Quantifying the Literature of Computer-aided Instruction in Medical Education

Mark D. Adler, MD, and Kevin B. Johnson, MD, MS

ABSTRACT

Purpose. To characterize the literature about computer-aided instruction (CAI) as it relates to medical education.
Method. A descriptive study using the Medline and ERIC databases, reviewing articles pertaining to CAI for medical professionals. The literature was evaluated with respect to type of article, journal, language, and year of publication.
Results. The search yielded 2,840 citations, 92% of which were in English. The number of journals with at least one citation was 747. Less than 5% of the 5,147 authors had three or more articles published in the CAI literature. Of the citations with abstracts, 60% were demonstrations of a CAI application, 11% were media-comparative studies, and 13% were analyses of the CAI field. While the pace of article publication increased markedly over time, the percentages of article types did not change

significantly over time. Less than 10% of CAI articles appeared in core medical journals.
Conclusion. Medical CAI is an increasingly popular topic of research and publication. However, these studies appear in journals with smaller circulations, are predominantly demonstration articles, and are generally written by authors with two or fewer publications. Evaluation articles remain less common. A series of analytic articles has appeared offering substantive suggestions for better research design. These suggestions appear to have gone unheeded. CAI investigators need be more aware of the gaps in the existing body of CAI publications, and the inherent difficulties of this type of research, if this literature is to move beyond this early stage of development. *Acad. Med.* 2000;75:1025–1028.

Medical educators face the challenge of providing students at different training levels with comprehensive educational content that covers an enormous knowledge base. One can understand why computer technology, a teaching tool that supports this task, has been embraced enthusiastically. Since the advent of computing systems, educators

Dr. Adler is a senior fellow in the Division of Pediatric Emergency Medicine and **Dr. Johnson** is assistant professor, Division of General Pediatrics, both at the Johns Hopkins University School of Medicine, Baltimore, MD.

Correspondence and requests for reprints should be addressed to Dr. Adler, Division of Pediatric Emergency Medicine, The Johns Hopkins Hospital, CMSC 144, 600 North Wolfe Street, Baltimore, MD 21287; telephone: (410) 955-6143; fax: (410) 614-7339.
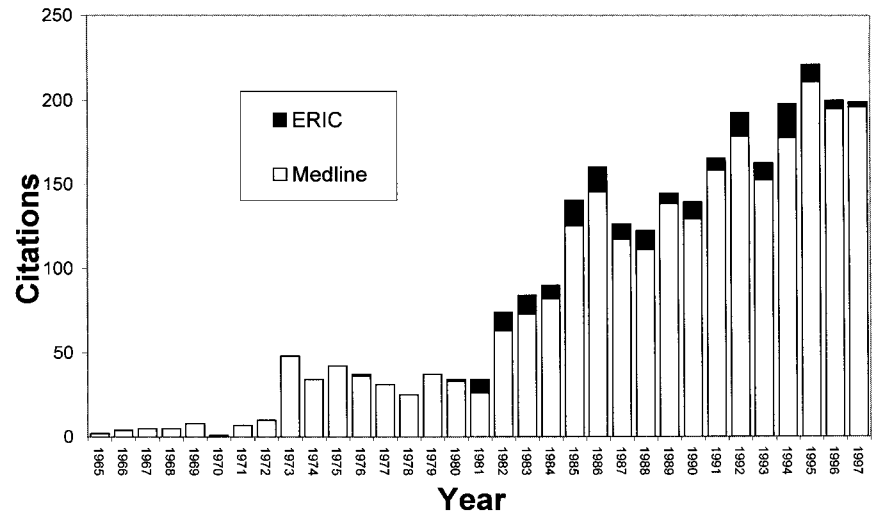
have used computer-aided instruction (CAI) to educate medical students and professionals in a broad variety of settings.

Published descriptions of medical CAI began to appear in the 1960s and have been categorized by three types of evaluation.[1] *Demonstration* articles describe, but do not evaluate CAI applications. *Media-comparative* articles evaluate a CAI application or applications against other teaching media or other CAI applications. Finally, *analytic articles* evaluate either an aspect of CAI (e.g., videodisc software for pathology instruction) or the literature as a whole.

A number of critical evaluations of the medical CAI literature have been published.[1–3] The authors have noted

that (1) demonstration studies have dominated the literature and (2) when comparisons have been made with traditional educational formats, confounders specific to the comparison of educational media have rarely been addressed. Although anecdotal evidence suggests there has been an increase in the use of instructional computing, and despite the emergence of articles that inform readers about the appropriate ways to evaluate CAI, comments from experts suggest that the proportions of evaluation articles have not changed. This study aims to quantify the relative proportions of CAI articles published, and how they have changed over time, in an effort to investigate the accuracy of the impressions cited in these critical evaluations.

METHOD

In August 1998, we conducted a search for citations in the Medline[4] and Educational Resources Information Center (ERIC)[5] databases using combinations of the words "computer," "aided," "assisted," "based," "instruction," and "teaching" and the date range 1966–1998. We classified the citations by language, year, and source of publication and calculated the number of citations per journal title and the number of articles published per year.

Citations with abstracts were categorized using information obtained from the title and abstract only. The following five categories—demonstration, media-comparative, analytic, other, and not applicable—were based on the classification scheme adapted from Friedman.[1] The media-comparative citations were further divided into two subcategories: CAI-versus-traditional, which describes articles that compare a CAI application with a lecture or other more common teaching method; and CAI-versus-CAI, which describes a direct comparison of two or more CAI applications. The classification "other" consisted of citations that did not fit in the three main categories but still pertained to CAI use in medical education. Articles unrelated to CAI for medical education (e.g., CAI for patient education only or for use as testing instruments) were categorized as "not applicable" and were excluded from the analysis. The first author (MA) made all categorization assignments. The second author (KJ) reviewed a 10% random sample of articles to assess interrater reliability.

RESULTS

The literature search using the Medline and ERIC databases yielded 2,840 citations from a total of 747 journals; 2,763 were individual citations (not compilations). Figure 1 illustrates the numbers of citations per year for the period 1966–1998. A majority of the articles, 2,540 (92%), were in English and 1,084 (43%) of these came from 25 (3.3%) of the journals. Of the 5,147 unique authors, 85% had only one citation, while 241 (4.7%) had three or more citations.

To address the accessibility of this literature to the general medical readership, we compared our search's results with a list of "core" journals compiled by an association of medical librarians[6] to establish a minimum complement of journals for a small medical library. In such a library, only 205 of the 2,540 CAI articles in English would be available. Consistent with this finding is the fact that only 23 citations have appeared in *JAMA*, the *Lancet*, or the *New England Journal of Medicine*.

Of the 2,763 total articles, 1,498 (54%) of their citations included abstracts. Of these, we excluded 427 citations (28.5%) for being unrelated to medical CAI, leaving 1,071 citations for analysis. Of these, 60% were demonstration articles, 11% were media-comparative studies, 13% were analytic ar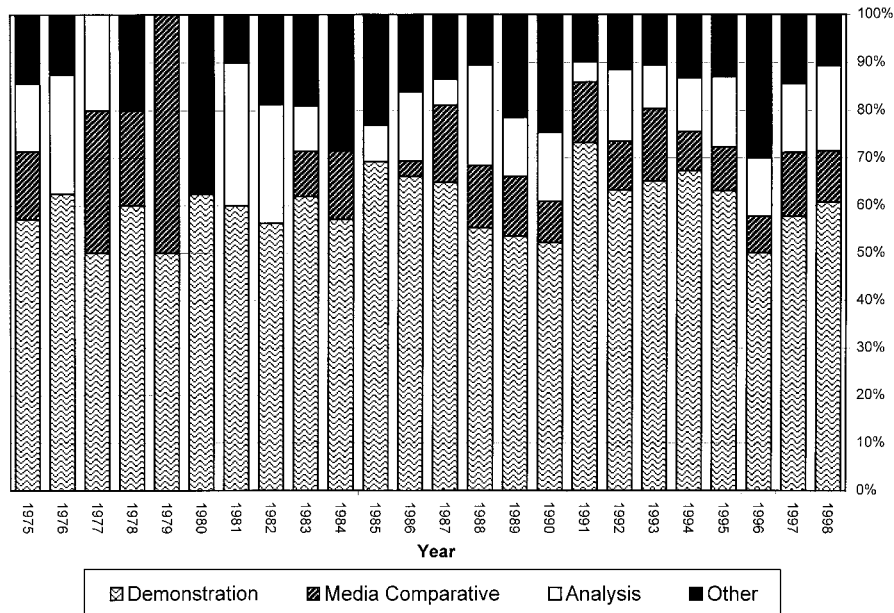ticles, and 16% were other. CAI-versus-CAI articles comprised 1% of the citations, and 9% of the media-comparative articles. The comparison of the authors' categorizations of the articles had an unweighted Cohen's kappa of 0.49, suggesting fair overall agreement among raters. In assigning articles into demonstration versus all other categories, the overall agreement was 84.7% with a kappa of 0.69.

Figure 2 illustrates the relative proportions of article types over time and it shows there was a consistent predominance of demonstration articles in the period spanning 1975–1997.

DISCUSSION

This study quantifies how the medical CAI literature did and did not change from 1966 to 1998. We found an increasing rate of production of medical CAI articles and a marked, continuing predominance of demonstration articles, which represented more than half of all the articles published per year since 1975 (see Figure 2).

There are a number of possible reasons why demonstration articles continue to dominate the medical CAI literature.

Figure 1. Numbers of computer-aided instruction citations found in a search of Medline and ERIC databases for using combinations of the terms "computer," "aided," "assisted," "based," "instruction," and "teaching" and the period 1966–1998.

Figure 2. Relative proportions of the types of computer-assisted instruction articles published from 1975 to 1998. Article types (overall percentage) were demonstration (60%), media comparative (11%), analysis (13%), and other (16%).

Demonstration articles are more easily produced than comparative studies. Without the need for a control group or the large number of participants required to develop statistical power, demonstrations can be conducted with fewer resources than more ambitious comparative studies.

The medical readership may be unaware of the ubiquitous nature of these articles. While CAI crosses medical fields and specialties, only a limited number of CAI articles appear in any one field's literature. Urologists publish articles about CAI for urologic education, anesthesiologists for their field, and so on. Hence, if the CAI literature in a particular medical field is not well done and CAI articles appear infrequently in core journals (as we have noted), no specific audience will encounter more than a few articles. A reader who does not actively seek out this topic will be unlikely to encounter a significant portion of the literature or the critical analyses and will remain unaware of the prevalence of demonstration articles. While this publication pattern is seen with other topics that bridge multiple fields, in this case it may contribute to the lack of maturation of the medical CAI literature.

Many authors are inexperienced medical CAI investigators. If we accept that most research endeavors begin with smaller demonstration studies, then those 85% of authors who publish only one work will be likely to publish a demonstration study. Whether these authors intended to stop after one publication, were unable to proceed to more mature investigation for lack of funds or infrastructure, or will publish more in the future is beyond the scope of this study.

Comparative studies of educational media are difficult to do well, given the implicit threat to internal validity when comparing educational media.[2] In comparing medical CAI with traditional media, such as lectures or textbooks, researchers seek to validate CAI and justify the development, use, and purchase of CAI materials. Media-comparative studies use as an outcome measure some comparison of educational improvement between control and experimental groups. Since the typical medical CAI program takes full advantage of the computing platform by incorporating video, audio, hyperlinked text, and other interactive features, a comparative study's two groups receive different content, which confounds any comparison of the two media. If CAI offers novel medical educational tools that cannot be replicated by other methods, then, as Friedman points out, there can be no true comparison group and the typical media-comparative study becomes "logically impossible."[1] Almost ten years ago, Keane summarized this point: "Any contribution to be derived from additional CAI–non-CAI studies is doubtful."[3]

Editors of medical journals may be unaware of the CAI literature. If journal editors are unfamiliar with the medical CAI, this promotes two unfortunate trends: (1) publication of more demonstration articles and (2) publication of comparative studies with "positive" results. This well-described publication bias[7] is especially problematic for this topic area given that evidence favoring CAI over other educational media would be confounded by differences in content, as discussed above.

Limited availability of research funding may have contributed to the lack of change in CAI literature. Demonstration articles are often about smaller-scale projects that may be easier to fund. The cost of small projects has fallen as hardware and software costs moderate and departments become increasingly well equipped. Comparative projects usually require more substantial resources and would be disproportionately discouraged if funding were a problem.

Implications

The impetus for conducting this study came from our reading of the "literature about the literature about medical CAI." Hagler and Knowlton invoked

Santayana when lamenting how technologic advances in educational tools seem doomed to follow this pattern of publication, citing a similar pattern that followed the introduction of motion pictures. We have quantified the trends that have been alluded to in the analytic literature. These same commentators offer some suggestions for better avenues of research. Essentially, they suggest that proponents of medical CAI stop trying to demonstrate, quantitatively, that CAI is a superior educational tool. Instead, the focus needs to be on improving the value of CAI research:

- CAI-to-CAI comparison—evaluations of how different CAI approaches compare with or complement one another. These studies suffer less from internal validity problems. Friedman et al. evaluated two different interfaces for presenting the same content.[8] Analysts, while critical of CAI–non-CAI studies, clearly point to CAI-to-CAI studies as important in improving approaches to application design.[1,3]
- Economic analyses—evaluations of the relative values of different teaching media from the standpoint of direct and indirect costs and efficiency from either the instructor's or the learner's standpoint(s). By justifying these applications and technologies, based on their economic value, the concerns of faculty and administrators responsible for implementing and supporting these efforts can be addressed.
- Curricular development—evaluations of how CAI integrates into the larger medical curriculum of a department or medical school.
- CAI in different learning settings—evaluations of how CAI is best incorporated into different learning environments and for use with different types of students.

A few changes may improve the quality of medical CAI literature. Organizations such as the American Medical Informatics Association and the Association of American Medical Colleges should promote education about how to conduct and obtain funding for better CAI studies. These efforts will hopefully result in better studies submitted for publication. In turn, increased awareness by journal editors of what constitutes meaningful contributions to this field would result in the publication of more useful studies. Most important, an effort to encourage publication of CAI research in representative journals is necessary. As more readers are exposed to the better models of research, we may finally see a change in the character of this literature.

Limitations

This study has a number of limitations. The citations, collected from two medical electronic databases, represent only a portion of the literature. The categorization scheme relied on citation abstracts that appeared in Medline after 1974 only. The trends seen in the citations with abstracts may not be generalizable to all medical CAI articles. Finally, the categorization approach that we employed is subjective. However, our interrater agreement was acceptable, which suggests our categorization scheme would be reproduced by others reviewing this literature.

The continued publication of demonstration articles will add little to the published body of work. The likelihood that only a small portion of the work reaches a significant proportion of readers is also of great concern. In the future, authors and editors should be encouraged to publish exemplary research projects in venues with the largest possible audience. Once the medical community at large begins to see more and better CAI articles, meaningful CAI research will become the norm rather than the exception.

REFERENCES

1. Friedman CP. The research we should be doing. Acad Med. 1994;69:455–7.
2. Hagler P, Knowlton P. Invalid assumption in CBI comparison research. J Comput Based Instruction. 1987;14(3):84–8.
3. Keane DR, Norman GR, Vickers J. The inadequacy of recent research on computer-assisted instruction. Acad Med. 1991;66:444–8.
4. Ovid Version 7.7. National Library of Medicine, National Center for Biotechnology Information. 1997.
5. Educational Resources Information Center [Producer]. ERIC (Summer 1998 Release). Silver Platter. 1998.
6. Brandon AN, Hill DR. Selected list of books and journals for the small medical library. Bull Med Libr Assoc. 1997;85(2):111–35.
7. Dickersin K. The existence of publication bias and risk factors for its occurrence. JAMA. 1990;263:1385–9.
8. Friedman CP, Wildemuth BM, Muriuki M, et al. A comparison of hypertext and Boolean access to biomedical information. In: Cimino JJ (ed). Proceedings of the American Medical Informatics Association Annual Fall Symposium, Washington, DC, 1996:2–6.