

Searching chiropractic literature: a comparison of three computerized databases. By: Aker PD, McDermaid C, Opitz BG, White MW, *Journal of Manipulative & Physiological Therapeutics*, 01614754, Oct96, Vol. 19, Issue 8

ABSTRACT

Purpose: To determine the efficiency of three computerized bibliographic databases in retrieving literature relevant to the chiropractor.

Methods: A cross-sectional design was used. English-language citations from 1990-92, on the topics of scoliosis, sciatica and neck pain, were searched in CHIROLARS, Index to Chiropractic Literature (ICL) and MEDLINE. Citations were assessed for relevance criteria by two assessors; a third assessor was used when both were unsure of relevance. Inter- and intraexaminer reliability of the relevance assessments was determined using the weighted Kappa statistic. The outcomes assessed were search time, search costs, number of citations, relevance, number of unique citations and the number of citations from refereed journals. Relative recall and cost per citation were used as primary measures of database efficiency.

Results: A total of 846 citations were retrieved. After exclusions, 786 citations were assessed for relevance. Interexaminer reliability of the relevance assessments was moderate [$K(w)$ (standard error) = 0.46 (0.03)]. Intra-examiner reliability was fair for each of the assessors [$K(w)$ (SE) = .36 (.10) and .35 (.10), respectively]. Of the 385 relevant citations, CHIROLARS retrieved 88 (relative recall = 23%) at a cost of CAN\$1.01 per relevant citation, ICL retrieved 37 (relative recall = 10%) at CAN\$.65 per relevant citation, and MEDLINE retrieved 260 (relative recall = 68%) at CAN\$.52 per relevant citation.

Conclusions: MEDLINE was found to be the most efficient database to search for literature relevant to the chiropractor; it retrieved the highest proportion of relevant citations and was the least expensive. CHIROLARS was the second most efficient of the three databases. No single database can function as a stand-alone source of information. For comprehensive searching, having an experienced reference librarian search MEDLINE in combination with at least one other database is recommended. (*J Manipulative Physiol Ther* 1996; 19:518-24).

Key Indexing Terms: Abstracting and Indexing; Databases, Bibliographic; Chiropractic; Comparative Study; Medline.

INTRODUCTION

To stay current with advances in health care, chiropractors, like other health care professionals, need to create time for personal education and advancement as part of a busy practice week. Lifelong, self-directed learning is an important component of professional conduct. Although reading scholarly periodicals or attending continuing education seminars are considered useful educational strategies, they may not satisfy a practitioner's need for clinically relevant, condition-specific information. As well, these traditional educational methods are resource intensive, in terms of time and money (1).

With recent advances in computer technology, it is now possible for the clinician to access information on a variety of topics from the home, office or nearest health sciences library. Computerized searching of literature databases has been shown to be more efficient than manual searching, not only in medicine (2) but also in chiropractic (3). Field practitioners and researchers should be encouraged to use these more modern forms of data gathering.

With the number of databases (including chiropractic databases) increasing, selection of a database can be an overwhelming task. Mychko-Megrin suggests 12 criteria to assess biomedical databases (4); however, these criteria were primarily designed to assess the coverage of the database, not whether the retrieved material was relevant to the searcher. The issue of relevance is pivotal in determining the effectiveness of a database search. However, relevance is determined subjectively and will vary with the expected use of the material. For example, literature relevant for use in diagnosis may differ from that used for treatment or prognosis. Establishing relevance has been an important factor in a number of studies of databases (5-8).

One study was retrieved that investigated chiropractic literature databases (3). The efficiency of CHIROLARS, a privately owned (by Action Potential, Denton, TX) computerized bibliographic database of chiropractic literature, Index to Chiropractic Literature (ICL), a bibliographic database of chiropractic literature produced by the Chiropractic Libraries Consortium, and MEDLINE, a general-purpose biomedical research literature database operated by the National Library of Medicine, were compared. Although Curl and Shapiro found CHIROLARS to be the most effective database, they did not determine whether the citations retrieved in their searches were relevant to the practicing chiropractor (3).

Previously performed studies on database efficiency either did not assess chiropractic literature or did not determine the relevance of the retrieved citations to chiropractors. We therefore found it necessary to initiate a new study. The purpose of this study was to determine which of three readily available computerized databases (CHIROLARS, ICL, MEDLINE) would be the most efficient source of relevant literature for chiropractors.

MATERIALS AND METHODS

A cross-sectional design was used. A test-retest reliability study was incorporated within this design to assess one of the outcome measures. The bibliographic databases selected were CHIROLARS, ICL and MEDLINE. These databases are readily available for access by students and field practitioners at the C. C. Clemmer Health Sciences Library at the

Canadian Memorial Chiropractic College (CMCC). MEDLINE and ICL were available on CD-ROM and fixed disk, respectively. CHIROLARS was accessed via modem.

Sciatica, scoliosis and neck pain were selected as topics to search. This selection was made by the senior investigator based on issues of relevance to clinical practice in chiropractic. Lists of subject headings unique to each database were used to generate the search strategies. "Sciatica" and "scoliosis" were search terms used for the first two topics. Because "neck pain" is not listed as a Medical Subject Heading (MESH) for use with MEDLINE, the terms "neck" or "cervical vertebrae" were combined with the term "pain" by using Boolean operators to create a search string to cover this subject. Searches were run for indexed material from 1990 to 1992 inclusive and were limited to English language citations involving human subjects.

All searches were conducted at the CMCC Library by a reference librarian experienced in searching all three databases. To ensure that the searching would be conducted with the same diligence between databases, the librarian was kept blinded to the purpose of the study. The citations generated from the search were printed while on-line, including the title, author, journal and abstract (if available). The citations were reviewed by one investigator for these exclusion criteria: conference proceeding citations were excluded because not all databases indexed them, and letters to the editor were excluded because their titles are reflective of the initial report, not of their content. The name of the database was removed from the printouts before distribution to the assessors.

A variety of outcome measures were selected to allow a thorough comparison among databases. The search time was determined by the total access time for each database. The search costs for MEDLINE and ICL were determined by multiplying the access time by a fixed rate as set by the CMCC Library. Estimates to the nearest minute were used to determine access time for ICL; precise access time was recorded by CHIROLARS and MEDLINE. The cost for CHIROLARS was as billed by the vendor, which included a combination of a flat-rate on-line charge, a charge per citation and a telecommunication charge (a 9600 bps modem was used). The cost per citation was calculated by dividing the cost of the search by the number of citations generated within each database.

Relevance was determined solely on the contents of the citation, including the abstract where available. This method is used commonly in the assessment of computerized bibliographic information (6, 7, 9). A citation was relevant if it met any one of these three criteria: (a) the citation specifically mentioned a chiropractic approach in terms of assessment or treatment of a contemporary or historical nature, regardless of the main focus of the citation; (b) the citation indicated a substantial (main focus) description of natural history, prognostic factors or etiologies; or (c) the citation indicated a substantial description of prognostic factors of other interventions with or without mention of chiropractic interventions. Two chiropractors, one a field practitioner and one a clinical instructor and field practitioner, each with more than 8 yr of clinical experience, acted as assessors of the relevance of the citations. Using the above criteria, they scored relevance using a 3-point scale: yes, no or unsure. Dichotomous and 7-point scales have been successfully used to assess relevance of literature citations in other studies (7, 8). If either one of the assessors indicated that the citation was relevant, it was scored as relevant. Only if both assessors indicated that the citation was not relevant was it scored as not relevant. If both assessors indicated they were unsure whether the citation was relevant, or if one assessor indicated the citation was not relevant and the other was unsure, the citation was sent to a third experienced clinician for another assessment. The third assessor was instructed to score the citations as relevant or not relevant, using the same relevance criteria.

The relevance scores of the initial two assessors were compared and tested for agreement. This allowed for a determination of interexaminer reliability. To assess intraexaminer agreement, the relevance scores of the citations identified by more than one database were analyzed. All citations were distributed in a nonrandom fashion, grouped by database. This, combined with the sheer bulk of citations being reviewed by the assessors, made memorizing previous scores improbable.

Relative recall (precision) and sensitivity are used to describe the ability of a database to retrieve relevant citations. Relative recall describes the proportion of relevant citations retrieved by the database compared with the relevant citations retrieved by all searches. This number, expressed as a percentage, relates the precision of the database in retrieving relevant citations and is calculated by dividing the number of relevant citations retrieved by the database by the total number of relevant citations from all searches. Sensitivity is calculated by dividing the number of relevant citations by the number of citations retrieved by the database and describes the proportion of relevant citations within a database. To assess the degree of overlap in the searches, the proportion of relevant citations unique to each database was determined. This was calculated by dividing the number of unique relevant citations per database by the number of relevant citations. Citations were described as unique if they were retrieved by only one database. A database with a high proportion of unique citations may be preferred as search results are less redundant. The citations were also assessed to determine the proportion arising from refereed journals, calculated as the number of relevant citations from refereed journals divided by the number of relevant citations ($\times 100\%$). Medical and Health Care Books and Serials in Print, 1994 was used as a source to identify whether a journal was refereed (10). Serials not listed in this text were scored as being not refereed. Nonrefereed sources may be of lower quality because their articles have not undergone a peer-review process (11). Finally, a retrieval rate for each database was calculated by dividing the number of relevant citations by the search time in minutes. This rate, similar to the time-efficiency index described by Curl and Shapiro (3), was used to measure the time taken to retrieve relevant citations from each database, expressed in citations per minute. High retrieval rates are preferred.

An efficient database yields relevant and comprehensive information at minimal cost (5). Because relative recall incorporates measures of relevance and comprehensiveness, we used it together with the cost per citation as primary measures of database efficiency. Other outcomes were used secondarily to describe the strengths and weaknesses of the databases.

Statistical Analysis

Inter- and intraexaminer reliability of the relevance assessments were determined using quadratic-weighted Kappa estimates. This statistic assigns weights to the magnitude of observed agreement and corrects for chance-expected agreement. It is the statistic of choice when measuring agreement with categorical or ordinal data (12, 13). Using this statistic, Kappa values of less than 0 are described as poor agreement, 0-.2 slight, .21-.4 fair, .41-.6 moderate, .61-.80 substantial and .81-1.00 as almost perfect (14). Standard errors were calculated to describe the variance about these agreement estimates.

RESULTS

A total of 849 citations were retrieved. Table 1 shows the distribution of the retrieved citations and the costs for each search topic before exclusion criteria were imposed. Sixtythree (7%) citations from conference proceedings or letters to the editor were excluded: 52 from MEDLINE, 9 from CHIROLARS and 2 from ICL. Of the 786 remaining citations, the majority were retrieved by MEDLINE (617; 78%) followed by CHIROLARS (130; 17%) and ICL (39; 5%). Table 2 summarizes the search outcomes across databases.

Slightly more than 90 min were spent in access time searching and printing from MEDLINE; nearly 25 min was spent in CHIROLARS, and 16 min in ICL. Although the cost for MEDLINE searching was the highest, at \$135.00 (all costs are in Canadian dollars), the high number of retrieved citations resulted in this database achieving the lowest cost per citation (\$.22). CHIROLARS had the highest cost per citation at \$.68 and ICL had a cost of \$.62 per citation. The relatively high costs per citation in CHIROLARS and ICL reflect the low number of citations retrieved by each of these databases.

All 786 citations were scored for relevance. For 666 (85%) citations, the assessors were able to agree on the determination of relevance: 306 (39%) were relevant and 360 (46%) were not relevant. For the remaining 120 citations (15%) at least one assessor was unsure of the relevance of the citation (see Table 3). Overall, the interexaminer reliability [weighted Kappa (standard error)] for these agreement scores was moderate at $K(w) (SE) = .46 (.03)$. The two assessors achieved fair intraexaminer reliability upon scoring 99 citations more than once. Assessor 1 rated the relevance of the citations the same for 43 of the 99 citations (43%) ($K(w) (SE) = .36 (.10)$); assessor 2 rated relevance the same for 51 of the 99 citations (52%) ($K(w) (SE) = .35 (.10)$).

The 120 citations that assessors scored as unsure were given to the third assessor to force a relevance score. From these, 79 were scored as relevant and 41 as not relevant, resulting in identification of a total of 385 relevant citations. Relative recall (precision) of each of the databases was: 68% for MEDLINE, 23% for CHIROLARS and 10% for ICL. Sensitivity within each database was 95% for ICL, 68% for CHIROLARS and 42% for MEDLINE. CHIROLARS and ICL were more costly to search, at \$.68 and \$.62 per relevant citation, respectively. MEDLINE was the least expensive at \$.52 per relevant citation. The database retrieving the highest proportion of citations from refereed journals was CHIROLARS (81%), followed by MEDLINE (68%) and ICL (19%). MEDLINE had the highest proportion of unique relevant citations (80%), ICL had 78% and CHIROLARS had 38%. CHIROLARS had the highest retrieval rate at 3.45 citations per minute, compared to retrieval rates of 2.85 for MEDLINE and 2.31 for ICL. Using relative recall and cost as primary determinants of efficiency, MEDLINE was found to be the most efficient database at finding relevant literature citations pertaining to the practicing chiropractor, followed by CHIROLARS then ICL.

DISCUSSION

Inferences from this study can only be made if the limitations of the study design are fully understood. A search of any computerized database is not likely to provide the searcher with all relevant literature on a subject. At best, a searcher will retrieve a portion of the available information. Currently, the standard for retrieving literature is still systematic hand-searching (15). In studies comparing systematic hand-searches to computer searches, computer searches have been found to miss between 20 to 60% of the available randomized controlled trials (RCTs) (16).

Retrieval can be improved by optimizing the search parameters and techniques according to the capabilities of the databases. For example, retrieval rates (sensitivity) for articles on treatment have been shown to improve to 99% when combinations of indexing terms and keywords are used (17). The search parameters that were used to collect the data in this study were not dynamic, and no attempt was made to explore or cross reference other search topics. As well, search techniques that may have been available through interface software unique to the database (for example: keyword searching) were not used. The use of more dynamic search parameters and techniques may alter the results.

The use of different search topics might also have changed the results. Selecting topics that were of known interest to chiropractors should favor the chiropractic databases, but this was not evident in this study. A searcher's expertise and experience also has a significant effect on search results (6, 8, 9). The individual performing the searches in this study was an experienced research librarian. We recommend clinicians use an experienced reference librarian to conduct their searches; only a clinician with good database searching skills might expect to achieve better retrieval (6). Without this

kind of expertise, it is unlikely that searching would be as successful. Individuals conducting searches on their own must be aware of the influence of experience and searching expertise on retrieval rates.

Table I. Search retrieval by database before exclusions

	CHIROLARS			ICL			MEDLINE		
	Neck pain	Scoliosis	Sciatica	Neck pain	Scoliosis	Sciatica	Neck pain	Scoliosis	Sciatica
Citations	22	78	39	11	26	4	129	444	96
Search time (min:sec)	6:51	9:11	8:52	4:00	10:00	2:00	17:36	52:33	20:11
Search cost (\$ Canadian)	18.21	43.30	27.29	6.00	15.00	3.00	26.25	78.75	30.00

Table II. Search outcomes

	CHIROLARS	ICL	MEDLINE
Search time (min:sec)	24:55	16:00	90:20
Search cost (\$) ^a	88.80	24.00	135.00
Citations	130	39	617
Cost per citation (\$)	0.68	0.62	0.22
Relevant citations	88	37	260
Relative recall (precision) (%)	23	10	68
Sensitivity (%)	68	95	42
Cost per relevant citation (\$)	1.01	0.65	0.52
Proportion from refereed journals	81	19	68
Proportion unique to database	38	73	80
Retrieval rate (citations per min)	3.45	2.31	2.85

^a All costs in Canadian \$.

The completeness or success of indexing by the source database may also limit a searcher's ability to retrieve citations. Incompleteness or errors in citation indexing or misuse of keywords may result in the citation not being retrievable despite the use of good search techniques. This has been described previously in a MEDLINE search for RCTs on hyperbilirubinemia (9). Using a sound search strategy, only 30% of the citations contained in MEDLINE were recovered. After exploratory searches, it was found that problems existed in the manner with which the papers had been indexed. The authors concluded that an effective search depended on the adequacy of the indexing policy, the proper interpretation of the policy by the indexers and the quality of the publication (9). In another study on database searching in manual medicine, Aker (18) compared the success of computerized searching to a reference standard of comprehensive searching (a combination of computerized, manual and content expert search strategies). Seventeen of the nineteen (89%) available RCTs were retrieved using searches of four computerized databases. The two clinical trials not retrieved were available on the databases, but were indexed using different keywords and subject headings (18). Indexing errors are beyond the control of the individual performing searches, but are important to consider when perceiving the thoroughness of a search.

Measuring the reliability of the relevance assessments is a crucial aspect of this study, because many of the outcome measures used to determine database efficiency are contingent on a citation's relevance. Our relevance assessors were selected to be representative of the field practitioner and academic clinician. Both had extensive fulltime clinical experience, and one had additional teaching and academic skills. The assessors agreed on the determination of relevance 85% of the time ($K(w) = .46$). This acceptable level of interexaminer reliability indicates that we can have good confidence in the interpretation of the other outcomes. It also indicates that both academic and nonacademic practitioners identify the relevance of citations similarly.

Because identifying features were removed from the printouts, it was unlikely that the assessors could identify a database from a printout. Observer bias, therefore, would be unlikely to influence the results of this study. Future studies can better control for observer bias by importing the search results into a word processor to prepare printouts using the same typeface. Although ICL does not contain abstracts, it did not seem to influence the determination of relevance, as agreement between assessors was similar across databases.

Assessing a citation in the context of the ones recently scored may influence an assessor's decision on relevance compared with examination of the citation in isolation. Fatigue may also cause an assessor to score a citation differently; assessments may be expected to be more cursory at the end of the package than at the beginning. Although both of these interaction biases may have existed in this study, it is difficult to predict their direction of effect or to measure their presence. However, they can be controlled in future studies by randomizing the order of delivery of the citations and delivering the citations in smaller packages.

Intraexaminer reliability was fair for both assessors. At the upper end of the confidence intervals, these scores can be considered to be within an acceptable range of reliability, although higher agreement scores would be desirable. Most of the disagreement within observers was accounted for by changing decisions from relevant or not relevant to unsure (or vice versa); in only 13% of citations for assessor 1 and only 17% of citations for assessor 2 did scores change from relevant to not relevant (or vice versa). Using larger sample sizes to assess intraexaminer reliability would more accurately portray the point estimates of agreement, because the confidence intervals would be narrowed. Low sample sizes can cause differing scores between assessments to be amplified, which may decrease the Kappa coefficients. Future studies should use a larger sample of repeat citations for calculation of intraexaminer reliability.

Although overall database efficiency in this study was determined by two primary measures, relative recall and cost per citation, it is essential to look at the other outcomes to fully understand the strengths and weakness of each database. MEDLINE retrieved most (78%) of the citations available on the three search topics. This is not surprising, in that it is a large and well-established bibliographic database. The overall cost of searching MEDLINE may seem expensive compared with the other databases, but when the number of citations retrieved is considered and cost is expressed in cost per citation or cost per relevant citation, it becomes apparent that MEDLINE is much less expensive than the other databases. Search time and costs for MEDLINE might be even lower if citations were output onto diskette instead of to a printer, because with this magnitude of printing, our printer proved to be the rate-limiting step. Other strengths of MEDLINE were its ability to retrieve citations relevant to the chiropractor (relative recall = 68%), to identify a large proportion of unique citations (80%) and to retrieve citations quickly (2.85 citations per minute). Two relative weaknesses of MEDLINE were identified. First, MEDLINE has low sensitivity (42%) compared with the other databases (68% and 95% for CHIROLARS and ICL, respectively). This should not be surprising because it was not designed as a chiropractic database. However, the low sensitivity for MEDLINE is outweighed by the large proportion of relevant articles the database was able to identify overall. The second weakness was that only 68% of the MEDLINE citations arose from refereed journals. As MEDLINE is often perceived as one of the best sources of robust on-line information, this finding was unexpected. In general, the strengths of MEDLINE overwhelm its weaknesses.

One might expect that CHIROLARS would retrieve more relevant citations than MEDLINE, because MEDLINE indexes only one chiropractic journal, *The Journal of Manipulative and Physiologic Therapeutics* (19). However, it found only 23% of the total number of relevant citations (relative recall) compared with 68% for MEDLINE. This result is probably reflective of the size of the MEDLINE database compared with the relatively newer and less developed CHIROLARS database. Substantial changes have been made to the CHIROLARS database in the time since these searches were conducted. Future study will be needed to determine the influence of these changes on the results. Compared with ICL, CHIROLARS retrieved considerably more citations and more relevant citations. The main strength of CHIROLARS was its ability to retrieve citations relevant to chiropractors. This should not be surprising; it was designed for this purpose. The sensitivity of CHIROLARS (68%) was higher than MEDLINE (42%), but not as high as ICL (95%). Another strength was that CHIROLARS retrieved the highest proportion of citations (81%) from refereed journals, indicating that most of its citations on these topics arise from good quality journals. CHIROLARS also had the fastest citation retrieval rate. Retrieval rate is not as powerful an outcome as the others. A weaknesses in CHIROLARS was that it had the lowest proportion of unique citations (38%) compared with ICL (78%) and MEDLINE (80%). Many of the citations retrieved from CHIROLARS will therefore also be accessible on the other databases. Another considerable weakness of CHIROLARS is its cost. CHIROLARS cost almost twice as much (per relevant citation) as MEDLINE, and about 1.5 times as much as ICL. Should CHIROLARS become available on CD-ROM or fixed disk, as are MEDLINE and ICL, or should a higher speed modem be used for access, costs for searching this database might be lowered.

Of the citations retrieved by ICL, 95% were relevant and 73% were unique. The sensitivity and uniqueness of the database may be important to a searcher interested in completeness and are considered strengths of the database. Another strength of ICL is that it is not expensive to search; it is only slightly more expensive than MEDLINE. ICL was considered a weaker database than the others, primarily because of its inability to identify relevant citations. It only identified 10% of the total number of relevant citations available from all databases. Another considerable weakness is that only 19% of the citations from ICL were from refereed journals; the remaining 81% are not likely to be high-quality information. ICL also had the slowest retrieval rate. Overall, ICL was considered to be of marginal value compared with the other databases.

When reviewing the strengths and weaknesses of all three databases, MEDLINE is clearly stronger than the others. Both CHIROLARS and ICL each hold unique strengths that might be called upon in searches combining more than one database. When resources are limited such that only one database can be searched, we recommend searching MEDLINE. Clinicians or researchers may select a database other than MEDLINE depending on the outcomes they feel would be most important. However, this should be done with the understanding that the search would probably not be as comprehensive and would cost more.

Searching two databases will result in a more comprehensive search than one alone. In this study, searches of both MEDLINE and CHIROLARS increased relative recall to 90% for an additional \$89 above the cost of the MEDLINE search alone. In another study (18), Aker compared CHIROLARS, the Cumulative Index to Nursing and Allied Health Literature (CINAHL) (Information Systems, Glendale, CA), EMBASE (Elsevier Science Publishers, Amsterdam, The Netherlands) and MEDLINE for their ability to identify clinical trials in manual medicine. MEDLINE had a relative recall of 63%. Searching CHIROLARS and MEDLINE increased relative recall to 79%; searching EMBASE and MEDLINE increased relative recall to 88%.

Either CHIROLARS or EMBASE could serve as useful databases to search in conjunction with MEDLINE. Used alone, both databases had similar levels of relative recall, where citations were relevant if they were RCTs (18). EMBASE was considerably more sensitive, and CHIROLARS cost considerably less. Recommendations on choice of second database therefore depend on the goals of the search and on the resources available to the searcher. If funds are limited, CHIROLARS would likely be the better selection, but at the expense of greater time to identify the relevant from the nonrelevant citations. If time or human resources are constrained, EMBASE would probably be a better selection, but at considerably higher cost. Results from both of these studies suggest that searching CINAHL or ICL is only necessary for

completeness or if access to other databases is restricted. Further study is needed to determine EMBASE's ability to identify citations relevant to the chiropractor and to compare its efficiency to that of CHIROLARS.

CONCLUSION

MEDLINE was found to be the most efficient database for retrieval of citations relevant to the chiropractor; it retrieved a higher number of relevant citations at a lower cost. However, because at best one database retrieved only 68% of the available relevant citations, no one database should function as a stand-alone source of literature. To retrieve relevant information efficiently, we recommend the use of an experienced reference librarian. Searching of MEDLINE in combination with at least one other database will yield the best results.

In future studies, the use of "free-style" search techniques and parameters may allow a better assessment of the benefits that individual databases may build into their search interface and may give a more accurate depiction of database efficiency. Comprehensiveness and accuracy of indexing will continue to pose limitations on all searches.

REFERENCES

- (1.) Haynes RB, Ramsden M, McKibbin KA, Walker CJ, Ryan NC. *A review of medical education and medical informatics. Acad Med* 1989; 64:207-12.
- (2.) Sackett DL, Haynes RB, Guyatt GH, Tugwell P. *Clinical epidemiology: a basic science for clinical medicine. 2nd ed. Toronto: Little, Brown, 1991:335-58.*
- (3.) Curl DD, Shapiro CS. *Literature searching by a field doctor: a comparison of manual versus computerized methods. Chiro Technique* 1993; 5:15-22.
- (4.) Mychko-Megrin AY. *A comparison of biomedical databases. Bull Med Libr Assoc* 1991; 79:302-8.
- (5.) Smith PJ, Knodel LC, Bierschenk NF, Klaus JR. *Qualitative and quantitative evaluation of on-line databases: EMBASE vs. MEDLINE. Drug Intell Clin Pharm* 1987; 21:3A.
- (6.) Haynes RB, Ramsden M, McKibbin KA, Walker CI, Ryan NC. *Online access to MEDLINE in clinical settings: a study of use and usefulness. Ann Intern Med* 1990; 112:78-84.
- (7.) Biarez O, Sarrut B, Doreau CG, Etienne J. *Comparison and evaluation of nine bibliographic databases concerning adverse drug reactions. Ann Pharmacother* 1991; 25:1062-5.
- (8.) Rovers JP, Janosik JE, Souney PF. *Crossover comparison of drug information on-line databases vendors: DIALOG and MEDLARS. Ann Pharmacother* 1993; 27:634-8.
- (9.) Dickersin K, Hewitt P, Mutch L, Chalmers I, Chalmers T. *Perusing the literature: comparison of MEDLINE searching with a perinatal trials database. Control Clin Trials* 1985; 6:306-17.
- (10.) *Medical and health care books and serials in print 1994. New Providence, NJ: RR Bowker, 1994.*
- (11.) Relman AS. *Peer review in scientific journals: what good is it? West J Med* 1990; 153:520-2.
- (12.) Cohen J. *Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. Psychol Bull* 1968; 70:213-20.
- (13.) Kramer MS, Feinstein AR. *Clinical biostatistics: LIV. The biostatistics of concordance. Clin Pharmacol Ther* 1981; 29:111-23.
- (14.) Landis RJ, Koch GG. *The measurement of observer agreement for categorical data. Biometrics* 1977; 33:159-74.
- (15.) Dickersin K, Scherer R, Lefebvre C. *Identification of relevant studies for systematic overviews. Br Med J* 1994; 309:1286-91.
- (16.) Adams CE, Power A, Frederick K, Lefebvre C. *An investigation of the adequacy of MEDLINE searches for randomized controlled trials (RCTs) of the effects of mental health care. Psychiatr Med* 1994; 24:741-8.
- (17.) Haynes RB, Wilczynski N, McKibbin KA, Walker C J, Sinclair J. *Developing optimal search strategies for detecting clinically sound studies in MEDLINE. J Am Med Inform Assoc* 1994; 1:447-58.
- (18.) Aker PD. *Database searching in manual medicine. Proceedings of the Second International Cochrane Collaboration; Hamilton, Ontario, Canada, 1994.*
- (19.) *National Library of Medicine. List of serials indexed for online users. Springfield, VA: U.S. Department of Commerce, National Technical Information Service, 1994.*