

Family Physicians' Use of Medical Abstracts To Guide Decision Making: Style or Substance?

Henry C. Barry, MD, MS, Mark H. Ebell, MD, MS, Allen F. Shaughnessy, PharmD, David C. Slawson, MD, MS, and Fern Neitzke, BS

Background: Many physicians rely on the abstracts of research articles to guide their clinical decision making. This need for expediency is one basis for many journals to reformat their abstracts.

Methods: To determine whether the format of medical abstracts affects physician decision making, we surveyed family physicians in Michigan, Pennsylvania, and Virginia. All participants were members of the American Academy of Family Physicians. The survey included three case scenarios (corneal abrasion, fibromyalgia, and hyperlipidemia) followed by structured and open-ended assessments of usual management. After assessing their usual management in each scenario, the respondents were provided with an abstract of a valid research paper. The format of abstracts (unstructured, IMRAD [introduction, methods, results, and discussion], structured, and POEM [patient-oriented evidence that matters]) were randomly assigned. After reading the abstract, we assessed changes in management of the case scenario.

Results: Two hundred eighty-nine family physicians responded to the survey. At baseline, 187 (65%) of physicians patched corneal abrasions. After reading the abstract, 142 (76%) would no longer use eye patches. Two hundred forty-five (83%) of physicians did not use the combination of fluoxetine and amitriptyline for managing fibromyalgia. After reading the abstract, 179 (73%) would use combination therapy. Two hundred thirty-four (84%) of physicians used "statins" when managing hyperlipidemia. After reading the abstract, 211 (90%) would continue using statins. The format of abstract had no significant effect on physicians' decision making.

Conclusions: Whereas the format of abstract in this study had no effect on physician decision making, having valid information available in the context of a clinical scenario appeared to influence decisions. (J Am Board Fam Pract 2001;14:437–42.)

Clinical practice, regardless of discipline, is information intensive. Primary care practice, by its breadth, can be particularly so; therefore, staying current with the medical literature and applying valid research into practice is a challenge. Although thousands of medical journals generate hundreds of thousands of articles, only a small proportion is relevant to daily clinical practice. Nonetheless, this small proportion still represents an unwieldy mass of information.

During a 6-month interval Ebell and colleagues¹ reviewed 85 clinical medical journals, selecting 211 primary care relevant research articles using clinical outcomes that had the potential to change practice. These POEMs (patient-oriented evidence that matters)² represented less than 2% of the articles but still averaged 35 important articles each month, or about 8 to 9 each week. Busy practitioners lack the time to read this many articles, and many lack the skill to read them critically if they had the time.³ Consequently, many physicians take shortcuts. They subscribe to abstracting services or they skim articles, reading only the abstracts for a quick hit on the information.

Traditionally, abstracts provide a summary of the article, allowing the reader to decide whether to wade into the full text. Abstracts were originally unstructured. To make them more useful to researchers and to practitioners, it has been proposed that journals change to a more structured format.⁴ The number of journals using structured abstracts has subsequently risen steadily.⁵ Although the use

Submitted, revised, 28 February 2001.

From the Department of Family Practice (HCB, MHE, FN), College of Human Medicine, Michigan State University, East Lansing, Mich; the Harrisburg Family Practice Center (AFS), Harrisburg, Pa; and the University of Virginia Health Science Center (DCS), Charlottesville, Va. Address reprint requests to Henry C. Barry, MD, MS, Department of Family Practice, College of Human Medicine, B-104B Clinical Center, Michigan State University, East Lansing, MI 48824-1315.

A preliminary version of this article was presented in abstract form at the North American Primary Research Group Annual Scientific Meeting in Montreal, November 1998.

of structured abstracts might have made it somewhat easier for researchers to retrieve pertinent studies,⁶ their effect on clinicians has heretofore been unreported. It is clear from other literature that the method of reporting research data has a direct influence on physician decision making.⁷⁻⁹ We therefore decided to explore the possible effect of different abstract formats on physician decision making through a survey of family physicians.

Methods

Participants

We randomly selected participants from the 5,229 active members of the American Academy of Family Physicians residing in Michigan, Pennsylvania, and Virginia. We randomly took 200 names from each state and sent them a survey instrument. The study methods and survey instruments were approved by the Michigan State University Committee on Research Involving Human Subjects.

Survey

The survey instrument included a request for basic demographic information, type of practice, and teaching activity. The survey instrument also assessed the physicians' rating of the usefulness and frequency of use of various sources of medical information. It further assessed a number of attitudes and beliefs about evidence-based medicine. We included in each survey instrument three case scenarios. In each scenario, the physician was asked, "The last time you treated a patient similar to this one, what did you do?" Several specific options, as well as an open-ended question about management issues not addressed in the forced choices, followed. After responding to each case, we provided an abstract of a methodologically valid article pertinent to the cases. We asked the participants to read the abstract and then assessed on a 5-point Likert scale the likelihood that their future management would change.

Case Scenarios

The scenarios and the accompanying abstracts were selected to represent three different possibilities. Assuming that rational physicians will adopt change based on valid research results, we selected articles and case scenarios that would result in negative change (abandon an ineffective approach), positive change (adopt a new effective approach), or

no change (research confirms that current practice is effective). The scenarios and abstracts included corneal abrasion for the negative change example,¹⁰ fibromyalgia for the positive change example,¹¹ and hyperlipidemia for the no change example.¹²

Abstracts

We developed four different abstract formats for each article: unstructured, IMRAD (introduction, methods, results, and discussion), structured (background, study design, population, data collection, results, conclusion), and a POEM (patient-oriented evidence that matters) review. A POEM review is a critical appraisal that begins with a focused clinical question, summarizes the key findings and their validity, and provides a clinical bottom line. Stratified by state, we randomly assigned each physician to receive one of the abstract formats.

Procedures

Each survey instrument was sent by mail. To improve response rate, we offered by lottery a \$100 gift certificate to one respondent from each state. We tracked nonresponders and sent second and third mailings at 3- to 4-week intervals.

Data Analysis

We used the Sample Size Expert software¹³ for sample size calculations and assumed $\alpha = 0.05$ and $\beta = 0.20$ for a two-tailed estimate. To determine an absolute difference of 20% in the physicians changing their behavior (75% vs 55%) would require 88 per group. To determine an absolute difference of 15% in the physicians changing their behavior (75% vs 60%) would require 152 per group. To detect a difference of 1 point on a Likert scale from 1 to 5, treating this variable as continuous and assuming a standard deviation of 2 points, would require 63 patients per group. Given an estimated response of 50% to 60%, an initial sample of 600 patients would provide a final sample of 300 to 360, or 75 to 90 per abstract type.

We performed data analyses with SAS (Cary, NC). Differences in willingness to change among the different abstract formats were compared using the Fisher exact test. We collapsed the five willingness-to-change categories into three (change, uncertain, no change). We stratified each scenario into two groups – those who before reading the abstract were already using a specific therapy and those who were not. For each stratum, we com-

Table 1. Effect of Abstract Format on Willingness to Change Practice of Patching Corneal Abrasions.

| Abstract Format | Not Willing to Change | | Undecided | | Willing to Change | |
|---------------------------|-----------------------|---------|-----------|---------|-------------------|---------|
| | Number | Percent | Number | Percent | Number | Percent |
| Not patching at baseline* | | | | | | |
| Unstructured | 23 | 82.14 | 1 | 3.57 | 4 | 14.29 |
| IMRAD | 15 | 78.95 | 1 | 5.26 | 3 | 15.79 |
| Structured | 28 | 93.33 | 1 | 3.33 | 1 | 3.33 |
| POEM | 20 | 90.91 | 0 | 0 | 2 | 9.09 |
| Patching at baseline† | | | | | | |
| Unstructured | 8 | 15.38 | 7 | 13.46 | 37 | 71.15 |
| IMRAD | 13 | 24.53 | 1 | 1.89 | 39 | 73.58 |
| Structured | 3 | 7.69 | 6 | 15.38 | 30 | 76.92 |
| POEM | 5 | 11.63 | 2 | 4.65 | 36 | 83.72 |

IMRAD—introduction, methods, results, and discussion.

POEM—patient-oriented evidence that matters.

*n = 99 (34.6%) Fisher exact test, $P = .6$.

†n = 187 (65.4%) Fisher exact test, $P = .06$.

pared willingness to change across the four abstract types.

Results

We mailed survey instruments to 600 physicians. After two additional mailings, we had 299 responders, of which 289 were usable (not blank). We were informed that one physician was dead. The mean age of the respondents was 44.7 years (95% confidence interval [CI] 43.6, 45.7). The physicians, on average, had been in practice for 14.5 years (95% CI 13.3, 15.6). Most (74%) were men and allopathic physicians (88%). More than 90% of the respondents were board certified, and nearly 90% had completed a residency. Although most practiced in the suburbs (147, 52.7%), many were rural practitioners (85, 30.5%). Most of the responders described their practice as a single-specialty group (171, 62.4%). One hundred thirty-five (45.8%) responders said they had no formal teaching affiliations.

We received responses from 109 physicians in Michigan, 89 from Pennsylvania, and 99 from Virginia. We received responses from 83 (28.0%) physicians randomized to receive unstructured abstracts, 76 (25.6%) who received IMRAD abstracts, 71 (23.9%) who received structured abstracts, and 67 (22.6%) who received a POEM.

Corneal Abrasions

Before reading the abstract, 187 of 286 respondents (65.4%) patched corneal abrasions and 99 (34.6%)

did not. Among those who patched abrasions, 142 (75.9%) were willing to change after reading the abstract. For those physicians not using eye patches, they were unlikely to use them in the future. There was no significant difference among the abstract formats (Table 1).

Fibromyalgia

Before reading the abstract, 245 of 286 respondents (85.7%) did not use the combination of amitriptyline and fluoxetine when treating fibromyalgia. Among these physicians, 179 (73.1%) were willing to use the combination in the future. There was no significant difference among the abstract formats (Table 2).

Hyperlipidemia

Two hundred thirty-four (83.6%) of the responders used lipid-lowering agents before reading the abstract. This group was unlikely to change (211, 90.1%). There was no significant difference among the abstract formats (Table 3). Interestingly, those who were not using lipid-lowering agents were still unlikely (30/46, 65%) to start using them!

Discussion

We believe that this study is the first to attempt a formal assessment of the effect of different abstract formats on physician decision making. In this study, family physicians responding to a survey instrument that assessed their management of three common clinical scenarios did not appear to be

Table 2. Effect of Abstract Format on Willingness to Use Combination of Amitriptyline and Fluoxetine to Treat Fibromyalgia.

| Abstract Format | Not Willing to Change | | Undecided | | Willing to Change | |
|-----------------------------|-----------------------|---------|-----------|---------|-------------------|---------|
| | Number | Percent | Number | Percent | Number | Percent |
| Not using combined therapy* | | | | | | |
| Unstructured | 12 | 17.91 | 3 | 4.48 | 52 | 77.61 |
| IMRAD | 8 | 12.50 | 7 | 10.94 | 49 | 76.56 |
| Structured | 13 | 19.70 | 7 | 10.61 | 46 | 69.70 |
| POEM | 8 | 16.67 | 8 | 16.67 | 32 | 66.67 |
| Using combined therapy† | | | | | | |
| Unstructured | 9 | 69.23 | 0 | 0 | 4 | 30.77 |
| IMRAD | 9 | 90.00 | 0 | 0 | 1 | 10.00 |
| Structured | 1 | 25.00 | 0 | 0 | 3 | 75.00 |
| POEM | 7 | 50.00 | 0 | 0 | 7 | 50.00 |

IMRAD—introduction, methods, results, and discussion.

POEM—patient-oriented evidence that matters.

*n = 245 (83.3) Fisher exact test, $P = .4$.

†n = 49 (16.7) Fisher exact test, $P = .07$.

influenced by the format of information. Interestingly, it appeared that merely providing valid summary information in the context of the care of the patient affected their management. Even more interestingly, positive change (starting a new approach to therapy) and negative change (abandoning an old approach) occurred with equal probability. We were surprised that the responding physicians were willing to change at such a high rate.

These findings are unanticipated based on other types of studies. Bucher et al⁷ and Cranney and Walley⁸ each showed that when research results are presented as relative risk or relative risk reduction, physicians were more impressed (and therefore more likely to use them in practice) than if they were presented as absolute risk or absolute risk reduction. Naylor et al⁹ similarly found that results presented as relative risk more profoundly influ-

Table 3. Effect of Abstract Format on Willingness to Use Lipid-Lowering Agents.

| Abstract Format | Not Willing to Change | | Undecided | | Willing to Change | |
|----------------------------------|-----------------------|---------|-----------|---------|-------------------|---------|
| | Number | Percent | Number | Percent | Number | Percent |
| Not using lipid-lowering agents* | | | | | | |
| Unstructured | 9 | 69.23 | 2 | 15.38 | 2 | 15.38 |
| IMRAD | 9 | 69.23 | 0 | 0.00 | 4 | 30.77 |
| Structured | 5 | 50.00 | 3 | 30.00 | 2 | 20.00 |
| POEM | 7 | 70.00 | 0 | 0.00 | 3 | 30.00 |
| Using lipid-lowering agents† | | | | | | |
| Unstructured | 61 | 93.85 | 1 | 1.54 | 3 | 4.62 |
| IMRAD | 56 | 93.33 | 2 | 3.33 | 2 | 3.33 |
| Structured | 49 | 87.50 | 0 | 0.00 | 7 | 12.5 |
| JFP POEM | 45 | 84.91 | 4 | 7.55 | 4 | 7.55 |

IMRAD—introduction, methods, results, and discussion.

POEM—patient-oriented evidence that matters.

*n = 46 (16.4%), Fisher exact test, $P = .4$.

†n = 234 (83.6%), Fisher exact test, $P = .1$.

Table 4. Issues With Applying Research Results Into Practice.

| |
|---|
| 1. Asking the right question* |
| 2. Finding the evidence* |
| 3. Setting vs patient characteristics |
| 4. Representatives of a trial population |
| 5. Framing of risk |
| 6. Acting on the evidence* |
| 7. Heterogeneity |
| 8. Baseline risk |
| 9. Tolerability |
| 10. Likelihood of being helped or harmed* |
| 11. Subgroup analyses |

*Items addressed in this study.

enced physicians than when presented as either absolute risk or numbers needed to treat.

We did not assess any of the potential barriers to applying the results of research in general practice. Fahey¹⁴ has summarized many of these barriers (shown in Table 4). Because we designed the case scenarios so that the patients resembled the patients included in the studies, we minimized the concerns about the external validity of the study. Additionally, we selected articles with valid methods to minimize the effect of internal validity. Finally, the scenarios represented problems that were either not controversial or for which therapies have not been very helpful, which might help account for the enthusiasm for change on the part of the respondents. One can only wonder what would occur had the cases been more controversial, such as the use of antibiotics for otitis media.

The low response rate in this study, while consistent with that found in other surveys of physicians, is a limitation and raises the potential for responder bias. Whether those physicians not responding might be more or less receptive to change or to the influence of research results is an important issue. An additional limitation of this study is the reliance on self-reported intention to change. Whether performance on case scenarios truly represents what occurs in real life is subject to question. Holmes et al¹⁵ compared internal medicine residents' responses to case scenarios with chart audits and billing records. They concluded that case scenarios might better reflect the patterns of use of information in a well-defined problem. Clearly, medical decision making and application of information are complex, and more rigorous assess-

ment (such as direct observation) is required to confirm what occurs in real practice.

The data presented in Tables 1 and 3 raise the possibility of a type II error (missing potentially significant differences as a result of insufficient numbers of participants). Our a priori sample size estimates were intended to detect modest differences in behavior changes. The differences we detected were smaller than predicted but appear to be clinically unimportant.

Close inspection of the data reveals a paradox. Several physicians reported their baseline approach to managing a case scenario was consistent with the data from the research. Nonetheless, after reading the abstract, they would have inappropriately changed their approach. The information might have been confusing and therefore incorrectly interpreted and applied. This paradox is not easily explained by the available data and should be addressed through a series of qualitative studies.

The goal of providing greater structure and easier access to using summary information is admirable and conceptually sound. Haynes et al¹⁶ report that more informative abstracts of this kind can facilitate peer review before publication, assist clinical readers to find articles that are both scientifically sound and applicable to their practices, and allow more precise computerized literature searches. Our study, despite its limitations, challenges the notion that clinicians will find highly structured abstracts more useful. Although we have not addressed the inaccuracies often found in abstracts,^{17,18} the means to assist busy clinicians to apply the results of valid research in the care of their patients is an area that clearly needs more rigorous evaluation. Rather, this study suggests that providing relevant and valid information in the context of a patient encounter has a high likelihood of influencing the care provided.

The authors thank Deb Richardson for her work in disseminating and tracking the survey instruments and for final manuscript preparation.

References

1. Ebell MH, Barry HC, Slawson DC, Shaughnessy AF. Finding POEMs in the medical literature. *J Fam Pract* 1999;48:350–5.
2. Shaughnessy AF, Slawson DC, Bennett JH. Becoming an information master: a guidebook to the medical information jungle. *J Fam Pract* 1994;39:489–99.

3. Slawson DC, Shaughnessy AF. What clinical information do doctors need? Few doctors are expert at evaluating information. *BMJ* 1997;314:904.
4. A proposal for more informative abstracts of clinical articles. Ad Hoc Working Group for Critical Appraisal of the Medical Literature. *Ann Intern Med* 1987;106(4):598–604.
5. Harbourt AM, Knecht LS, Humphreys BL. Structured abstracts in MEDLINE, 1989–1991. *Bull Med Libr Assoc* 1995;83:190–5.
6. Wilczynski NL, Walker CJ, McKibbin KA, Haynes RB. Preliminary assessment of the effect of more informative (structured) abstracts on citation retrieval from MEDLINE. *Medinfo* 1995;8(Pt 2):1457–61.
7. Bucher HC, Weinbacher M, Gyr K. Influence of method of reporting study results on decision of physicians to prescribe drugs to lower cholesterol concentration. *BMJ* 1994;309:761–4.
8. Cranney M, Walley T. Same information, different decisions: the influence of evidence on the management of hypertension in the elderly. *Br J Gen Pract* 1996;46:661–3.
9. Naylor CD, Chen E, Strauss B. Measured enthusiasm: does the method of reporting trial results alter perceptions of therapeutic effectiveness? *Ann Intern Med* 1992;117:916–21.
10. Kaiser PK. A comparison of pressure patching versus no patching for corneal abrasions due to trauma or foreign body removal. Corneal Abrasion Patching Study Group. *Ophthalmology* 1995;102:1936–42.
11. Goldenberg D, Mayskiy M, Mossey C, Ruthazer R, Schmid C. A randomized, double-blind crossover trial of fluoxetine and amitriptyline in the treatment of fibromyalgia. *Arthritis Rheum* 1996;39:1852–9.
12. Shepherd J, Cobbe SM, Ford I, et al. Prevention of coronary heart disease with pravastatin in men with hypercholesterolemia. West of Scotland Coronary Prevention Study Group. *N Engl J Med* 1995;333:1301–7.
13. Ebell MH, Neale AV, Hodgkins BJ. An expert system for the calculation of sample size. *Fam Pract Res J* 1994;14:183–95.
14. Fahey T. Applying the results of clinical trials to patients to general practice: perceived problems, strengths, assumptions, and challenges for the future. *Br J Gen Pract* 1998;48:1173–8.
15. Holmes MM, Rovner DR, Rothert ML, Schmitt N, Given CW, Ialongo NS. Methods of analyzing physician practice patterns in hypertension. *Med Care* 1989;27:59–68.
16. Haynes RB, Mulrow CD, Huth EJ, Altman DG, Gardner MJ. More informative abstracts revisited. *Ann Intern Med* 1990;113:69–76.
17. Pitkin RM, Branagan MA, Burmeister LF. Accuracy of data in abstracts of published research articles. *JAMA* 1999;281:1110–1.
18. Winker MA. The need for concrete improvement in abstract quality. *JAMA* 1999;281:1129–30.