An SDILINE Profile Oriented to Patient Care

By Mark E. Funk, *Clinical Medical Librarian*

University of Missouri–Kansas City
School of Medicine
Kansas City, Missouri

ABSTRACT

Hospital physicians naturally require current medical information in order to treat their patients. An objectively obtained SDILINE storesearch was devised to obtain this information. A rank-ordered list of hospital discharge diagnoses, coded with *ICDA* terms, was translated into MeSH. The top two Bradford zones of journals defined as useful at the University of Missouri–Kansas City were ANDed to these MeSH terms. Citations retrieved from the storesearch are input into an in-house computerized data base. The method can be easily used by other medical or hospital libraries without access to computers.

CURRENT information from the medical literature is very important to practicing physicians. Time factors, however, often allow these physicians to scan only a few medical journals. Various devices and services have been created to get new and pertinent information to physicians. These range from current awareness journals such as *Current Contents*® to the concept of clinical medical librarians (CMLs) where librarians "live with the user" to actively provide needed information in the form of bibliographies or photocopied documents. The CMLs at the University of Missouri–Kansas City (UMKC) Medical Library have developed a current awareness system designed for the specific needs of a specific hospital. While used at UMKC in an in-house computerized retrieval system, the method could be used in a variety of ways by other medical libraries.

Background

As reported earlier [1], the CMLs at UMKC have developed a Clinical Document Citation File (CDCF).* This is an indexed list of journal articles chosen by the CMLs which have proven useful in filling information needs occurring on in-patient rounds, such as treatment protocols, disease complications, and so on. CDCF exists in a printed form, and in an on-line form, called DOCLINE (DOCuments-on-LINE). CDCF and DOCLINE are created by using the FAMULUS system of computer programs.

FAMULUS was developed at the Pacific Southwest Forest and Range Experimental Station of the U.S. Department of Agriculture at Berkeley, California. It provided for computerized, batch-mode maintenance of personal reference collections maintained by researchers. Various programs allowed numerous ways of sorting, printing, and thus indexing a file. Additionally, a file could be searched using Boolean logic to combine concepts, but only in the batch mode [2]. The interactive version of FAMULUS, called Conversational FAMULUS, is suitable for on-line searching. It was developed by the University of Missouri–Columbia School of Library and Information Science.

Input to DOCLINE consists of citations found manually, from MEDLINE searches, and from DOCLINE searches (indicating reuse of items). About 200 to 300 citations are added monthly. At regular intervals, older citations that have not been reused are weeded. These procedures have led to a fairly stable file of 5,000 to 6,000 citations.

Figure 1 shows the unit record of a citation in DOCLINE. Field labels for the unit record include author, title, source, journal code (which includes indication of *Abridged Index Medicus* status), publication year, document number (indicating which CML submitted the citation and when; multiple numbers indicate reuse of the item), and MeSH terms. All of the fields are searchable, although most searches use combinations of title text words and MeSH terms. Because of the limited subject scope of the file, only print MeSH terms (those under which an article appears in *Index Medicus*) are included in the unit record, as well as any nonprint heading deemed vital by the CML submitting the citation.

*Development of the Clinical Document Citation File was supported in part by Public Health Service Grant No. 5 R01 LM01574-03 from the National Library of Medicine.

CIT#	0171
AU	BORDEN C W
TI	THE CURRENT STATUS OF THERAPY WITH ANTICOAGULANTS. (REVIEW)
SO	MED CLIN NORTH AM 56 235-53 JAN 72
JC	LU6 / AIM
PY	72
DN	740476002 / 740496003 / 740879001 / 741247001
MH	ANTICOAGULANTS/THERAPEUTIC USE, THROMBOEMBOLISM/DRUG THERAPY, CEREBROVASCULAR DISORDERS/DRUG THERAPY, HEPARIN/ADMINISTRATION & DOSAGE, PULMONARY EMBOLISM, HEPARIN, WARFARIN

FIG. 1.—DOCLINE unit record.

The guiding principle behind the construction and use of DOCLINE is that a small file of selected, *clinically* oriented documents is more useful to health care team members than MEDLINE alone. The CMLs search DOCLINE to provide citations on a wide variety of clinical information needs. Additionally, faculty and students have access to DOCLINE for personal searching on seven CRT terminals throughout the medical school.

While a search of DOCLINE yields relatively recent articles, problems arise that raise questions about the efficiency and effectiveness of the system. To avoid missing the most recent articles on a subject, a CML would usually do a follow-up MEDLINE search. And often, because of time restraints and some software limitations of the DOCLINE system, MEDLINE would be the only data base used. This underutilization of DOCLINE (a relatively inexpensive and more clinically oriented data base) and overutilization of MEDLINE was cause for growing concern.

Another impediment to more active use of DOCLINE has been the rather cumbersome method of adding citations to the data base. CMLs have had to initiate additions to DOCLINE by filling out a citation form. This is tedious, and unfortunately has become somewhat neglected. A clerk then uses this form to pull unit records from MEDLINE for indexing, a process that takes approximately three to four hours of on-line searching per month. The MEDLINE unit records are then adapted to the FAMULUS format and entered on-line into a TSO (Time Sharing Option) dataset in the university's computer. A TSO dataset is a temporary "holding" file maintained by the computer. Data may be entered, modified, and retrieved, but are not directly searchable. Once the dataset is large enough, it is edited for corrections and eventually added to the printed and on-line version for searching. This long process has only served to lengthen the time gap between DOCLINE and MEDLINE, further discouraging DOCLINE use.

To improve the system, we needed to (1) cut down the use of MEDLINE for searching and unit record pulling; (2) decrease the time gap between MEDLINE and DOCLINE; and (3) reduce the amount of time spent filling out forms.

To do all this we have devised a method of automatically selecting recent, clinically oriented articles (and unit records) for input to DOCLINE by using SDILINE and the storesearch capability of MEDLARS at the National Library of Medicine. The search strategy used was designed to fit the specific needs of our user group.

Search Strategy Formulation

As DOCLINE contains citations resulting from over three years of information requests obtained on rounds, it necessarily reflects the long-range patient population of the hospital. This population is unlikely to change drastically, so it can be used to predict future information requests. An analysis of the disease states of this population would produce an objective, easily obtained search strategy.

From the Medical Records Department we obtained a frequency count of discharge diagnoses from the internal medicine units for three months. This was felt to be a representative sample. From this was produced a rank-ordered list of diseases as coded from the *International Classification of Diseases, Adapted for Use in the United States,* eighth revision (*ICDA*) [3]. The entire list was not used. Decreasing frequencies of diseases were used to a point where the CMLs felt that very few disease information requests would be made. As closely as possible, applicable *ICDA* descriptors were then translated into MeSH. Interpretations, rather than transliterations, often had to be used. *ICDA* will often code a nonspecific entity such as "pneumonia, unspecified." Liberal use of explosions was made. Table 1 shows the rank-ordered list of *ICDA* diseases and the MeSH terms and explosions we used.

When entered into the computer, these terms and explosions were ORed together logically, forming a search "hedge," or horizontally related

TABLE 1

RANK-ORDER *ICDA* CODES AND MeSH TERMS

Diabetes mellitus without mention of acidosis or coma	EXP diabetes mellitus
Essential benign hypertension	EXP hypertension
Congestive heart failure	EXP heart failure, congestive
Urinary tract infection NEC	EXP urinary tract infections
Chronic ischemic heart disease without mention of hypertensive disease	EXP coronary disease
Alcoholic addiction	EXP alcoholism
Pneumonia, unspecified	EXP pneumonia
Electrolyte disorders	EXP water-electrolyte imbalance
Anemia, unspecified	EXP anemia
Other diseases of the lung	EXP lung diseases, obstructive
Asthma	Asthma
Osteoarthritis and allied conditions	Osteoarthritis
Diabetes mellitus with mention of acidosis or coma	—
Obesity not specified as of endocrine origin	EXP obesity
Convulsions	Convulsions
Iron deficiency anemias	—
Atrial fibrillation or flutter	EXP arrhythmia
Chronic nephritis	Nephritis Glomerulonephritis
Generalized and unspecified arteriosclerosis	Arteriosclerosis
Bronchopneumonia, unspecified	—
Emphysema	—
Other renal disease	Kidney diseases Kidney failure, acute Kidney failure, chronic
Other and unspecified alcoholism	—
Chronic ischemic heart disease	—
Other and unspecified heart disease	—
Diverticula of the colon	Diverticulosis, colonic Diverticulitis, colonic
Other pyelonephritis, pyelitis, and pyelocystitis	Pyelonephritis
Pulmonary embolism and infarction	Pulmonary embolism
Chronic bronchitis	Bronchitis Bronchial diseases
Pulmonary heart disease	Pulmonary heart disease
Other and unspecified ulcer of the duodenum	Duodenal ulcer
Hemoptysis	Hemoptysis
Pulmonary tuberculosis NOS	Tuberculosis

NOTE: Absence of corresponding MeSH term indicates that this concept was used in an earlier explosion.

group of MeSH terms, as opposed to the hierarchical "tree" structure. All terms were preceded by an asterisk to indicate central concepts only, and no subheadings were applied.

Obviously, using the hedge exactly as it is shown in Table 1 would retrieve thousands of citations each month from SDILINE. Many would not be clinically oriented, and many would be from journals not available in the UMKC Medical Library. We needed to restrict retrieval to those journals "proven" to be useful in DOCLINE. An analysis of the 3,161 citations contained in DOCLINE in 1976 showed that 410 journals had been used. Many of these journals, however, had contributed very few citations to DOCLINE, and to include all of these titles would retrieve many articles of little potential usefulness.

To identify the most productive journals for our search strategy, a Bradford distribution analysis was performed. The Bradford distribution is a phenomenon describing the dispersion over a set of journals of articles relating to a certain subject. Bradford described his discovery in 1948 in this way:

> 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 or zones containing the same numbers of articles as the nucleus, when the numbers of periodicals in the nucleus and succeeding zones will be as $1:n:n^2 \ldots$ [4].

Bradford's law was originally applied to bibliographies on a specific subject, such as vitamins. Since then, the distribution has been found to hold for general periodical circulation in a medical library [5]. We felt that Bradford's law would work well in identifying the nucleus of journals and succeeding zones in DOCLINE. Table 2 lists in rank order zones 1 and 2, as derived from apply-

ing a simple three-zone division of the journals in DOCLINE. Zone 1 journals contributed one-third of the citations in DOCLINE. Likewise, zone 2 journals also contributed one-third. All journals were available in the UMKC Medical Library.

After identifying the disease states about which we wanted information, and a set of useful journals, there followed several months of experimentation. In the first trial, the list of MeSH terms was ANDed to zone 1 of the journals, and the search was run against SDILINE. Off-line prints containing author, title, source, and main headings were routed to the five CMLs serving the Medicine units. They initialed citations which were felt to be "typical" of citations used in the past to satisfy information requests, and they also added comments on the search strategy as a whole. An analysis of the uninitialed citations as well as the subjective comments led us to modify the search strategy. Because of their low likelihood of being used, we negated letters, editorials, and case reports. This action drastically reduced the number of citations retrieved, and so allowed us to expand our search by adding zone 2 of our Bradford journal list.

TABLE 2

RANK-ORDER JOURNAL FREQUENCY USE

Zone 1	N. Engl. J. Med.
	Am. J. Obstet. Gynecol.
	JAMA
	Ann. Intern. Med.
	Br. Med. J.
	Lancet
	Med. Clin. North Am.
	Obstet. Gynecol.
	Am. J. Med.
	Chest
	Am. J. Surg.
	Postgrad. Med.
	Arch. Intern. Med.
	Circulation
Zone 2	Cancer
	South. Med. J.
	Arch. Surg.
	Med. J. Aust.
	Pediatrics
	Am. J. Cardiol.
	Ann. Surg.
	Surg. Clin. North Am.
	Am. Heart J.
	Am. Rev. Respir. Dis.
	Gastroenterology
	J. Pediatr.
	Diabetes
	DM
	Am. Surg.
	Surg. Gynecol. Obstet.
	West. J. Med.
	Br. J. Surg.
	Can. Med. Assoc. J.
	Geriatrics
	Adv. Intern. Med.
	Am. Fam. Physician
	Am. J. Nurs.
	N.Y. State J. Med.
	Prog. Cardiovasc. Dis.
	Am. J. Dis. Child.
	Cardiovasc. Clin.
	Practitioner
	Br. J. Obstet. Gynaecol.
	Obstet. Gynecol. Surv.
	S. Afr. Med. J.
	Am. J. Gastroenterol.
	Mayo Clin. Proc.
	Radiology
	Annu. Rev. Med.
	Drugs
	J. Reprod. Med.
	Minn. Med.
	Am. J. Roentgenol.
	Br. Heart J.
	Surgery
	Curr. Med. Res. Opin.
	J. Neurosurg.
	J. Thorac. Cardiovasc. Surg.
	Semin. Hematol.

SEARCH STORAGE

The final version of our search strategy was run on-line for several months, until we were satisfied with its workability. Since then we have taken advantage of the automatic SDI service at NLM. This is a service available to MEDLINE centers since August 1976. Searches are input one time, stored by the computer, and automatically run against the new SDILINE file each month. (See *Library Network/MEDLARS Technical Bulletin* no. 86, June 1976, for details.)

Three different off-line print formats are available from this storesearch capability. We chose the PRINT FULL option, which includes Medical Subject Headings as well as the necessary author, title, and source. This option has eliminated the need to pull unit records from MEDLINE at a later date.

As the monthly SDILINE off-line prints arrive, the printout is first checked by a CML for obvious errors. It is then handed to a clerk who starts putting the results into a TSO dataset. Within a short time, the results are available for searching.

Results

The number of citations retrieved each month with our stored search has stayed close to 200. The number will of course vary because of different journal publication frequencies. The cost of our search is very reasonable. Storage of the search itself costs $1.00 a month, and off-line print page costs have been averaging $5.00 a month. If these automatically selected citations become the major part of our DOCLINE input, as we hope, our on-line MEDLINE costs should substantially decrease. At any rate, we seem to have achieved our goals of keeping DOCLINE current, reducing the amount of time spent on MEDLINE, and reducing the time spent filling out forms.

We are currently working on evaluating the use of the automatically retrieved citations as compared to those input under the old system, and hope to be able to report on their effectiveness. The search strategy itself will also need to be reevaluated on a periodic basis. As with any acquisitions profile, it should not remain static, but always be responsive to users' needs. Yearly examination of medical records and journal usage will allow us to see shifts in disease states and journal usefulness.

Other Applications

While the technique described here was developed for a specific library, variations of it can be used by other medical libraries. All medical libraries have a large population of interns, residents, house staff, and other physicians who seldom have time to make use of library services. If a library cannot support a clinical medical librarian program, an adaptation of our patient care SDILINE profile may prove useful in reaching these users. A custom-designed SDI search which matches patient population *and* useful available journals would make an attractive outreach package.

It is a relatively easy matter to obtain statistics from the Medical Records Department, especially if these records are computerized. A list of useful journals can be obtained from circulation statistics, from recommendations, or from standard lists.

Even if a medical library does not have an in-house computerized retrieval system, other options remain for information access. Monthly printouts could be routed to staff or simply left in a staff lounge. Printouts could be cut up, mounted on cards, and maintained in a file box for ready reference. Smaller, more tailored search strategies could be created for specific departments, such as obstetrics, pediatrics, or surgery. Use could also be made of subheadings and checktags.

Conclusion

Two devices are described which may be used to predict and satisfy future information needs of hospital physicians: analysis of disease states and identification of useful journals. By using these devices, a medical library can take full advantage of the concepts behind selective dissemination of information and current awareness by emphasizing both selectivity and currency.

Acknowledgments

The author gratefully acknowledges the assistance of Gary D. Byrd, Chief Medical Librarian, for his editorial assistance and encouragement; and Carolyn Anne Reid, Senior Clinical Medical Librarian, for her analyses of disease frequencies and journal usage.

References

1. Algermissen, V. Biomedical librarians in a patient care setting at the University of Missouri–Kansas City School of Medicine. Bull. Med. Libr. Assoc. 62: 354–358, Oct. 1974.
2. Lancaster, F. W., and Fayen, E. G. Information Retrieval On-Line. Los Angeles, Melville Publishing Co., 1973.
3. International Classification of Diseases, Adapted for Use in the United States. 8th rev. Washington, D.C., National Center for Health Statistics, 1967.
4. Bradford, S. C. Documentation. London, Crosby Lockwood, 1948.
5. Goffman, W., and Morris, T. G. Bradford's law and library acquisitions. Nature 226: 922–923, June 6, 1970.

Received September 2, 1977; revision accepted December 7, 1977.