subsequently widely adopted in medical journals, with little demonstration of its superiority over unstructured abstracts. Previous attempts to assess quality differences between structured and unstructured abstracts have compared different periods[4] or different journals,[5] while others have tested the results of rewriting abstracts in a structured format,[6] therefore subjecting all of these studies to confounding bias. We aimed to evaluate and compare the quality of abstracts of articles published in 2000 in 3 major clinical dermatology journals, 2 of which combined structured and unstructured abstracts during this period. This sample allowed comparison of structured and unstructured abstracts during the same period within the same journals.

From the Service de Dermatologie, Hôpital Saint-Louis, Paris, France.
The authors have no relevant financial interest in this article.

METHODS

SAMPLE OF ABSTRACTS

For our sample, we chose the 3 leading clinical dermatology journals: *Archives of Derma-*

Abstract Quality Scale

Category	Item No.	Criterion
Objective	1	Was any information on the objective given?
	2	Was the objective explicitly stated?
	3	Was the main objective distinguished from secondary ones?
Design	4	Was any information on the research design given?
	5	Were technical descriptors* used?
	6	If a follow-up study, was the duration given?
Setting	7	Was any information on the setting given?
	8	Was the level of clinical care (eg, primary care) indicated?
Subjects	9	Was any information on the subjects given?
	10	Were common demographic characteristics given?
	11	Were technical descriptors* of subject selection (eg, random sample) used?
	12	Was the number of subjects indicated?
	13	Were the response and refusal rates indicated?
	14	Was the number of dropouts and losses indicated?
	15	If the samples were matched, were matching characteristics given?
Intervention†	16	Was any information on intervention given?
	17	Was a description given?
	18	Was the duration indicated?
Measurement of variables	19	Was any information on the measures given?
	20	Were the variables explicitly given?
	21	Was the source of the data given?
	22	If the measurements were subjective, was blinding (or nonblinding) of the observer mentioned?
Results	23	Were any results given?
	24	Were they directly related to the objective?
	25	Were appropriate numeric data given?
Conclusions	26	Were any conclusions drawn?
	27	Were they directly related to the objective?
	28	Were they consistent with the results?
	29	Were the study's limitations mentioned?
	30	Were the study's implications mentioned?

*Technical descriptors refer to those listed by the Ad Hoc Working Group for Critical Appraisal of the Medical Literature.[2]
†Intervention category was quoted only for therapeutic trials.

tology (*Arch Dermatol*), *The British Journal of Dermatology* (*Br J Dermatol*), and *Journal of the American Academy of Dermatology* (*J Am Acad Dermatol*). Selected articles reported a clinical study (excluding case reports, case series, and reviews), dealt with patients or volunteers (excluding predominantly pathological or biological work), had an abstract, and were published during 2000. The year was chosen because it allowed comparison between structured and unstructured abstracts in *Br J Dermatol* and *J Am Acad Dermatol*. For *Br J Dermatol*, 2000 was transitional between its publishing primarily unstructured (January-June) and primarily structured (July-December) abstracts. During 2000, *J Am Acad Dermatol* published structured and unstructured abstracts, and *Arch Dermatol* published structured abstracts exclusively. The MEDLINE database was searched on the PubMed Web site (http://www.ncbi.nih.gov/entrez/query.fcgi) of the US National Library of Medicine using the following query: "((Arch Dermatol[ta] OR J Am Acad Dermatol[ta] OR Br J Dermatol[ta]) AND 2000[dp] AND hasabstract AND (clinical trials[mh] OR clinical trial[pt] OR epidemiologic studies[mh])) NOT case report[mh]." Of the 228 retrieved references, 31 were excluded (12 observations or case series, 12 reviews, and 7 describing predominantly pathological or biological work), leaving 197 abstracts for evaluation. Relying on data from previous studies,[4,7] and given a ratio of structured-unstructured abstracts in *Br J Dermatol* and *J Am Acad Dermatol* in 2000 of 1.85, we estimated that 25% of the abstracts would yield a 95% power to detect a one third difference in scores between structured and unstructured abstracts (estimated mean ± SD, 0.6 ± 0.15; $\alpha = .05$; bilateral test). From the final list of 197 articles (45 in *Arch Dermatol*, 75 in *Br J Dermatol*, and 77 in *J Am Acad Dermatol*), 25% of articles in each journal were selected for evaluation using computer-generated random numbers, resulting in a list of 49 articles: 11 in *Arch Dermatol*, 19 in *Br J Dermatol,* and 19 in *J Am Acad Dermatol*.

ABSTRACT RATING

Abstracts were considered structured if they were broken down by headings into 5 or more parts. Two assessors (A.D. and K.K.) independently rated each abstract using a slightly modified version of the quality scale established by Narine et al.[8] This 30-item scale is presented in the **Table**; the criteria were classified according to 8 categories: objective, design, setting, subjects, intervention, measurement of variables, results, and conclusions. Each item was first rated using the following 2 questions: (1) Is this item applicable to the study? (2) If yes, is this piece of information reported in the main text of the article (as well as in the abstract)? If the answer to both of these questions was yes, the item was considered eligible and was rated yes or no based on the content of the abstract. If one of the questions was answered negatively, the item was considered ineligible and was not rated. A quality score (range, 0-1) was obtained for each abstract by calculating the proportion of criteria rated yes among the eligible criteria. Therefore, the global score evaluated the proportion of important information in the article that was also present in the abstract. A score was also calculated for each category. Disagreements between the 2 raters were resolved by discussion. The interrater agreement was good ($\kappa = 0.71$). Assessors were not blinded to the journal names. The length of the abstract (number of words, including headings for structured abstracts) was automatically calculated by a word processor count.

STATISTICAL ANALYSIS

One-way analysis of variance was used to compare mean scores among journals and to compare structured and unstructured abstracts. Correlation between abstract length (number of words) and score was calculated by Pearson correlation coefficient. Tests were 2-sided, and $P = .05$ was considered significant. Commercially available software was used for the statistical analysis (Excel97 for Windows, Microsoft, Redmond, Wash; and SAS version 8.0, SAS Institute, Cary, NC).

RESULTS

Quality scores of abstracts are presented in **Figure 1**. The mean ± SD abstract scores were 0.78 ± 0.07 for *Arch Dermatol*, 0.67 ± 0.17 for *Br J Dermatol*, and 0.64 ± 0.15 for *J Am Acad Dermatol* ($P = .045$, difference across the 3 journals).

Scores by category are presented in **Figure 2**. The journal that obtained the best global score (*Arch Dermatol*) received the best scores in almost all categories. Information on setting was notably missing from the 2 lower-scored journals.

Archives of Dermatology requested a structured abstract format. *The British Journal of Dermatology* and *Journal of the American Academy of Dermatology* published structured and unstructured abstracts. Ten abstracts (53%) in *Br J Dermatol* and 13 (68%) in *J Am Acad Dermatol* had a structured format. For these 2 journals, the mean ± SD score for structured abstracts (0.71 ± 0.11) was significantly higher than the score for unstructured abstracts (0.56 ± 0.18) ($P = .002$). Differences between structured and unstructured abstract scores were more pronounced when the 3 journals were considered together ($P < .001$).

Structured abstracts were longer on average than unstructured abstracts (mean ± SD, 256 ± 77 vs 169 ± 65 words; $P < .001$). A strong positive correlation between length and score was observed for unstructured abstracts (Pearson correlation coefficient, 0.75; $P = .002$), while no such significant correlation was observed for structured abstracts (Pearson correlation coefficient, 0.30; $P = .08$) (**Figure 3**).

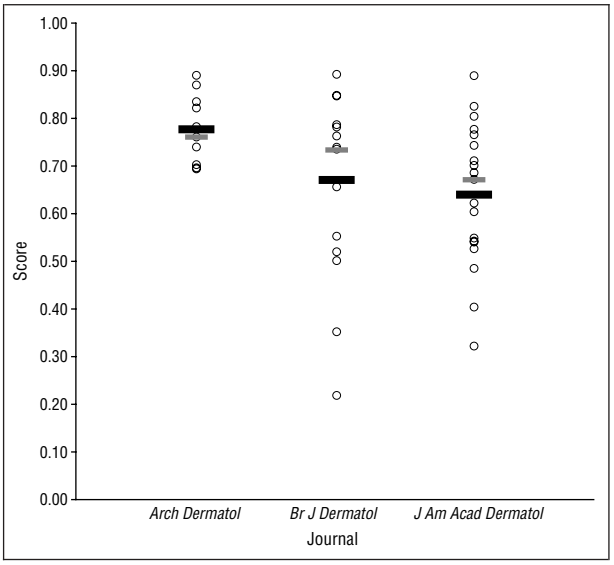

Figure 1. Abstract scores of selected articles. Mean ± SD scores, 0.78 ± 0.07 for *Archives of Dermatology (Arch Dermatol)*, 0.67 ± 0.17 for *The British Journal of Dermatology (Br J Dermatol)*, and 0.64 ± 0.15 for *Journal of the American Academy of Dermatology (J Am Acad Dermatol)* ($P = .045$). Median scores, 0.76 for *Arch Dermatol*, 0.73 for *Br J Dermatol*, and 0.67 for *J Am Acad Dermatol*. Thick bars represent means; thin bars, medians.

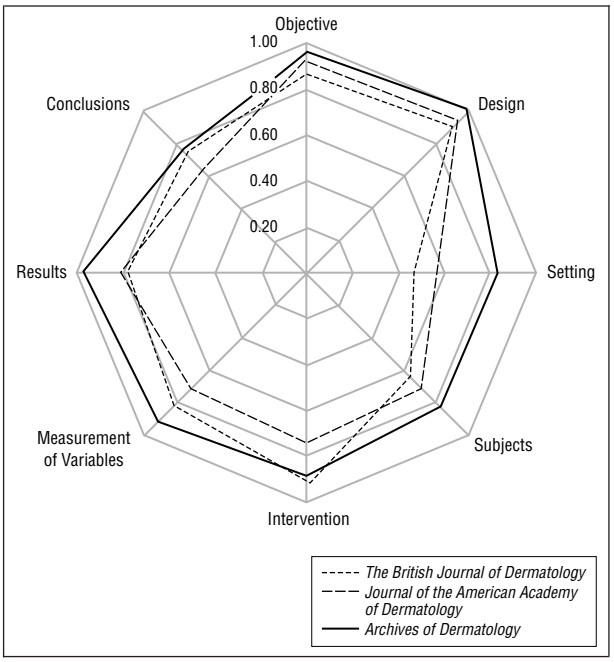

Figure 2. Abstract scores by category.

COMMENT

By comparing the quality scores of abstracts in 3 journals, we found significant differences among journals, and we demonstrated the superiority of the structured format over the unstructured format.

Clarification of the quality scale we used and the rating modalities is needed. First, the uniform weight of each criterion in the final score could be questioned, as it may be more important for an abstract to mention the objective (item 1) than the implications (item 30) of the study. However, a consensus on the weighting or on the choice of criteria would be hard to achieve. Nonetheless, we were satisfied with this scale because it offered good interrater reproducibility and satisfactorily distinguished different levels of scores. Second, we chose to rate a criterion in the abstract only if the related information was present in the main text of the article. A good abstract should be in conformity with the information contained in the article and should be in as concise a format and be as informative as possible. It does not make sense to require the abstract to mention items missing from the article, even if those items should have been addressed. We realize that a poorly informative abstract could have scored well, just as it summarized a poorly informative article. However, we believe that such an abstract deserves a good score because it allows the reader to answer (albeit negatively) the most important question when reading an abstract: "Is the

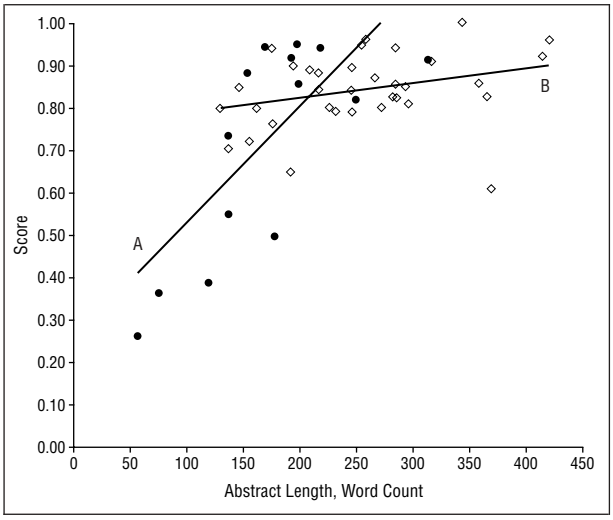

Figure 3. Correlation between abstract length and quality score. A, Regression line; solid circles represent unstructured abstracts. B, Regression line; open diamonds represent structured abstracts.

article worth reading?" Because of our rating modalities, the rating scores in this study cannot be directly compared with those calculated in the 2 other studies[4,8] using the same quality scale.

The mean quality scores for abstracts were different among the 3 journals. The assessors were not blinded to the journal titles. However, because 2 journals combined structured and unstructured abstracts, assessment bias is unlikely, although it cannot be ruled out. Also as a consequence of our rating modalities, differences in scores cannot be explained by quality differences among articles. We believe that the superiority of structured abstracts over unstructured abstracts is the main explanation for the observed differences. For the 2 journals publishing structured and unstructured abstracts, there was a significant difference in quality scores, favoring the structured format. In *J Am Acad Dermatol*, the choice of the abstract format was left to the author; in *Br J Dermatol*, 2000 was transitional between publication of primarily unstructured (January-June) and primarily structured (July-December) abstracts. A confounding bias for "better" or more compulsive authors paying more attention to the quality of their publication and, therefore, choosing structured rather than unstructured abstracts cannot be excluded. However, such a bias cannot be suspected in *Br J Dermatol*.

Structured abstracts have been widely adopted in medical journals. However, some editors of medical journals made an explicit decision not to require structured abstracts,[9] and numerous nonmedical scientific journals have not adopted structured abstracts.[10] Opponents to the structured format generally make 2 main points: (1) The widespread adoption of structured abstracts is supported by little demonstration of their superiority. (2) Structuring makes abstracts longer and less readable.[11] Our study addresses these 2 points. First, we provide evidence that structured abstracts were more informative than unstructured abstracts in 2 clinical journals, during the same period. Few studies have compared the quality of structured and unstructured abstracts. Taddio et al[4] documented improvement after the adoption of the structured format in 3 journals (*British Medical Journal*, *Canadian Medical Association Journal*, and *Journal of the American Medical Association*), but they could not exclude confounding because of the long duration of their study. Comans and Overbeke[5] could not exclude confounding by journal quality. Hartley and Benjamin[12] reported that rewritten abstracts in a structured format were more informative than original unstructured ones. Other studies[8,13-15] have assessed abstract quality, with no direct comparison of structured and unstructured abstracts. Addressing the second point made by opponents, regarding abstract length, we found that structured abstracts were longer than unstructured ones. Quality score was positively correlated to length only for unstructured abstracts. Obviously, some of them were too short to give sufficient information. Use of structured abstracts might have avoided this inconvenience by forcing authors to provide some otherwise overlooked items. We did not assess readability because it is a subjective concept, including many factors, such as typography and layout.[16] We believe that a precise piece of information is quicker to scan in a structured abstract, because the headings help in locating it. Easy access to relevant information is part of the readability. Besides modifying readability, another theoretical consequence of lengthening the abstract might be more space in which to interject inaccuracies. Discrepancies between text and accompanying abstract are known to occur.[14] However, discrepancies were often minor. We did not assess accuracy in this study; it would be interesting to compare the rate of inaccuracies relative to the format of the abstracts. Finally, as shown by our results expressed by categories, important information on the setting (essential to evaluate external validity) was lacking in 41% (9 of 22 items) and 53% (10 of 19 items) of the 2 lower-scored journals (*J Am Acad Dermatol* and *Br J Dermatol*, respectively). We believe that structured abstracts help to ensure inclusion of certain information, such as the setting, by the addition of a specific heading to a structured format.

The abstract of a medical publication is often the only part that is read. Decisions in clinical care may result from reading them alone.[1] The proposal for structuring abstracts echoed widespread enthusiasm among most editors of medical journals. The need for improvement in abstract quality has been acknowledged, and editors of *JAMA* have recently implemented quality criteria.[17] This seems to have led to improvement.[18] From 2001 onward, structured formats have been more widely adopted in all 3 dermatology journals studied herein, and consistency in the quality of abstracts should be tested in a further study. We believe that the commitment of editors is essential to improve abstract quality and that the structured abstract format can help in this task.

Accepted for publication October 25, 2002.

This study was presented as poster 2196 at the 20th World Congress of Dermatology, Paris, France, July 4-5, 2002.

Corresponding author: Alain Dupuy, MD, Service de Dermatologie, Hôpital Saint-Louis, 1, avenue Claude Vellefaux, 75010 Paris, France (e-mail: alain.dupuy@sls.ap-hop-paris.fr).

REFERENCES

1. Haynes RB, McKibbon KA, Walker CJ, Ryan N, Fitzgerald D, Ramsden MF. Online access to MEDLINE in clinical settings: a study of use and usefulness. *Ann Intern Med*. 1990;112:78-84.
2. Ad Hoc Working Group for Critical Appraisal of the Medical Literature. A proposal for more informative abstracts of clinical articles. *Ann Intern Med*. 1987; 106:598-604.
3. Haynes RB, Mulrow CD, Huth EJ, Altman DG, Gardner MJ. More informative abstracts revisited. *Ann Intern Med*. 1990;113:69-76.
4. Taddio A, Pain T, Fassos FF, Boon H, Ilersich AL, Einarson TR. Quality of nonstructured and structured abstracts of original research articles in *The British Medical Journal*, the *Canadian Medical Association Journal* and the *Journal of the American Medical Association*. *CMAJ*. 1994;150:1611-1615.
5. Comans ML, Overbeke AJ. The structured summary: a tool for reader and author [in Dutch]. *Ned Tijdschr Geneeskd*. 1990;134:2338-2343.
6. Hartley J. Are structured abstracts more or less accurate than traditional ones? a study in the psychological literature. *J Information Sci*. 2000;26:273-277.
7. Khosrotehrani K, Dupuy A, Lebbé C, Rybojad M, Morel P. Qualité des abstracts des articles publiés dans les Annales de Dermatologie. *Ann Dermatol Venereol*. 2002;129:1271-1275.
8. Narine L, Yee DS, Einarson TR, Ilersich AL. Quality of abstracts of original research articles in *CMAJ* in 1989. *CMAJ*. 1991;144:449-453.
9. Spitzer WO. The structured sonnet [editorial]. *J Clin Epidemiol*. 1991;44:729.
10. Kostoff RN, Hartley J. Structured abstracts for technical journals [letter]. *Science*. 2001;292:1067.
11. Heller MB. Structured abstracts: a modest dissent. *J Clin Epidemiol*. 1991;44: 739-740.
12. Hartley J, Benjamin M. An evaluation of structured abstracts in journals published by the British Psychological Society. *Br J Educ Psychol*. 1998;68:443-456.
13. Froom P, Froom J. Deficiencies in structured medical abstracts. *J Clin Epidemiol*. 1993;46:591-594.
14. Pitkin RM, Branagan MA, Burmeister LF. Accuracy of data in abstracts of published research articles. *JAMA*. 1999;281:1110-1111.
15. Trakas K, Addis A, Kruk D, Buczek Y, Iskedjian M, Einarson TR. Quality assessment of pharmacoeconomic abstracts of original research articles in selected journals. *Ann Pharmacother*. 1997;31:423-428.
16. Hartley J. Clarifying the abstracts of systematic literature reviews. *Bull Med Libr Assoc*. 2000;88:332-337.
17. Winker MA. The need for concrete improvement in abstract quality. *JAMA*. 1999; 281:1129-1130.
18. Pitkin RM, Branagan MA, Burmeister LF. Effectiveness of a journal intervention to improve abstract quality [letter]. *JAMA*. 2000;283:481.