

# Original Investigations

# JAMIA

Research Paper ■

## The Structure of Medical Informatics Journal Literature

---

THEODORE A. MORRIS, MLS, KATHERINE W. McCAIN, PhD

**Abstract** **Objective:** Medical informatics is an emergent interdisciplinary field described as drawing upon and contributing to both the health sciences and information sciences. The authors elucidate the disciplinary nature and internal structure of the field.

**Design:** To better understand the field's disciplinary nature, the authors examine the intercitation relationships of its journal literature. To determine its internal structure, they examined its journal cocitation patterns.

**Measurements:** The authors used data from the Science Citation Index (SCI) and Social Science Citation Index (SSCI) to perform intercitation studies among productive journal titles, and software routines from *SPSS* to perform multivariate data analyses on cocitation data for proposed core journals.

**Results:** Intercitation network analysis suggests that a core literature exists, one mark of a separate discipline. Multivariate analyses of cocitation data suggest that major focus areas within the field include biomedical engineering, biomedical computing, decision support, and education. The interpretable dimensions of multidimensional scaling maps differed for the SCI and SSCI data sets. Strong links to information science literature were not found.

**Conclusion:** The authors saw indications of a core literature and of several major research fronts. The field appears to be viewed differently by authors writing in journals indexed by SCI from those writing in journals indexed by SSCI, with more emphasis placed on computers and engineering versus decision making by the former and more emphasis on theory versus application (clinical practice) by the latter.

■ **JAMIA.** 1998;5:448–466.

---

Affiliations of the authors: University of Cincinnati, Cincinnati, Ohio (TAM); Drexel University, Philadelphia, Pennsylvania (KWM).

This work was supported in part by a Title II fellowship from the U.S. Department of Education (TAM).

Correspondence and reprints: Theodore A. Morris, University of Cincinnati Medical Center Academic Information Technology and Libraries, 231 Bethesda Avenue, P.O. Box 670574, Cincinnati, OH 45267-0574. e-mail: <theodore.morris@uc.edu>.

Received for publication: 11/25/97; accepted for publication: 4/30/98.

Medical informatics draws on, and contributes to, multiple disciplines in the health sciences and information sciences.<sup>1</sup> While many definitions of the field can be found, most share two characteristics: reference to health sciences, biomedicine, and the healing arts; and reference to the use of information management techniques and technologies in support of those pursuits.<sup>2</sup> For example, Lincoln and Korpman<sup>3</sup> considered medical informatics to be “the hybrid child of medicine and those logical sciences that are suggested by computer technology.” Shortliffe<sup>4</sup> reiterated the infor-

mation side of medical informatics, noting a National Library of Medicine statement that "Medical informatics is the study of biomedical information, data, and knowledge—their storage, retrieval, and optimal use for problem solving and decision making." Greenes and Siegel<sup>5</sup> saw medical informatics as "the field concerned with the cognitive, information processing, and information management tasks of medical and health care, and biomedical research, and with the application of information sciences and technology to those tasks." Lincoln<sup>6</sup> later expanded the list of fields linked to medical informatics, noting that it "draws on various branches of logic, mathematics, computer science and behavioral science as well as focused disciplines such as decision theory, artificial intelligence, systems analysis, and industrial psychology." Blois<sup>7</sup> discussed the structure of the field, noting that "Medical information science (medical informatics) may be viewed as a discipline with several subfields, some of which (e.g., biostatistics) are already in the standard medical curriculum, while others (e.g., inference methods, decision theory) are not." From these definitions one expects to find a great deal of borrowing from other disciplines in medical informatics research and practice.

Klein<sup>8</sup> suggests four basic types of interdisciplinary interaction, and through its history medical informatics has shown them all: borrowing, solving of problems, increased consistency of subjects or methods, and emergence of an interdiscipline. For example, biomedical and health disciplines are by their nature problem-oriented, and medical informatics has sought to establish rigorous research questions.<sup>9–12</sup> Many cross-disciplinary informatics curricula and several professional societies have developed in the last 30 years. Medical informatics is now seen by most people in the field as a discipline of its own.<sup>9,11–14</sup>

Because medical informatics is an interdisciplinary field, however, it offers special challenges for study. New disciplines spring from old ones when paradigms followed by existing disciplines no longer meet the needs of their researchers.<sup>15</sup> In contrast, interdisciplines are formed by merging ideas from existing disciplines.<sup>8</sup> At first, an interdisciplinary field will maintain, through their literatures, close links with the core specialties from which it has arisen. Eventually, the new field will develop a core literature of its own, which can be consulted to maintain contact with major movements in the field.<sup>16</sup> However, the demarcation where the original fields "leave off" and the new one starts may be unclear. Greenes and Siegel<sup>5</sup> point out that "emerging research and development fields—such as medical informatics—present special

problems for NLM [National Library of Medicine] inasmuch as they are characteristically in a state of flux and there is a lack of generally agreed upon definitions of the boundaries and structures of the field. Moreover, as is characteristic of work that is highly interdisciplinary, publications may appear in many fields not necessarily connected with medicine or health."

Our goal in the research reported here was to take the first steps toward understanding how the interdiscipline of medical informatics is structured and how it relates to neighboring fields.

### **Studying Interdisciplines through Citation Analysis**

Much can be learned about the context and direction of a field through studies of its literature and the use made of the literature by its practitioners. At least three approaches may be employed: productivity analyses (where medical informatics researchers frequently publish), user surveys and use statistics (what medical informatics researchers read and the literature they value), and citation studies (what medical informatics researchers choose to reference in their writings). We have chosen the last approach to consider medical informatics based on the way its knowledge markers are built upon to create new knowledge. We explored the patterns of citations within the core journal literature of medical informatics to identify its internal structure and relationships to other fields. Specifically, we examined the links between journals made when an article in one journal cites an article in another journal (or in a previous issue of its own back run) and when articles from two different journals are included in the same reference list. The first relationship is termed "intercitation"; the second, "cocitation."

Citations may be given for many reasons.<sup>17,18</sup> When authors publish articles citing other articles, they establish evidence of the importance they perceive the cited articles have in supporting their own work. When an author cites two different articles in the same paper, these two articles are linked as being together of some importance to the author's thesis. Even citations given to illustrate opposing sides of a discussion are linked through the arguments of the (co)citing article, serving as concept symbols for the opposing views.<sup>19</sup> This perceived link among articles can be extended to the journals that publish those articles. As the cocitation relationships build within a field's scholarly output, the relative importance of particular journals to one another in supporting different re-

search topics can be ascertained. Research themes and major subject orientations within the field can be determined through analysis of the citations that journals receive jointly. These patterns are often elucidated by examination of results of cluster analysis, factor analysis (especially principal components analysis), and multidimensional scaling of the cocitation data.

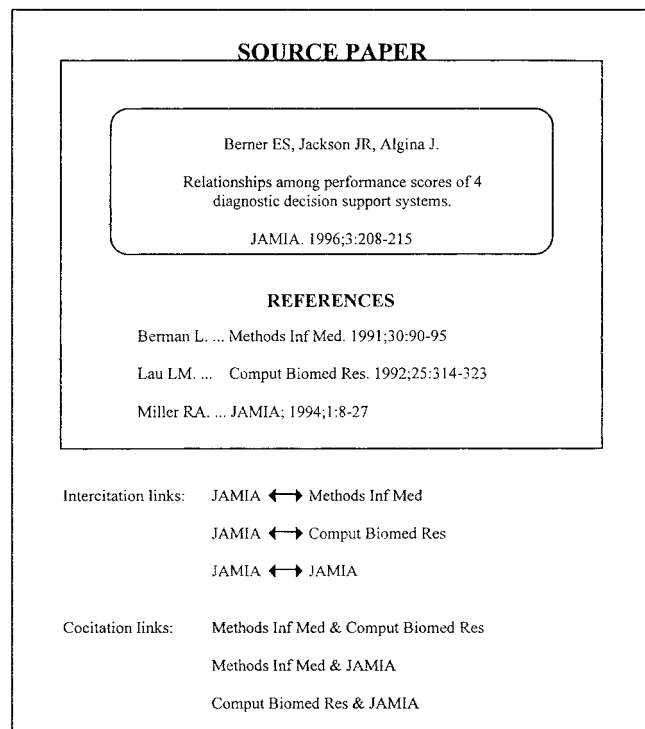
Our focus is on the identification and analysis of a “core journal literature” in medical informatics. Core journals are those likely to be central to the interests of researchers in the field because a high proportion of the articles they contain are relevant—these are the journals that are frequently chosen as publication outlets and whose articles are frequently cited. Additional articles of interest are scattered among many more journals, with proportionately fewer appearing per journal as the number of journals considered is increased. This characteristic “core and scatter” mirrors the cumulative advantage distribution noted by Price<sup>20</sup> and described as Bradford’s Law.<sup>21,22\*</sup>

The primary sources of data for core journal identification through intercitation and cocitation analyses are the online databases published by the Institute for Scientific Information (ISI)—SCISEARCH and SOCIAL SCISEARCH—and the related annual statistical compilations of journal productivity and citation visibility—the *Journal Citation Reports (JCR) for Science Citation Index (SCI)* and *Journal Citation Reports (JCR) for Social Science Citation Index (SSCI)*. In 1993, SCISEARCH covered just over 4,500 journals in the natural sciences, engineering, and medicine, while SOCIAL SCISEARCH indexed approximately 1,450 journals in the social sciences. There is some overlap of coverage in areas such as information science, neuroscience, geriatrics and gerontology, public health, and psychology. The *JCRs* contain information concerning journal productivity, prominence, and subject relationships. The data include the number of source items published annually (for the most recent three years), the number of citations received annually (for the most recent ten years plus a residuum), lists of journals by ISI subject category, and frequency-ranked lists of journals linked by intercitation.

\*Bradford<sup>21</sup> states that “If scientific journals are arranged in order of decreasing productivity of articles on a given subject, they may be divided into a nucleus of periodicals more particularly devoted to the subject and several groups of zones containing the same number of articles as the nucleus, when the number of periodicals in the nucleus and succeeding zones will be as 1:n:n<sup>2</sup>.” Although the precise mathematics of the relationship has been the subject of ongoing discussion since Bradford first proposed it in 1934, the general operation remains strikingly applicable across disciplines.

The best known measure of journal prominence in the *JCRs* is the impact factor—the ratio between the number of articles published in two years in a given journal and the number of current citations to those works from all indexed sources.<sup>23</sup> Other *JCR*-based approaches to assessing journal prominence and identification of a core journal literature include counting the citations from journals within a specific discipline rather than from all science,<sup>24</sup> calculating the proportion of citations given and received,<sup>25</sup> and taking into account the number of articles within each citing journal as well as the cited journal.<sup>26,27</sup> This last approach—intercitation network analysis—moderates the effect of large or general science journals such as *Science* or *Nature* (or, in our case, *Journal of the American Medical Association [JAMA]*), which, by virtue of the size of their corpus, will carry some appreciable number of relevant articles. McCain has used intercitation network analysis to identify core journal networks in genetics,<sup>27</sup> marine sciences,<sup>28</sup> biotechnology,<sup>29</sup> and fisheries and aquatic sciences.<sup>30</sup>

The two different journal linkages, intercitation and cocitation, are shown in Figure 1. The source article from *JAMIA* includes in its bibliography citations to an article from *Methods of Information in Medicine*, one from *Computers and Biomedical Research* and one from



**Figure 1** Relationships among a source paper, its cited journals, and their intercitation and cocitation links.

its own back run. *JAMIA* (as a source journal) is linked to each of the three cited journals via intercitation. Each pair of cited journals (i.e., *Methods of Information in Medicine* and *JAMIA*) is linked via cocitation. In this study, we used intercitation network analysis in our examination of intercitation data within the journal literature of medical informatics over ten years (1984–93) to identify potential core medical informatics journals. To explore major focus areas and research specialties in medical informatics, we examined journal cocitation patterns among articles published in 20 core medical informatics journals over approximately two and a half years (January 1993–July 1995).

## Methods

### Identifying a “Core” Journal Set

The list of journals identified as the “core” medical informatics titles was developed on the basis of an evaluation of candidate core journals primarily in terms of their propensity to publish articles relevant to medical informatics and the patterns of citations they give and receive. In this way we hoped to define the core as a set of journals that were relatively highly productive, in terms of publishing medical informatics articles, and relatively strongly linked through the referencing patterns of the article authors.

Initially, we searched seven Dialog databases covering biomedicine or information and computer science, or both. These databases were EMBASE (Excerpta Medica), Information Science Abstracts (ISA), INSPEC, Library and Information Science Abstracts (LISA), MEDLINE, SCISEARCH, and SOCIAL SCISEARCH. For each, we tried to create a database-appropriate search strategy that combined a suite of medical and health-related terms with a suite of information science and computing terms. We reasoned that a “core” medical informatics journal would be likely to be covered by both biomedical and information science indexing services, have relevant language in the article title, abstract, or indexing term fields, and have sufficient citation visibility to be covered by one or both citation index databases. Comparisons were made easier through the use of the Dialog RANK command to list the journal titles in terms of number of articles in the retrieval set.<sup>31</sup>

We identified 27 journals that were highly ranked (in terms of number of retrieved articles related to medical informatics) in four or more of the seven databases. Most of these titles could be placed easily into one of two groups—general medical journals (e.g., *JAMA*, *American Journal of Public Health*, *Annals of Internal Medicine*) or clearly relevant medical informatics

journals (e.g., *International Journal of Bio-Medical Computing*, *Methods of Information in Medicine*). Two titles, *Information Processing & Management* (an information science title) and *Academic Medicine* (formerly *Journal of Medical Education*) did not fit well into either category. Lowering the threshold for inclusion from the databases to three, and choosing journals with high rankings within those databases, added six additional medical informatics, information science, and non-medical informatics titles.

We were able to eliminate the non-medical informatics general medical journals from the list using intercitation network analysis on the journal-to-journal citation data published in the 1993 *JCR for SCI*.<sup>27</sup> As noted earlier, this technique distinguishes between journals linked through substantial proportions of citations made and received (which are thus part of a core literature) and those that may receive many citations from one or more core journals but rarely make any in return. We retained all apparent medical informatics journals, information science journals, and *Academic Medicine* for the cocitation mapping. To this list we added *JAMIA*, since its recent publication date prevented it from ranking highly in the searches or being covered in the 1993 *JCR*. At this stage we also added the *American Health Information Management Association Journal* and two nursing journals (*Journal of Advanced Nursing* and *Journal of Nursing Administration*), on the basis of their ranking in the database searches for terms relevant to medical informatics. A total of 29 possible medical informatics and related titles were submitted to cocitation analysis (Table 1).

### Cocitation Analysis

Journal cocitation data are compiled as counts of current (source) papers that cite at least one article from two different journals. In the single source paper shown in Figure 1, for instance, *JAMIA* (as a cited journal title) has a cocitation count of 1 with *Computers and Biomedical Research* and a cocitation count of 1 with *Methods of Information in Medicine*.† These cocitation data can be collected by searching the online version of the citation indexes (in this case both SCISEARCH and SOCIAL SCISEARCH) and specifying the journal pairing of interest. See McCain for details on the search methodology.<sup>27</sup>

In each analysis, we retrieved cocitation counts for all unique pairs of potential core journals in that portion

---

†Had there been several different *JAMIA* articles cited in this source paper, the cocitation count for *JAMIA* with each of the other two journal titles would still have been 1.

Table 1 ■

## Titles in the 29-Journal Data Set

Journal Title	Journal Abbreviation
<i>Academic Medicine*</i>	Acad Med
<i>American Health Information Management Association Journal†</i>	AHIMAJ
<i>Bulletin of the Medical Library Association</i>	BMLA
<i>Computer Applications in the Biosciences</i>	CABIOS
<i>Computer Methods and Programs in Biomedicine‡</i>	Comput Methods Programs Biomed
<i>Computerized Medical Imaging and Graphics</i>	Comp Med Imag Graph
<i>Computers and Biomedical Research</i>	Comput Biomed Res
<i>Computers in Biology and Medicine</i>	Comput Biol Med
<i>Computers in Nursing</i>	Comput Nurs
<i>IEEE Transactions on Biomedical Engineering</i>	IEEE Trans Biomed Eng
<i>Information Processing and Management</i>	Inf Process Manage
<i>International Journal of Bio-Medical Computing§</i>	Int J Bio-Med Comput
<i>International Journal of Clinical Monitoring and Computing</i>	Int J Clin Monit Comput
<i>International Journal of Man-Machine Studies</i>	Int J Man-Mach Stud
<i>Journal of Advanced Nursing</i>	J Adv Nurs
<i>Journal of Clinical Computing</i>	J Clin Comput
<i>Journal of Clinical Monitoring</i>	J Clin Monit
<i>Journal of Computer-Assisted Tomography  </i>	J Comput Assist Tomogr
<i>Journal of Medical Systems</i>	J Med Syst
<i>Journal of Nursing Administration</i>	J Nurs Adm
<i>Journal of the American Medical Informatics Association</i>	JAMIA
<i>Journal of the American Society for Information Science</i>	JASIS
<i>MD Computing¶</i>	MD Computing
<i>Medical and Biological Engineering and Computing#</i>	Med Biol Eng Comput
<i>Medical Decision Making</i>	Med Decis Mak
<i>Medical Education**</i>	Med Educ
<i>Medical Informatics</i>	Med Inform
<i>Methods of Information in Medicine</i>	Methods Inf Med
<i>Statistics in Medicine</i>	Stat Med

\*Including *Journal of Medical Education*.

†Including *American Medical Record Association Journal/Journal of the American Medical Record Association* and *Medical Record News*.

‡Including *Computer Programs in Biomedicine* and *Journal of Biomedical Measurement, Informatics and Control*.

§Title changed in 1997 to *International Journal of Medical Informatics*.

||Including *CT Journal of Computerized Tomography* and *Journal of Computerized Tomography*.

¶Including *Medical Computer Journal*, *Medcomp*, and *Computers in Medicine*.

#Including *Medical and Biological Engineering and Biological Engineering*, but not *Medical Electronics*, as there is also a current journal with that title.

\*\*Including *British Journal of Medical Education*.

of the SCISEARCH and SOCIAL SCISEARCH databases covering the indexing period January 1993–July 1995.‡ We were careful to incorporate both variant journal abbreviations in the “cited work” field and journal title changes in our search strategies.<sup>27,28</sup> The cocitation counts were recorded as square matrices (with journal titles marking the rows and columns) for further analysis.

‡All searches were limited to the accession number ranges 12004611–14023577 for SCISEARCH and 02440512–02773348 for SOCIAL SCISEARCH by using Dialog’s LIMITALL command.

Cluster analysis, factor analysis (including principal components analysis), and multidimensional scaling are among a group of techniques that can be used to explore the underlying structure in a data set. All these techniques rely on a transformation of the original observations into “proximity” data, indicating the similarity or dissimilarity of the pair of individuals or objects being considered.§

Our analyses are based on the similarity between cocitation *profiles*—the patterns of high and low cocita-

§See McCain<sup>28</sup> for additional discussion of this problem.

Table 2 ■

## Titles in the 20-Journal Data Set

Journal Titles	JCR Categories
<i>Academic Medicine</i> .....	Medical Informatics Medical, General & Internal Education, Scientific Disciplines
<i>American Health Information Management Association Journal</i> .....	Not indexed
<b>Bulletin of the Medical Library Association</b> .....	Information & Library Science (SSCI)
<b>Computer Applications in the Biosciences</b> .....	Computer Science, Interdisciplinary Applications Biology, Miscellaneous
<b>Computer Methods and Programs in Biomedicine</b> .....	Medical Informatics Computer Science, Interdisciplinary Applications Computer Science, Theory & Methods Engineering, Biomedical
<b>Computers and Biomedical Research</b> .....	Computer Science, Interdisciplinary Applications Engineering, Biomedical
<b>Computers in Biology and Medicine</b> .....	Computer Science, Interdisciplinary Applications Engineering, Biomedical
<b>Computers in Nursing</b> .....	Not fully indexed (SSCI)
<b>International Journal of Bio-Medical Computing</b> .....	Medical Informatics Computer Science, Interdisciplinary Applications Computer Science, Theory & Methods Engineering, Biomedical
<b>International Journal of Clinical Monitoring and Computing</b> .....	Medical Informatics Computer Science, Interdisciplinary Applications Medical Laboratory Technology
<i>Journal of Clinical Monitoring</i> .....	Medical Laboratory Technology
<b>Journal of Medical Systems</b> .....	Not fully indexed
<i>Journal of the American Medical Informatics Association</i> .....	Medical Informatics Computer Science, Information Systems Information Science & Library Science
<b>MD Computing</b> .....	Medical Informatics Computer Science, Interdisciplinary Applications
<b>Medical and Biological Engineering and Computing</b> .....	Medical Informatics Computer Science, Interdisciplinary Applications Engineering, Biomedical
<b>Medical Decision Making</b> .....	Medical Informatics
<i>Medical Education</i> .....	Medical Informatics Education, Scientific Disciplines
<b>Medical Informatics</b> .....	Medical Informatics Computer Science, Information Systems Medicine, Miscellaneous
<b>Methods of Information in Medicine</b> .....	Medical Informatics Computer Science, Information Systems Medicine, Miscellaneous
<i>Statistics in Medicine</i> .....	Medical Informatics Medicine, Research & Experimental Public Health Statistics & Probability

NOTE: The JCR categories are taken from the 1995 *Journal Citation Reports (JCR) for Science Citation Index* except where otherwise noted. Journal titles shown in **bold** also appear on the Sittig list (see reference 46).

tion counts—of the journals in the two core lists, rather than their raw cocitation counts. For each pair of journals, the cocitation frequencies were converted to Pearson correlations (using the SPSS program CORRELATIONS) and a new matrix generated. Using the correlation between two profiles as a similarity measure decreases the effect of scale (size of journal, num-

ber of citations received) and emphasizes patterns and subject relationships. The structure of the correlation matrix (profile similarities) was then investigated with cluster analysis (SPSS CLUSTER, Complete Linkage option), principal components analysis (SPSS FACTOR) and multidimensional scaling (SPSS ALSICAL). Individual differences scaling (INDSCAL options in ALSICAL) of

the two correlation matrices taken together was used to explore any systematic differences between the structures of the SCI and SSCI matrices.

Preliminary analysis of the cocitation data showed that three information science or systems journals (*Journal of the American Society for Information Science*, *Information Processing and Management*, and *International Journal of Man-Machine Studies*) and two medical imaging journals (*Journal of Computer-assisted Tomography* and *Computerized Medical Imaging and Graphics*) were extreme outliers, having little or no connection to the remainder of the journal set and severely distorting the maps. These were eliminated for the final analysis reported below. Also eliminated were the *IEEE Transactions on Biomedical Engineering*, which was not cocited with any of the remaining journals in SOCIAL SCISEARCH, and *Journal of Clinical Computing*, which was cocited with one-third or fewer of the other journals in SCISEARCH and SOCIAL SCISEARCH.<sup>32</sup> Finally, the two nursing journals added at the beginning of the study (*Journal of Advanced Nursing* and *Journal of Nursing Administration*) were removed from the data set, since they tended to form a single coherent, highly isolated cluster. The results reported in the next section are based on the final 20-title journal set listed in Table 2. This table also includes the ISI subject categories for the journals. All but four are covered by SCISEARCH; *Bulletin of the Medical Library Association* is indexed solely in SOCIAL SCISEARCH (Information Science and Library Science) and the other three (*AHIMA Journal*, *Computers in Nursing*, and *Journal of Medical Systems*) are partially indexed or not indexed at all by ISI. The interdisciplinary nature of medical informatics is well illustrated by the ISI subject category assignments in SCISEARCH. *Medical Informatics* is the only title solely assigned to medical informatics as a subject category, and *JAMIA* is the only medical informatics journal with an additional SCISEARCH categorization in information science. Four journals identified as core titles relevant to medical informatics in the intercitation network analysis are not categorized as medical informatics journals by ISI.

### Problems and Limitations

As illustrated above, some judgment was required regarding adjustments to the core journal set selections derived from the results of our original search strategy. Adding or subtracting journals from a data set may change the overall representation of the cocitation patterns for the journals in the data set. The intercitation relationships that we calculated could also have diminished or overlooked the relationships among important medical informatics journals because of the selective reporting of data for source and

nonsource, citing and cited journal data in the *JCRs*. Some of the journals studied began publishing only within the last ten years and could not garner “their share” of citations compared with older journals (e.g., *Computer Applications in the Biosciences* did not begin publication until 1985, and *JAMIA* did not begin until 1994, making them unavailable for the intercitation analysis and available for citation only during a small portion of the period included in the cocitation analysis). The length of a citable back run must be balanced, however, against the demonstrable citing of recent works more heavily than older works. However, Price’s Index suggests that “soft science” journals have less focus on citing recent works than “hard science” journals, so that a comparison of citations given within the same time period might not consider the “lag” in citation behavior in the soft sciences.<sup>33</sup> Finally, only a subset, or core set, of journals are being considered—not all scientific journals that might contribute or receive medical informatics citations, or all journals relevant to medical informatics overall, or even all journals that might be of interest to medical informatics researchers.

## Results and Discussion

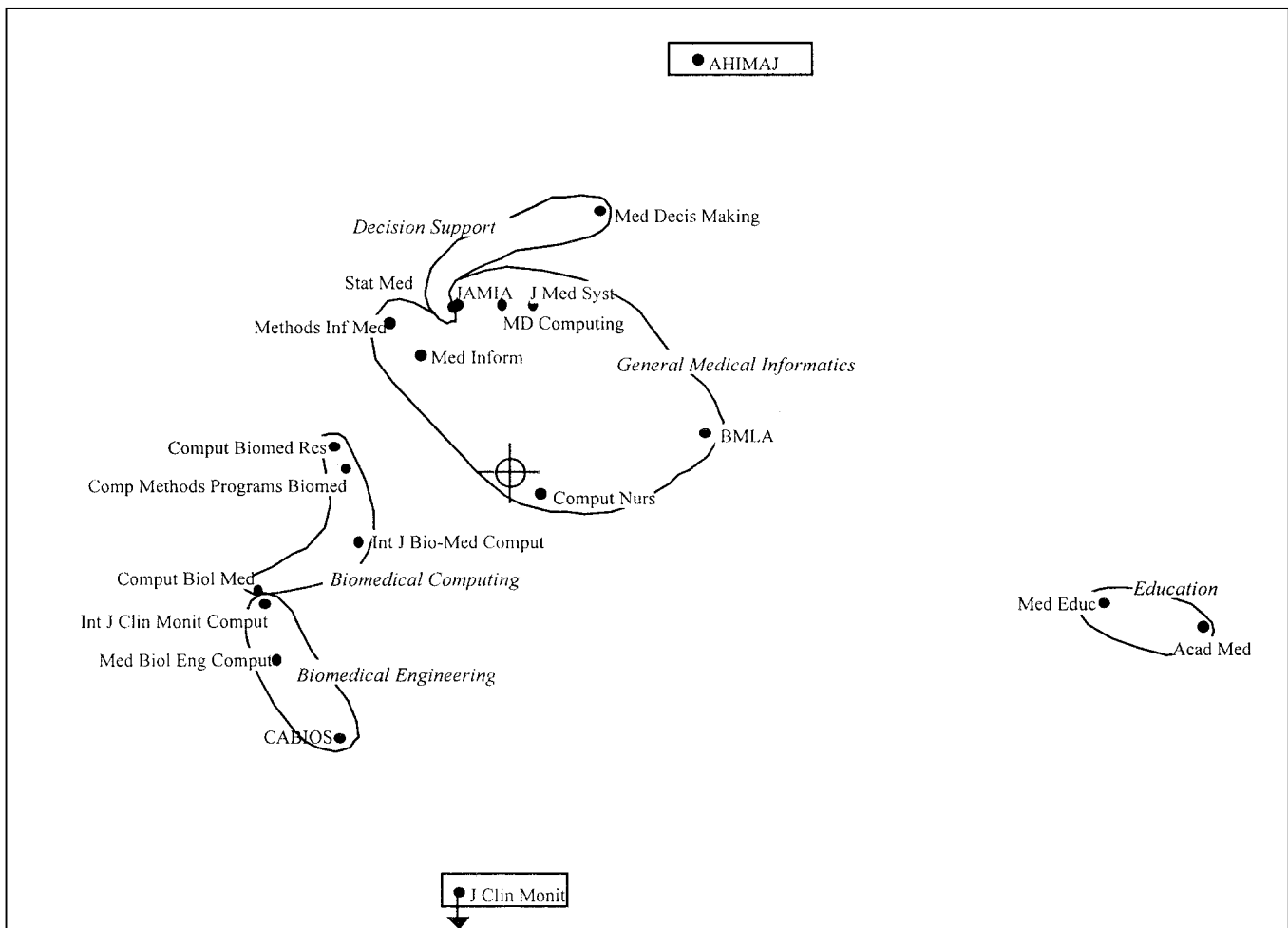
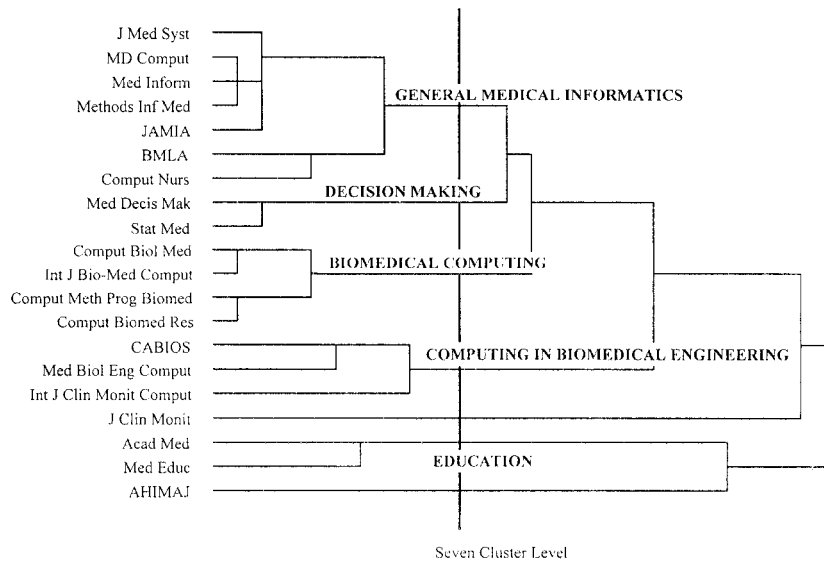
We ran separate analyses on the cocitation data gathered from SCISEARCH and SOCIAL SCISEARCH, and factor analyzed, mapped, and clustered the two data sets separately. Following this, we mapped the two data sets jointly to identify any difference in emphasis or perspective based on differences in the source journal coverage and subject orientation between SCI and SSCI. We discuss each of these analyses separately in this section.||

### SCISEARCH

Figure 2 (dendrogram) and Figure 3 (multidimensional scaling map) show the cluster structure for the SCI data. Based on several “stopping rules” for clustering, we determined that the seven-cluster model (five multijournal clusters and two singletons) had good explanatory potential.<sup>34,35</sup> In Figure 2, one can trace the formation of clusters from individual journals (far left) to a single cluster (far right). The vertical line near the right side of the figure points to the seven-cluster solution. Labels assigned to the clusters are impressionistic, reflecting the authors’ understanding of journal and article content. The five multijournal clusters suggest cohesive areas of research

||Copies of all data sets and SPSS analyses are available from Mr. Morris.

**Figure 2** Cluster solution for SCI citation data.



**Figure 3** SCI multidimensional scaling display with cluster boundaries.



and practice within medical informatics, broadly construed. *JAMIA* forms a fairly tight cluster with four other focused medical informatics titles as the main part of the “General Medical Informatics” cluster. Also part of the cluster at this stage are two journals with related content relevant to medical informatics: *Bulletin of the Medical Library Association* and *Computers in Nursing*. Two other substantial clusters—“Biomedical Computing” and “Computing in Biomedical Engineering”—include journals with a broader view of the intersection of computing/information systems/information technology and biomedicine. The partitioning apparently reflects the relative technologic content of highly cited articles—journals in both clusters are classified as “computer science” or “biomedical engineering” journals, or both, by ISI. The two remaining clusters focus on “Decision Making” and “Education,” respectively. At this clustering level, *Journal of Clinical Monitoring* and *American Hospital Information Management Association Journal* are isolates.

Multidimensional scaling techniques allow one to discover relationships among variables by representing their similarities or proximities in a spatial map.<sup>36</sup> Typically, an array of data points in  $n$ -dimensional space (where  $n$  equals the number of variables) is remapped into two or three dimensions for easier interpretation. As the number of dimensions is reduced, the relative positioning of two data points differs somewhat from their “true” proximities; when taken together these differences are considered to add “stress” to the model—a measure of goodness-of-fit, where lower values are better.<sup>37</sup> A balance must be struck among the desire for a small number of dimensions (for interpretability), explanation of a large amount of variance within the data set, and low stress in the model.<sup>38</sup>¶ Journals whose cocitation patterns are similar (highly correlated) will generally be positioned closer together in the diagram; those with less similar citation patterns are mapped farther apart. Journals placed near the center of the diagram (indicated by the compass rose) share similarities with many other journals. Journals appearing further from the center are less strongly connected overall or are well-linked to only part of the journal set. These may have a specialized focus or may serve as “environmental” or boundary-spanning journals, linking medical informatics to related disciplines.<sup>27,29</sup> The boundaries drawn around the journal names and the

assigned labels are taken from the seven-cluster model in Figure 2.

The placement of the journals and journal clusters in Figure 3 highlights important dimensions of scholarly activity in medical informatics. The most striking feature is the relative isolation of the “Education” cluster from the four remaining multijournal clusters and the two singleton journals, which are also placed peripherally on the map. This suggests a clear distinction in the minds of authors writing in journals indexed by SCI between “research and practice” and “education” as they relate to medical informatics. The main body of journals appears to be arrayed along a continuum that could be interpreted as “more engineering and tool oriented” versus “more clinical practice oriented.”

Principal components analysis, a type of factor analysis, offers a different view of structure and integration within the correlation matrix by grouping the journals into combinations that would account for the largest proportion of variance in the data.<sup>39</sup> An advantage of factor analysis over cluster analysis is the ability to detect multiple relationships among journals, as opposed to the single cluster assignment shown in Figure 2. Using oblique rotation (SPSS OBLIMIN option) to allow multiple journal interrelationships to emerge, four factors, accounting for 81.0 percent of the variance, were extracted based on the “eigenvalue  $\geq 1$ ” criterion. Table 3 lists the journals loading with absolute value greater than 0.5 on at least one factor. Negative factor loadings in this table (and Table 4) reflect the degree of obliqueness of the factor; as the factor’s loading values are more negative, the factor is less oblique.<sup>40</sup> The factor, or component, labels were assigned on the basis of the content of those journals with high factor loadings (generally above 0.70). For example, *Medical Education* and *Academic Medicine* show high loadings on Component 3 (along with *Bulletin of the Medical Library Association*), and it seems reasonable to label this component “Education.” Component 1 includes all the “General Medical Informatics” journals from the cluster analysis, and we append that same label to this component. Many of these journals also have substantial loadings on Component 4, “Decision-making Support,” ranking just below the two journals clustered under the similar title, “Decision Making,” in Figure 2 and highlighting the importance of this topic across medical informatics. Component 3 integrates journals from the biomedical computing and engineering clusters; all these titles are classified by ISI in both Computer Science and Biomedical Engineering (Table 2).

In addition to the integration of topics discussed

¶For the two-dimensional SCI representation in Figure 3, the  $R^2$  (proportion of explained variance in the model) is 0.913 and the stress—given as Kruskal’s Stress 1<sup>37</sup>—is 0.152. The SSCI representation in Figure 5 illustrates a model with  $R^2 = 0.896$  and stress = 0.175.

above, two kinds of "boundary spanning" can be seen in the factor loadings. Five journals (*Bulletin of the Medical Library Association*, *Computer Methods and Programs in Biomedicine*, *Computers and Biomedical Research*, *Computers in Nursing*, and *Medical Informatics*)

Table 3 ■

Factor Loadings of Principal Components from the Science Citation Index (SCI)

Journal Title	Loading
<b>Component 1: General Medical Informatics</b>	
<i>Journal of Medical Systems</i>	0.92
<i>MD Computing</i>	0.90
<i>Journal of the American Medical Informatics Association</i>	0.81
<i>Computers in Nursing</i>	0.74
<i>American Health Information Management Association Journal</i>	0.70
<i>Bulletin of the Medical Library Association</i>	0.69
<i>Computer Methods and Programs in Biomedicine</i>	0.51
<i>Methods of Information in Medicine</i>	0.73
<i>Medical Informatics</i>	0.78
<i>Computers and Biomedical Research</i>	0.51
<b>Component 2: Biomedical Engineering and Computing</b>	
<i>Computers in Biology and Medicine</i>	-0.91
<i>Medical and Biological Engineering and Computing</i>	-0.88
<i>International Journal of Bio-Medical Computing</i>	-0.87
<i>Computer Applications in the Biosciences</i>	-0.81
<i>Computer Methods and Programs in Biomedicine</i>	-0.80
<i>International Journal of Clinical Monitoring and Computing</i>	-0.72
<i>Computers and Biomedical Research</i>	-0.61
<i>Computers in Nursing</i>	-0.57
<i>Journal of Clinical Monitoring</i>	-0.55
<i>Medical Informatics</i>	-0.55
<b>Component 3: Education</b>	
<i>Medical Education</i>	0.98
<i>Academic Medicine</i>	0.77
<i>Bulletin of the Medical Library Association</i>	0.53
<b>Component 4: Decision-making Support</b>	
<i>Statistics in Medicine</i>	-0.89
<i>Medical Decision Making</i>	-0.86
<i>Methods of Information in Medicine</i>	-0.79
<i>Medical Informatics</i>	-0.78
<i>Computer Methods and Programs in Biomedicine</i>	-0.74
<i>Journal of the American Medical Informatics Association</i>	-0.73
<i>Computers and Biomedical Research</i>	-0.73
<i>MD Computing</i>	-0.65
<i>Journal of Medical Systems</i>	-0.64
<i>International Journal of Bio-Medical Computing</i>	-0.62
<i>Bulletin of the Medical Library Association</i>	-0.57
<i>Computers in Biology and Medicine</i>	-0.54
<i>Computers in Nursing</i>	-0.53

NOTE: Only component loadings above 0.50 (absolute value) are listed.

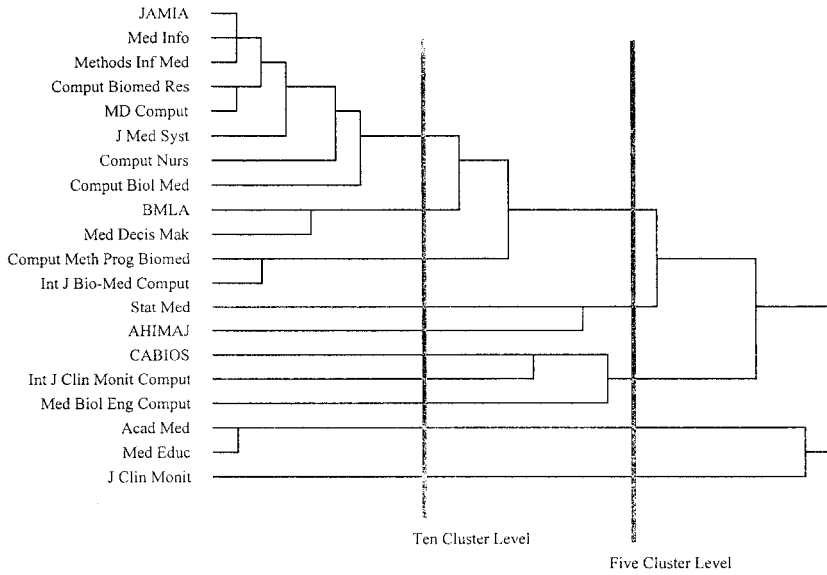
have substantial loadings on three of the four factors, suggesting that their cited content has greater breadth and is relevant to several different research topics within medical informatics.<sup>41</sup> (These journals are also positioned near the compass rose in Figure 3.) The cluster and map isolates *American Health Information Management Association Journal* and *Journal of Clinical Monitoring* have relatively substantial loadings on components 1 and 2, respectively. This suggests that, while they are not integrated into the core medical informatics literature, they are likely to be serving as links between medical informatics and related fields of research and practice.

#### SOCIAL SCISEARCH

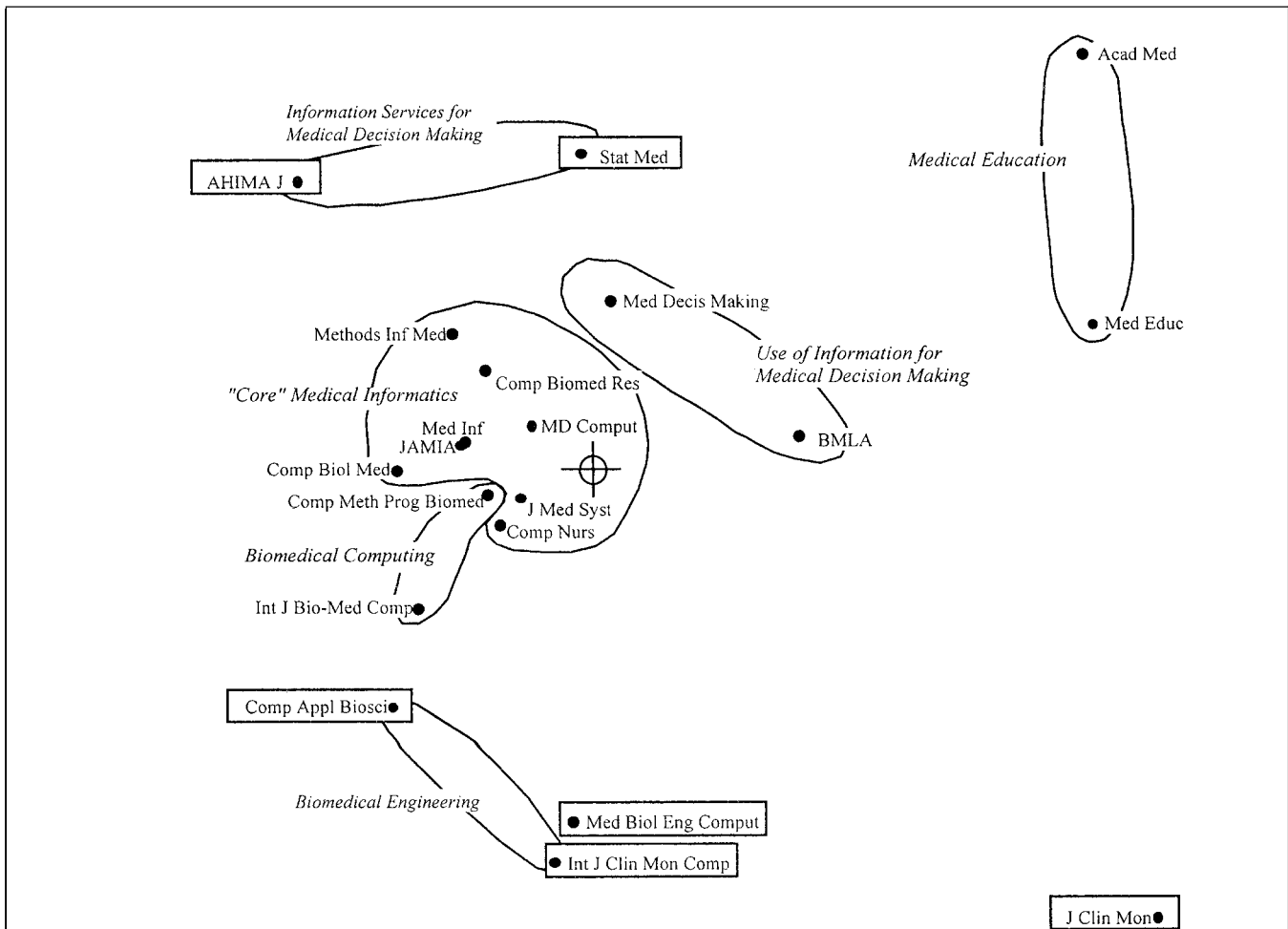
SOCIAL SCISEARCH indexes articles in social science journals. As noted earlier, there is only a partial overlap of journals and subjects covered by SSCI and SCI, and no overlap in indexing coverage of the 20 core titles in medical informatics. Thus, one might reasonably anticipate that the cocitation patterns will be somewhat different from those seen in the SCISEARCH data.

The cluster analysis for SSCI, as illustrated in Figure 4 and mapped in Figure 5, shows a looser structure than that for the SCI data (see footnote on p. 456). Two cluster levels are of interest here. At the five-cluster level, the major distinction seems to be between "medical and biomedical informatics" and "other." Here we find one large cluster containing all the general medical informatics and biomedical computing titles discussed earlier, three smaller clusters, and one singleton title (*Journal of Clinical Monitoring*). The biomedical engineering journals are still grouped, though somewhat more loosely than in Figure 2. The two education journals are tightly linked and remain separate until the last stages of clustering. The grouping of *Statistics in Medicine* and *American Health Information Management Association Journal* may simply be a result of their overall relative similarity to the medical informatics and biomedical titles as opposed to those in biomedical engineering or education. There is no strong connection between the two journals in the citing literature of this time period.

Earlier in the clustering process, at the ten-cluster level, most of the journals outside the "medical/biomedical informatics" grouping are isolates. Within this large cluster we can see a core grouping of medical informatics titles (along with two journals from the SCI biomedical computing category), the remaining pair of biomedical computing journals, and a separate pairing of *Bulletin of the Medical Library Association* and *Medical Decision Making*. These last two



**Figure 4** Cluster solution for SSCI co-citation data.



**Figure 5** SSCI multidimensional scaling display with cluster boundaries

journals were cocited 13 times by all journals covered in SSCI, with most of these cocitations given by *JAMIA* (six) and the *Bulletin* itself (three), in articles covering such topics as information access, concept representation, and retrieval. In SCI, the 22 cocitations for *Medical Decision Making* and *Statistics in Medicine* during the same period focus on outcomes research and meta-analysis of diagnostic performance measures. The *Bulletin of the Medical Library Association/Medical Decision Making* link appears to be "information services for medical decision makers," while the *Statistics in Medicine/Medical Decision Making* pairing suggests "use of information for medical decision making."

In the map showing multidimensional scaling results for SSCI (Figure 5), the placement of journals is fairly similar to that in Figure 3. Here, the underlying dimensions appear to be "patient care theory" versus "patient care practice" (upper left to lower right) and "clinical information gathering" versus "clinical decision making" (lower left to upper left).

A principal components analysis (oblimin rotation) yielded five factors, or components, accounting for 85.8 percent of the explained variance (Table 4). Ten journals had high loadings on Component 1, led by *American Health Information Management Association Journal*, *Methods of Information in Medicine*, *MD Computing*, *JAMIA*, *Medical Informatics*, and *Journal of Medical Systems*, all loading at 0.80 or higher. We have again labeled this factor "General Medical Informatics." It is interesting to notice the inclusion of *American Health Information Management Association Journal* in this group, with the highest loading (0.89). It has no other substantive loadings on other factors; essentially, all its variance is explained by its linkage with the medical informatics journals in Component 1. The second component, "Education," is defined by *Medical Education*, *Bulletin of the Medical Library Association*, and *Academic Medicine*, all loading at 0.78 or higher. The bipolar nature of Component 3, "Patient Monitoring," appears to point to two contrasting aspects of data management: data acquisition and data analysis. The visibility of this topic in the SSCI data (as opposed to the SCI data) may be an indication of greater emphasis on human contact aspects versus instrumentation and technology aspects of health care in the writings of social science authors. Components 4 and 5 represent the research and application aspects of biomedical computing and engineering, respectively. The "General Medical Informatics" component alone accounts for 52.6 percent of the total variance. That so many journals load heavily on multiple components reflects the integration and interdisciplinarity of medical informatics in the eyes of authors writing in social science journals.

Table 4 ■

## Factor Loadings of Principal Components from the Social Science Citation Index (SSCI)

Journal Title	Loading
<b>Component 1: General Medical Informatics</b>	
<i>American Health Information Management Association Journal</i>	0.89
<i>Methods of Information in Medicine</i>	0.88
<i>MD Computing</i>	0.83
<i>Journal of the American Medical Informatics Association</i>	0.81
<i>Medical Informatics</i>	0.80
<i>Journal of Medical Systems</i>	0.80
<i>Computers and Biomedical Research</i>	0.80
<i>Computers in Biology and Medicine</i>	0.73
<i>Computers in Nursing</i>	0.65
<i>Medical Decision Making</i>	0.54
<b>Component 2: Education</b>	
<i>Medical Education</i>	1.03
<i>Bulletin of the Medical Library Association</i>	0.79
<i>Academic Medicine</i>	0.79
<i>Medical Decision Making</i>	0.64
<b>Component 3: Patient Monitoring</b>	
<i>Journal of Clinical Monitoring</i>	0.76
<i>International Journal of Clinical Monitoring and Computing</i>	0.71
<i>Statistics in Medicine</i>	-0.55
<b>Component 4: Biomedical Computing and Engineering Research</b>	
<i>Computers and Biomedical Research</i>	-0.82
<i>Medical Informatics</i>	-0.75
<i>Computer Methods and Programs in Biomedicine</i>	-0.74
<i>Computers in Biology and Medicine</i>	-0.73
<i>Journal of Medical Systems</i>	-0.72
<i>Journal of the American Medical Informatics Association</i>	-0.70
<i>International Journal of Bio-Medical Communications</i>	-0.69
<i>Computer Applications in the Biosciences</i>	-0.68
<i>Medical Decision Making</i>	-0.65
<i>MD Computing</i>	-0.64
<i>Methods of Information in Medicine</i>	-0.61
<i>Bulletin of the Medical Library Association</i>	-0.57
<i>Statistics in Medicine</i>	-0.56
<i>Computers in Nursing</i>	-0.50
<b>Component 5: Biomedical Computing and Engineering Application</b>	
<i>Computer Methods and Programs in Biomedicine</i>	0.92
<i>Computers in Nursing</i>	0.84
<i>Computer Methods and Programs in Biomedicine</i>	0.83
<i>Computers in Biology and Medicine</i>	0.73
<i>Medical Informatics</i>	0.69
<i>Journal of the American Medical Informatics Association</i>	0.68
<i>International Journal of Bio-Medical Communications</i>	0.67
<i>MD Computing</i>	0.64
<i>Journal of Medical Systems</i>	0.62

NOTE: Only component loadings above 0.50 (absolute value) are listed.

Table 5 ■

Relative Weights Assigned for Individual Differences Scaling

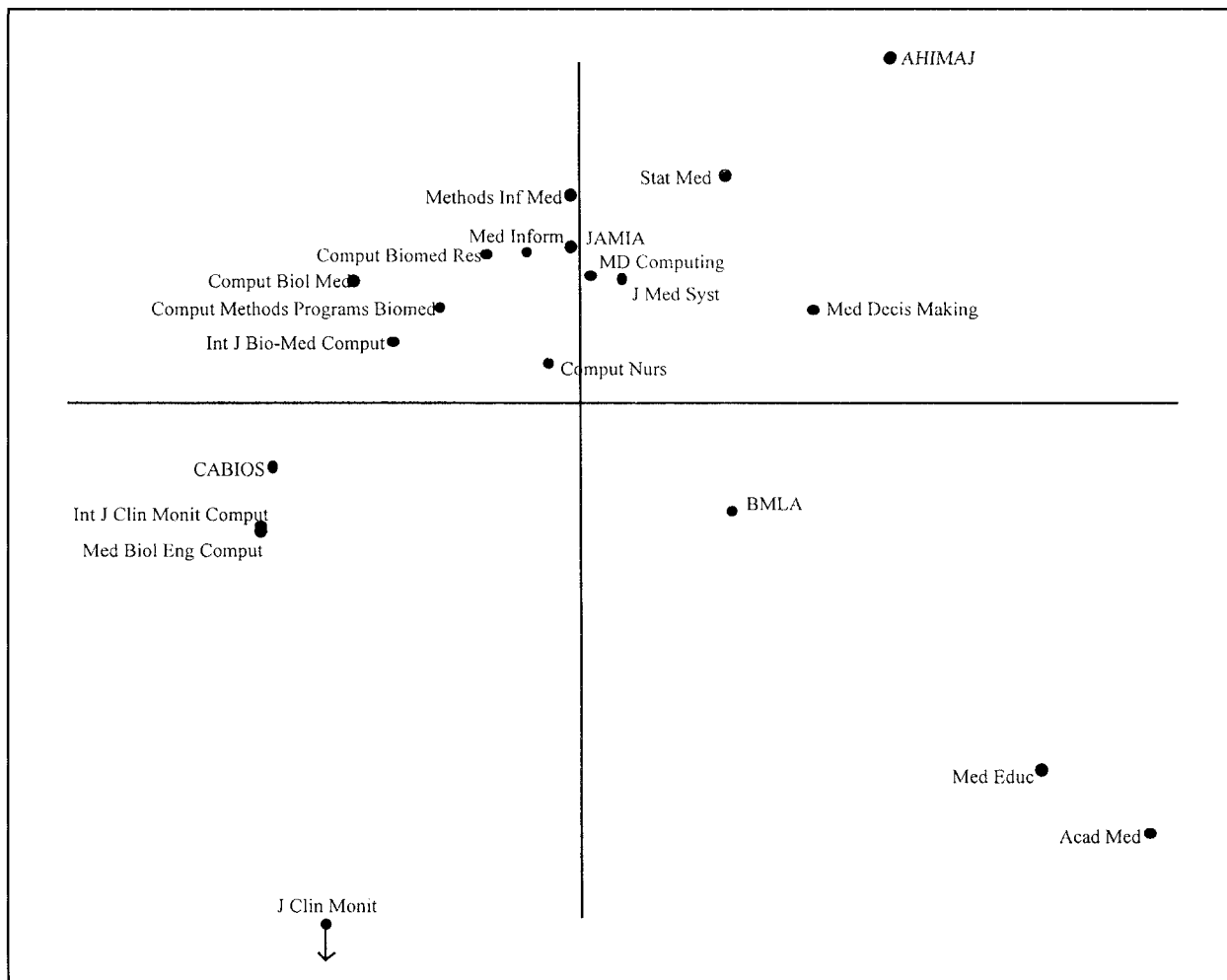
Data Set	Weight Assigned to:	
	Dimension 1 (X)	Dimension 2 (Y)
SCI	0.81	0.50
SSCI	0.56	0.73

**Comparison of SCISEARCH and SOCIAL SCISEARCH**

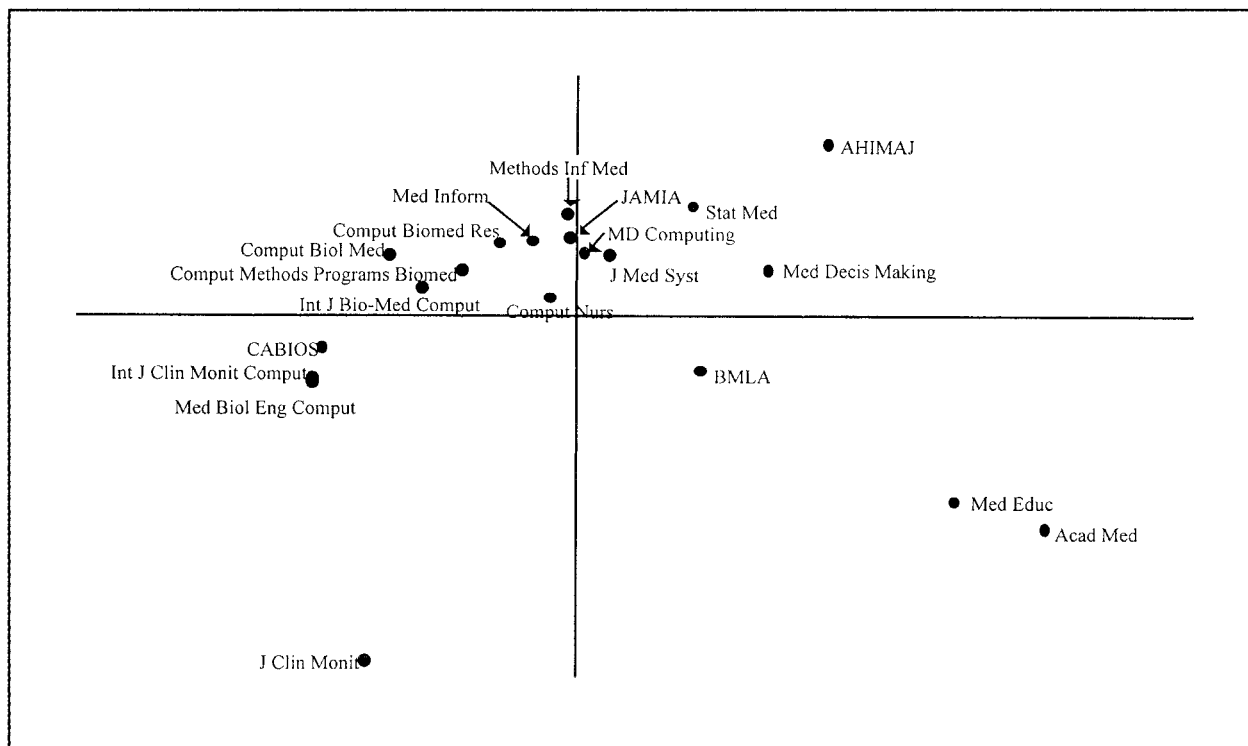
Individual differences scaling (INDSCAL) performed on the SCI and SSCI data sets jointly highlights the different perspectives on medical informatics provided by authors writing in science and social science journals. The output of an INDSCAL analysis includes a map (see Figure 6), representing the “joint subject

space” (a summary display of the best dimensional configuration when all input matrices are considered together), and a set of weights (Table 5) that represents the different emphases given to the dimensions of the group subject space in the separate cocitation correlation matrices. These weights can be used in combination with the Cartesian coordinates of the joint map to produce maps showing the “individual differences” of SCI versus SSCI journal cocitation patterns.<sup>42</sup> # The axes of INDSCAL maps may not be rotated further and must be interpreted as calculated by the program; for that reason we include the axes rather than a compass rose (Figures 6 to 8).

#The phrase “individual differences” comes from the original purpose of this approach—to collect proximities data from, say, 12 experimental subjects, map the 12 data sets jointly, and then examine each subject’s personal perspective map as a weighted version of the summary map.



**Figure 6** Combined (unweighted) individual differences scaling display.



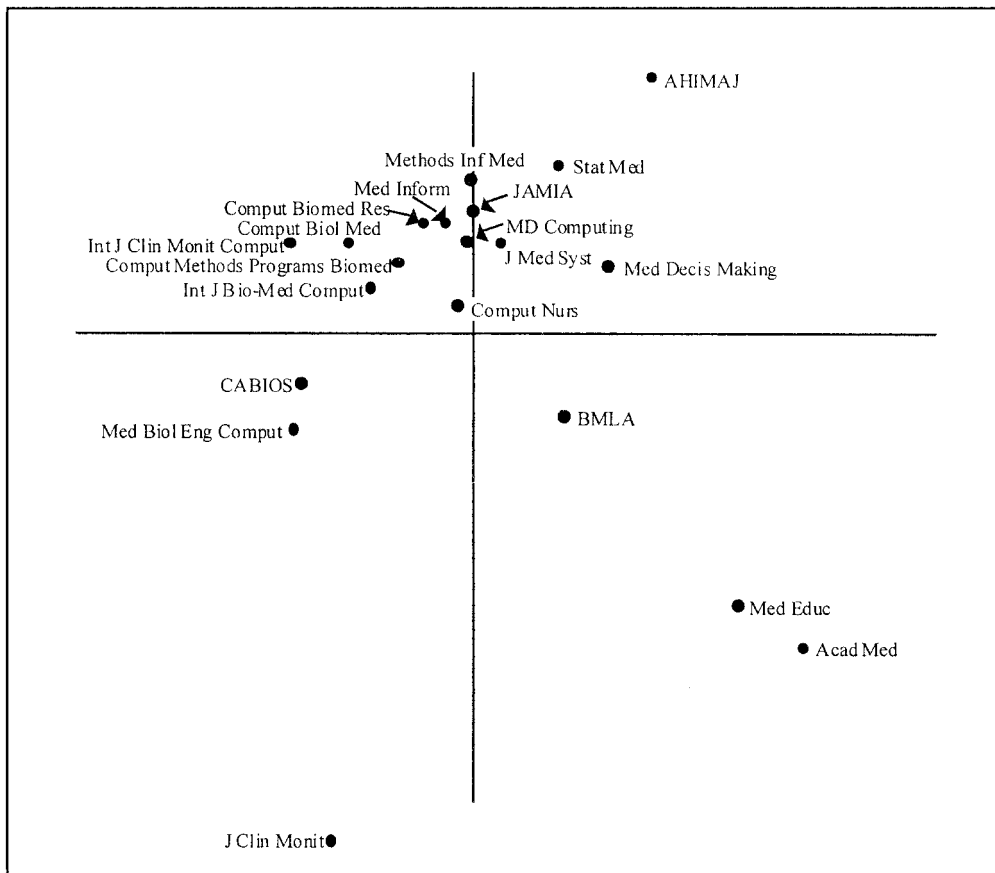
**Figure 7** SCI weighted individual differences scaling display.

In Figure 6 we can see the same linear array of medical informatics and related titles along a rough diagonal from lower left to upper right, as well as the separated "Education" grouping and the outlier *Journal of Clinical Monitoring*. The vertical axis separates journals with computational and engineering content (to the left) from those with a stronger focus on information management and education. The horizontal axis appears to distinguish between research or theoretic content (above) and an orientation toward application and clinical practice (below). This interpretation of the dimensions echoes that seen in some other multidimensional scaling studies of scientific literature cocitation data, where one axis represents a dimension of greater or lesser emphasis on mathematics or formal methods and another represents a domain-specific subject orientation.<sup>43</sup>

The differences contributing to the structure of the display, derived from the SCI and SSCI data, are expressed as weights that adjust the coordinate axes. These weights represent the different emphases given to the underlying dimensions of the data structure by the two sets of citing authors. The effect of these underlying dimensions can be made clearer when the display is redrawn taking the weights into account. These are shown as Figure 7 (SCI) and Figure 8 (SSCI).

Comparing the unweighted display (Figure 6) with the corresponding weighted displays (the three figures are drawn to the same scale), we see the overall shape of the SCI display (Figure 7) widened and flattened, suggesting that finer distinctions are being made among topics in journals appearing along the computer/engineering::information management/education continuum than along the theory/research::application/clinical practice continuum. The SSCI display (Figure 8) is narrowed and elongated vertically, with the opposite emphasis; here, relatively greater weight is given to distinguishing between theory and clinical practice. This suggests that authors writing in journals indexed by SSCI make a greater distinction between theory and application when building their arguments and reporting their research. Although not as visually striking, the effect is similar to that seen in Helm's study of subjects' comparisons of color chips.<sup>44\*\*</sup>

\*\*In Helm's study,<sup>44</sup> the observations of subjects with normal color sight mapped as a circle corresponding to the color wheel, with the orthogonal axes of the two-dimensional map anchored by red and green and by blue and yellow, whereas color-blind subjects' observations mapped as ellipses—they did not consider the red-green (or blue-yellow) information as strongly when making color-matching decisions.



**Figure 8** SSCI weighted individual differences scaling display.

It should be remembered that in the SSCI data set, cocitations are not being provided by most of the journals on the map—the majority of the journals in the core are indexed online in SCI. Thus, the picture should be expected to be different, since the core journals are not “voting,” as it were. More of the SSCI picture is being provided by authors writing outside the medical informatics core and citing into the medical informatics literature than is the case in the SCI data set.

### Relative Role of *JAMIA*

*JAMIA* began publication fairly recently, and its citation position—both within medical informatics and in the view of authors discussing topics relevant to medical informatics in other journals—is only now beginning to be established. In the period under study, *JAMIA* sits close to the midpoint of the horizontal axis in the INDSICAL maps, balancing between the engineering and information management/education poles of the display. The relatively central position of *JAMIA* within the “general medical informatics” clusters indicates that its cocitation pattern is similar to those of well-established titles such as *MD Computing*

and *Medical Informatics*. In addition, the journals most frequently cocited with *JAMIA* in the time period studied were (in order) *Methods of Information in Medicine*, *Computers and Biomedical Research*, *MD Computing*, and *Medical Informatics*. This was true for both the SCI and SSCI data sets. The substantial loading of the *American Health Information Management Association Journal* on the general medical informatics factor in both the SCI and SSCI factor analyses points to a stronger content linkage than is apparent in the cluster analyses and maps. As the American Health Information Management Association and American Medical Informatics Association pursue more inter-organizational activities, it will be interesting to see how the two societies’ journals are (co)cited in the future.

### Comparison with Prior Studies

#### Core Journal Sets in Medical Informatics

Our list of 20 core titles in medical informatics, based on citation linkages and indexing data from several bibliographic databases, is strongly congruent with previously published lists. Sittig and Kaalaas-Sittig<sup>45</sup> ranked biomedical informatics serials on the basis of

such productivity measures as impact factor, number of citations received, citations given by several informatics specialty texts, library holdings, circulation, and interlibrary loan requests. They also conducted a popularity survey among medical informatics fellows. Their set of journals included all those indexed by the National Library of Medicine with terms from the medical informatics hierarchy of the Medical Subjects Headings (MeSH) for Index Medicus, those indexed by the International Yearbook of Medical Informatics, and several other selected journals from standardized lists and considered opinion. The journals were ranked on each of several scales and on a combined scale. Of 34 journals ranked, 14 of the 17 are included in this study.†† (In addition, 2 of the top 20 sources they included were conference proceedings, which we did not consider for this study.) In a later paper, Sittig<sup>46</sup> examined journals indexed by the National Library of Medicine and ranked them according to the number of articles they carried that were indexed with terms from the MeSH medical informatics hierarchy and those that published within their total output the greatest percentage of medical informatics articles indexed. Our study included 11 of the top 14 journals in terms of publishing medical informatics articles (see Table 2).‡‡

Greenes and Siegel<sup>5</sup> considered a group of 64 journals and 12 volumes of proceedings for indexing by the National Library of Medicine as being important to medical informatics. Their journal list included 12 titles in this study, 10 of which were included in our final 20-journal set. This list also included 27 general medical or medical specialty journals that were excluded from this study by intercitation analysis. The remainder of the journals pertained to such topics as biometrics and public health along with several computing and information science titles. A group of American College of Medical Informatics fellows

---

††In order of ranking from Sittig and Kaalaas-Sittig,<sup>45</sup> these journals are *Computers and Biomedical Research*, *MD Computing*, *Methods of Information in Medicine*, *Medical Decision Making*, *Computers in Biology and Medicine*, *Journal of Chemical Information and Computer Sciences*, *Computer Methods and Programs in Biomedicine*, *Medical and Biological Engineering and Computing*, *International Journal of Clinical Monitoring and Computing*, *Bulletin of the Medical Library Association*, *Medical Informatics*, *International Journal of Bio-Medical Computing*, *Computers in Nursing*, *Computer Applications in the Biosciences*, and *Journal of Medical Systems*.

‡‡The journals, in order of ranking by Sittig,<sup>46</sup> are *Computer Applications in the Biosciences*, *Computer Methods and Programs in Biomedicine*, *International Journal of Bio-Medical Computing*, *Computers in Biology and Medicine*, *Computers in Nursing*, *Computers and Biomedical Research*, *Methods of Information in Medicine*, *MD Computing*, *Medical Informatics*, *Journal of Medical Systems*, and *International Journal of Clinical Monitoring and Computing*.

found four of the suggested criteria to be key for evaluating journals for possible inclusion or retention in MEDLINE: quality of articles, readership by medical informatics professionals, importance as a source of ideas arising outside medical informatics, and desirability of indexing in MEDLINE. For these four criteria, none of our "core" medical informatics titles ranked in the top ten for quality of articles; seven of our medical informatics journals ranked in the top ten for readership by medical informatics professionals; one ranked in the top ten for source of ideas arising outside medical informatics; and three ranked in the top ten for desirability of indexing in MEDLINE. On the basis of these and other results, Greenes and Siegel concluded that while "purely bibliographic measures" such as immediacy and journal impact factor might be useful in examining the broad field of medicine, they did not help determine a journal's importance to medical informatics. Ninety-four percent of the journals ranked above the median for desirability of indexing by the National Library of Medicine were already covered by it. That this included 27 titles not in our core selections suggests how wide-ranging are medical informatics professionals' interests. That so few information science journals were included in the top rankings suggests that the borrowing relationship of medical informatics with information science is rather one-sided, and not deemed particularly citeworthy.

Garfield<sup>47</sup> mapped the narrower research area of biomedical engineering through cocitation analysis using the biomedical engineering journals then indexed in the 1984 SCI. Of those 19 journals, four journals or their successors are found in our core journal set (*Computers in Biology and Medicine*, *Computers and Biomedical Research*, *Computer Programs in Biomedicine*, which became *Computer Programs and Methods in Biomedicine*, and *Medical and Biological Engineering and Computing*). A fifth (*IEEE Transactions on Biomedical Engineering*) was considered but not included in the final 20-journal set discussed here. The four journals group into the "computers in biomedical engineering" and "biomedical computing" clusters in our analysis of SCI data and in the engineering and "general medical informatics" clusters in SSCI. Comparing the number of citations the journals made to one another with the number of articles included in ISI's biomedical engineering "research fronts," Garfield suggested that it was clear that many papers in the field were published in other than the core journals.

Considering the congruence of the lists of journals considered by the first three of the above studies with our journal lists, and despite the observations of



Greenes and Siegel regarding bibliographic measures, we believe our findings bring a new perspective to the discussion. Garfield's findings are echoed by our impressions that any subset area of medical informatics research very quickly expands into the wider literature of adjacent subject fields.

#### The Subject Structure of Medical Informatics

Levy, Shortliffe, Lincoln and Korpman, and others commenting on the organization of medical informatics and its subject structure have focused on fairly fine-grained themes and research areas.<sup>3,4,13,48</sup> Levy<sup>13</sup> identified four distinct categories of components of medical informatics: those related to computing, to systems analysis, to health care organization, and to biology and physiology. Shortliffe,<sup>4</sup> listing major research areas in medical informatics needing attention at the time, included knowledge representation, knowledge and data acquisition, medical decision making, cognitive sciences, human-machine interfaces, information storage and retrieval, and evaluation methodologies. Lincoln and Korpman<sup>3</sup> asserted that there are three overlapping domains in health care—clinical medicine, health management and statistics, and fundamental sciences—and that medical informatics concentrates specifically on the overlaps. The International Medical Informatics Association Yearbook (subtitled "Advances in an Interdisciplinary Science") groups its chapters into sections titled "Health and Clinical Management," "Computer-based Patient Records," "Information Systems," "Image and Signal Processing," "Decision Support Systems," "Knowledge Processing," and "Education."<sup>48</sup> These echo the various themes brought out in our study—in journal clusters and subclusters, factor loadings, and map placements.

Our use of the scholarly journal (rather than individual documents or key authors) as the unit of analysis and representation results in a more global view of medical informatics and related fields than the specific focuses mentioned by earlier medical informatics commentators. In addition to a cluster of journals representing "general medical informatics," (the "core of the core," so to speak), we were able to identify related topic areas in biomedicine, biomedical engineering, decision making, and education that had their own core journal subset. Whether viewed from the perspective of authors writing in the sciences (including journals in medical informatics, computer science, information systems, and medicine) or the social sciences, the overall perceived structure of the field remained reasonably consistent.

#### Future Research

Medical informatics is a discipline that is still developing and changing. We have provided one snapshot of the discipline as it existed in the early 1990s and as portrayed through one unit of analysis—the highly (co)cited scholarly journal. The results of these investigations provide one viewpoint of the use and distribution of the literature of medical informatics. They must be independently validated with other techniques and other units of analysis. Journals could be mapped, clustered, and factor-analyzed on the basis of subject indexing profiles in order to provide a second, complementary view of subject structure. The relationship between other measures of journal prominence (article productivity, visibility in library collections, circulation counts, etc.) and citation-based measures should also be examined.

Analysis of the patterns of cocitation of authors' works (author cocitation analysis), coupled with title content terms and citations to key works, could provide the finer-grained results necessary to explore specialties in research and practice within medical informatics. Analyses over subsequent time periods can be expected to highlight changes in the intellectual structure of medical informatics in general and the role of *JAMIA* as a major medical informatics core journal in particular. Additional cross-validation of these citation and indexing results through correlation with comparable data collected from studies of other disciplines can provide additional insights.<sup>49</sup>

We began this study with the assumption that we would discover connections between the domains of information science and medical informatics. It was clear from preliminary analyses that, at least at the level of the journal (and based on aggregate citation data from SCI and SSCI), information science and medical informatics are linked weakly, if at all. The information science journals containing articles relevant to medical informatics distorted the maps so badly (because of their lack of cocitation linkages) that they had to be eliminated for the structure of medical informatics to emerge at all. The paths by which information science information enters medical informatics can be addressed by investigating the roles of boundary-spanning journals, such as the *American Health Information Management Association Journal* and the *Journal of Clinical Monitoring*, and journals in neighboring disciplines. We anticipate that this will yield a better picture of the structure of medical informatics within the context of the greater medical journal corpus and reveal its links to predecessor and

neighboring disciplines. Although medical informatics may borrow from the information sciences, it does not appear to cite them to the same degree. Perhaps this is an indication that computers and computing technology are so ubiquitous in modern medicine as to cease to be noteworthy. Also, the history of medical informatics shows it to have been instigated by researchers and practitioners who adopted or adapted computing technologies for their own biomedical applications. Early results were published primarily in the medical literature,<sup>50</sup> and reporting of medical informatics results have apparently remained closer to that venue than to the information science literature.

We originally included two nursing journals in the potential core list on the basis of their inclusion of terms relevant to medical informatics in database records. However, they were removed before the final analyses. No other divisions within the health care arena that identify their own slant on informatics, including pharmacy, dentistry, and veterinary medicine, had any highly ranked journals in our exploratory database searches leading toward the intercitation analyses. Neither did we include such titles heuristically when conducting the cocitation analyses. It may be that nursing has sought more actively to differentiate its own unique and particular information needs. Further validation of these findings would consider journals in pharmacy, dentistry, and veterinary medicine that can be independently judged to contain informatics-oriented papers.

## Summary

Medical informatics is an interdisciplinary field, claiming ties to biomedical research, clinical practice, medical education, and information and computer science. We found evidence of its interdisciplinarity through the factor analysis of our cocitation data, especially in the SSCI data set where many journals loaded heavily on multiple factors. Medical informatics extends across hard and soft science boundaries, and its literature is used differently by authors in those arenas—the former emphasizing engineering versus information management and education issues; the latter emphasizing issues of theory versus clinical practice. Emerging multidisciplinary and interdisciplinary fields have ties to neighboring disciplines. However, expected strong ties to information science journals were not found. Other relationships between medical informatics and related disciplines have yet to be delineated through citation analysis and will be addressed in further studies.

## References ■

- Shortliffe EH. The state of the art in medical information sciences. In: Kuhn RL (ed). *Frontiers of Medical Information Sciences*. New York: Praeger, 1988;11–8.
- Protti DJ. The synergism of health information science/health informatics. *CRC Crit Rev Med Inform*. 1986;1:149–63.
- Lincoln TL, Korpman RA. Computers, health care, and medical information science. *Science*. 1980;210(4467):257–63.
- Shortliffe EH. Medical informatics: the emergence of a discipline. In: Schoolman HS (ed). *Past, Present, and Future of Biomedical Information*. Bethesda, Md.: National Library of Medicine, 1987;73–87. NIH Publication 88–2911.
- Greenes RA, Siegel ER. Characterization of an emerging field: approaches to defining the literature and disciplinary boundaries of medical informatics. *Symp Comp Appl Med Care*. 1987;411–5.
- Lincoln TL. Medical informatics: the substantive discipline behind health care computer systems. *Int J Biomed Comput*. 1990;26:73–92.
- Blois MS. What is medical informatics? *Western J Med*. 1986;145:776–7.
- Klein JT. *Interdisciplinarity*. Detroit, Mich.: Wayne State University Press, 1990.
- Greenes RA, Shortliffe EH. Medical informatics: an emerging academic discipline and institutional priority. *JAMA*. 1990;263:1114–20.
- Brennan PF. On the relevance of discipline to informatics. *J Am Med Inform Assoc*. 1994;1:200–1.
- Friedman CP. Where's the science in medical informatics? *J Am Med Inform Assoc*. 1995;2:65–7.
- Warner HR. Medical informatics: a real discipline? *J Am Med Inform Assoc*. 1995;2:207–14.
- Levy AH. Is medical informatics a basic medical science? In: Shires DB, Wolf H (eds). *Medinfo*. 1977;979–81.
- American Medical Association, Council on Scientific Affairs and Council on Long Range Planning and Development. *Medical informatics: an emerging medical discipline*. *J Med Syst*. 1990;14:161–79.
- Kuhn TS. *The Structure of Scientific Revolutions*. Chicago, Ill.: University of Chicago Press, 1962.
- McCain KW, Whitney PJ. Contrasting assessments of interdisciplinarity in emerging specialties: the case of neural networks research. *Knowledge: Creation, Diffusion, Utilization*. 1994;15:285–306.
- Garfield E. Can citation indexing be automated? In: Stevens ME, Guilaiano VE, Heilprin L (eds). *Statistical Association Methods for Mechanized Documentation: Symposium Proceedings*. Washington, D.C.: Government Printing Office, 1965:189–92. National Bureau of Standards Miscellaneous Publication 269.
- McCain KW, Turner KW. Citation context analysis and aging patterns of journal articles in molecular genetics. *Scientometrics*. 1989;17:127–63.
- Small H. Cited documents as concept symbols. *Social Studies of Science*. 1978;8:327–40.
- Price DdS. A general theory of bibliometric and other cumulative advantage processes. *J Am Soc Inf Sci*. 1976;27:177–91.
- Bradford SC. *Sources of information on specific subjects*. Engineering. 1934;137:85–6.

22. White HD, McCain KW. Bibliometrics. *Annual Review of Information Science & Technology*. 1989;24:119–86.
23. Institute for Scientific Information. 1993 SCI Journal Citation Reports. Philadelphia: Institute for Scientific Information. 1993:10.
24. Hirst G. Discipline impact factors: a method for determining core journal lists. *J Am Soc Inf Sci*. 1978;29:171–2.
25. He C, Pao ML. A discipline-specific journal selection algorithm. *Information Processing and Management*. 1986;22:405–16.
26. Servi PN, Griffith BC. A method for partitioning the journal literature. *J Am Soc Inf Sci*. 1980;31:36–40.
27. McCain KW. Core journal networks and cocitation maps: new bibliometric tools for serials research and management. *Libr Q*. 1991;61:311–36.
28. McCain KW. Core journal networks and cocitation maps in the marine sciences: tools for information management in interdisciplinary research. In: Shaw D (ed). *Proceedings of the 55th American Society for Information Science Annual Meeting*. Pittsburgh, Pa. 1992:3–7.
29. McCain KW. Biotechnology in context: a database-filtering approach to identifying core and productive non-core journals supporting multidisciplinary R&D. *J Am Soc Inf Sci*. 1995;46:306–17.
30. McCain KW. Islands in the stream: mapping the fisheries and aquatic sciences literatures. *Fisheries*. 1994;19:20–7.
31. McCain KW. Bibliometric tools for serials collection management in academic libraries. *Adv Ser Manage*. 1997;6:105–46.
32. White HD, Griffith BC. Author cocitation: a literature measure of intellectual structure. *J Am Soc Inf Sci*. 1981;32:163–71.
33. Price DjdS. Citation measures of hard science, soft science, technology, and nonscience. In: Nelson C, Pollock D (eds). *Communication among Scientists and Engineers*. Lexington, Mass.: Heath, 1970;3–22.
34. Aldenderfer MS, Blashfield RK. *Cluster Analysis*. Beverly Hills, Calif.: Sage Publications, 1984.
35. Everitt BS. *Cluster Analysis*, 3rd ed. New York: Halstead Press, 1993.
36. Schiffman SS, Reynolds ML, Young FW. *Introduction to Multidimensional Scaling: Theory, Methods, and Applications*. Orlando, Fla.: Academic Press, 1981.
37. Kruskal JB, Wish M. *Multidisciplinary Scaling*. Beverly Hills, Calif.: Sage Publications, 1978.
38. McCain KW. Mapping authors in intellectual space. *J Am Soc Inf Sci*. 1990;41:433–43.
39. Gorsuch RL. *Factor Analysis*. 2nd ed. Hillsdale, N.J.: Erlbaum Associates, 1983.
40. Norusis MJ. *SPSS Base System User's Guide*. Chicago, Ill.: SPSS, 1990.
41. White HD, Griffith BC. Authors as markers of intellectual space: cocitation in studies of science, technology and society. *J Documentation*. 1982;38:255–72.
42. Coxon APM, Davies PM. *The User's Guide to Multidimensional Scaling: With Special Reference to the MDS(X) Library of Computer Programs*. Exeter, N.H.: Heinemann Educational Books, 1982.
43. McCain KW. Mapping authors in intellectual space: population genetics in the 1980s. In: Borgman CL (ed). *Scholarly Communication and Bibliometrics*. Newbury Park, Calif.: Sage Publications, 1990:194–216.
44. Helm CE. Multidimensional ratio scaling analysis of perceived color relations. *J Optical Soc Am*. 1964;54:256–62. Described in Carroll JD, Kruskal JB. *Scaling, multidimensional*. In: Kruskal WH, Tanur JM (eds). *International Encyclopedia of the Statistics*. New York: Free Press, 1978;2:892–907.
45. Sittig DF, Kaalaas-Sittig J. A quantitative ranking of the biomedical informatics serials. *Methods Inf Med*. 1995;34:397–410.
46. Sittig DF. Identifying a core set of medical informatics serials: an analysis using the MEDLINE database. *Bull Med Libr Assoc*. 1996;84:200–4.
47. Garfield E. *Mapping the world of biomedical engineering: Alza Lecture (1985)*. *Ann Biomed Eng*. 1986;14:97–108.
48. Bommel JH, McCray AT (eds). *Yearbook of Medical Informatics: Advanced Communications in Health Care*. Geneva, Switzerland: International Medical Informatics Association; Stuttgart, Germany: Schattauer, 1994. 1995 issue also examined.
49. McCain KW. *Longitudinal Cocited Author Mapping and Intellectual Structure: A Test of Congruence in Two Scientific Literatures* [PhD thesis]. Philadelphia, Pa.: Drexel University, 1985.
50. Collen MF. Origins of medical informatics. *West J Med*. 1986;145:778–85.



## ERRATUM

### Omission of Author's Name

The following name was omitted from the list of authors of "Development of the Logical Observation Identifier Names and Codes (LOINC) Vocabulary," published in the May/June issue of *JAMIA*<sup>1</sup>:

Douglas Martin, MD  
 Associate Professor of Medicine  
 Indiana University School of Medicine and Department of Veterans Affairs  
 Indianapolis, Indiana

1. Huff SM, Rocha RA, McDonald CJ, et al. Development of the Logical Observation Identifier Names and Codes (LOINC) Vocabulary. *J Am Med Inform Assoc*. 1998;5:276–92.