

The mini MEDLINE SYSTEM™*: A Library-Based End-User Search System†

BY NAOMI C. BROERING

*Dahlgren Memorial Library
Georgetown University Medical Center
3900 Reservoir Road, N.W.
Washington, D.C. 20007*

ABSTRACT

The miniMEDLINE SYSTEM, a user-friendly search system developed in 1981 at the Georgetown University Medical Center, has been operational since 1982. The system is designed to meet the immediate educational and clinical information needs of students, residents, and faculty. This article focuses on system planning and design, database creation through "downloading," hardware adaptation, and system use. The database is a subset of the NLM's MEDLINE file; it includes over 180,000 citations to articles indexed in over 160 journals from 1982 to the present. With only a few keystrokes in a two-step process it allows users to conduct bibliographic searches. The system is being replicated at eight other medical center libraries.

COMPUTER-ASSISTED bibliographic searches have traditionally been conducted by librarians and information analysts because the complex file structure of large databases has required trained searchers for successful retrieval. However, end-user products currently becoming available enable library patrons to conduct their own searches. The increasing number of users with home or office computers represent a previously untapped market for new information products. In addition, specialized institutional databases are being designed to meet in-house user needs for frequent searching from selected sources. In-house systems are being created by "downloading" (transferring records from major databases) or by acquiring tapes of selected portions of files from database developers. Libraries can capitalize on and benefit from these recent trends by making user-friendly database searching available at the

library and by teaching users how to access information through computers.

Georgetown University became involved in both these activities by implementing the miniMEDLINE SYSTEM, a self-service bibliographic search component of the Library Information System (LIS). The miniMEDLINE SYSTEM has been operational at the Dahlgren Memorial Library of Georgetown University since 1982, and it seems appropriate during its second year to share our experiences by describing pertinent aspects of the system, including database creation, hardware and software operations, and end-user instruction. Sufficient data gathered since inception of the system allow us also to report some observations of end-user reactions.

DESCRIPTION OF THE MINIMEDLINE SYSTEM

The miniMEDLINE SYSTEM is a very user-friendly online bibliographic database, developed as a separate component of LIS. It contains approximately 180,000 citations to articles indexed in over 160 key clinical and basic science journals from January 1982 to the present. It is a subset of the National Library of Medicine's (NLM) MEDLINE file, and is updated monthly from that master file.

The system is self-instructional. It can be learned in only a few minutes because it accepts natural language. This is made possible through the interface with the NLM MeSH (Medical Subject Headings) terms for retrieval of article citations. The miniMEDLINE SYSTEM is offered as a local in-house search service to meet immediate health information needs rather than to address the more intensive, scholarly needs of researchers. It is used by students, young researchers, residents, and faculty members seeking quick answers to educational or clinical questions that might not justify the costly and less convenient MEDLINE search. Scientific researchers prefer and, in fact, require

*The miniMEDLINE SYSTEM is a trademark of Georgetown University.

†This manuscript is an expanded and updated version of "miniMEDLINE: Observations to User Reaction," a paper presented June 2, 1983, at the Eighty-third Annual Meeting of the Medical Library Association, Houston, Texas.

THE miniMEDLINE SYSTEM

traditional, sophisticated searches from more comprehensive databases such as MEDLINE that are performed by trained information analysts and librarians. Because the system is maintained in the library's dedicated computer, it is available all hours the library is open and practically twenty-four hours a day for users with home computers. It is provided free to registered library borrowers, including Georgetown University faculty, staff, students, residents and fellows [1].

The miniMEDLINE SYSTEM is not meant to be comprehensive; the parameters of coverage and scope are intentionally narrow. Journal holdings,

for example, are limited, but users are assured that those journals are available in the library's collection. The journal list in the database was formulated by a Steering Committee of clinical and basic sciences faculty and reference librarians. The list was then reviewed by deans of the Schools of Medicine, Dentistry and Nursing to assure that key sources in their respective disciplines were not overlooked (Figure 1).

HOW TO USE THE MINIMEDLINE SYSTEM

The miniMEDLINE SYSTEM is similar to the PaperChase concept reported in 1981 [2]. How-

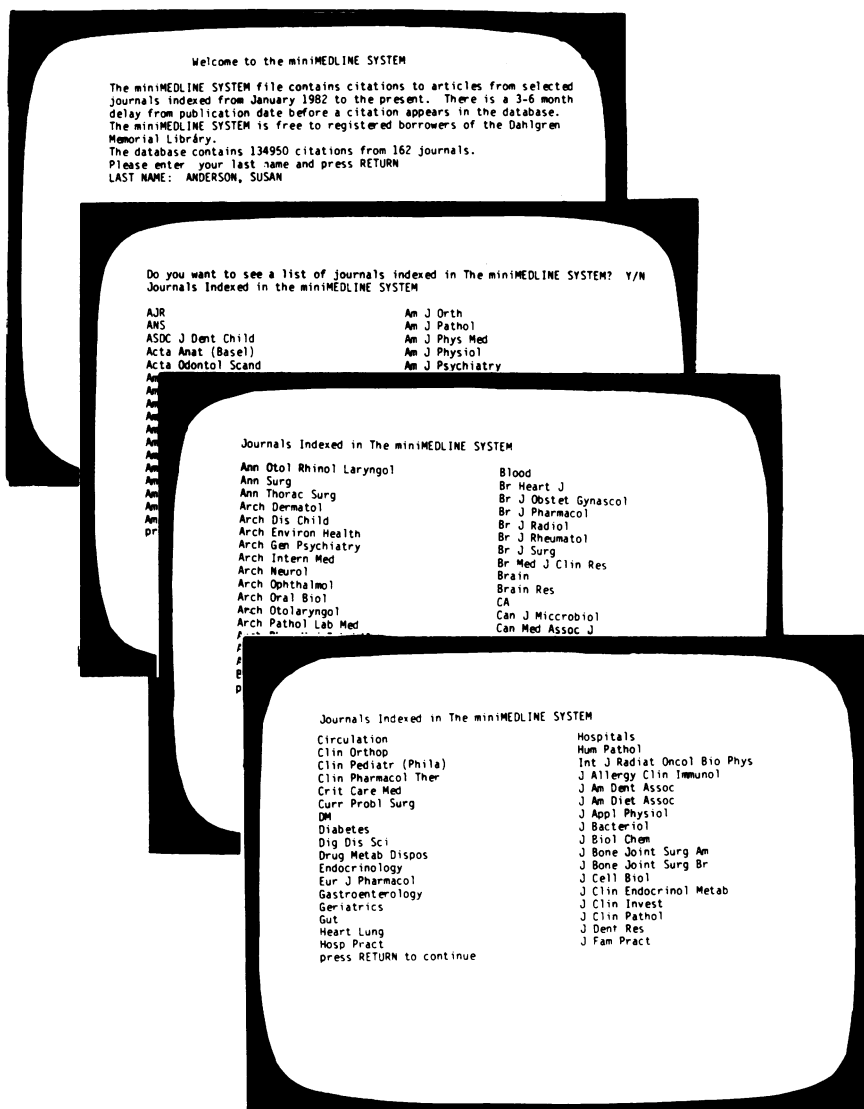


FIG. 1.—Sample of journal list available to users of the miniMEDLINE SYSTEM.

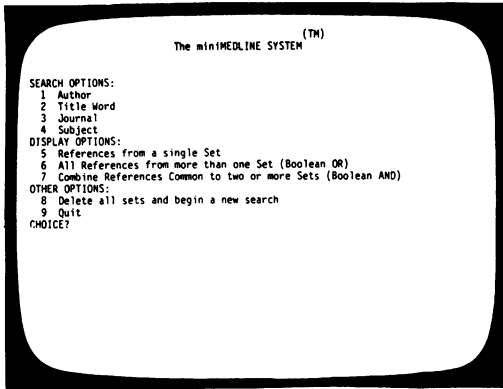


FIG. 2.—Menu used to generate searches.

ever, it has unique features and requires little typing and few keystrokes. It is menu-driven (Figure 2). The menu consists of nine options with only two search steps and one printing step, which is automatically generated for the user at the end of the search (Figure 3). In Step 1 the sets of citations can be built by searching any one of the four index fields, options 1–4: author, title word, journal, or subject. Truncated keyword searching can be carried out directly on authors and titles. The NLM MeSH vocabulary file and a journal title file serve as intermediary authority files to improve the user's efficiency in identifying subject headings and journal titles. In Step 2 sets can be displayed singly or combined, as shown in illustrated options

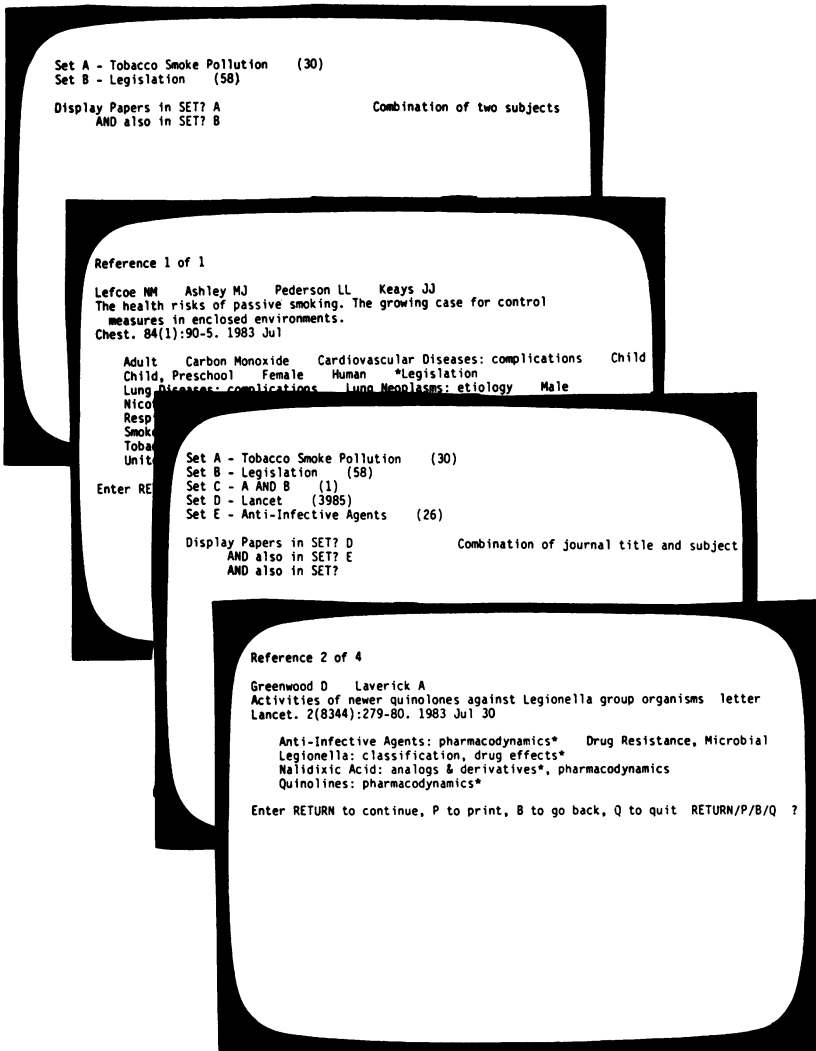


FIG. 3.—Two step approaches to developing sets for a bibliography.

THE miniMEDLINE SYSTEM

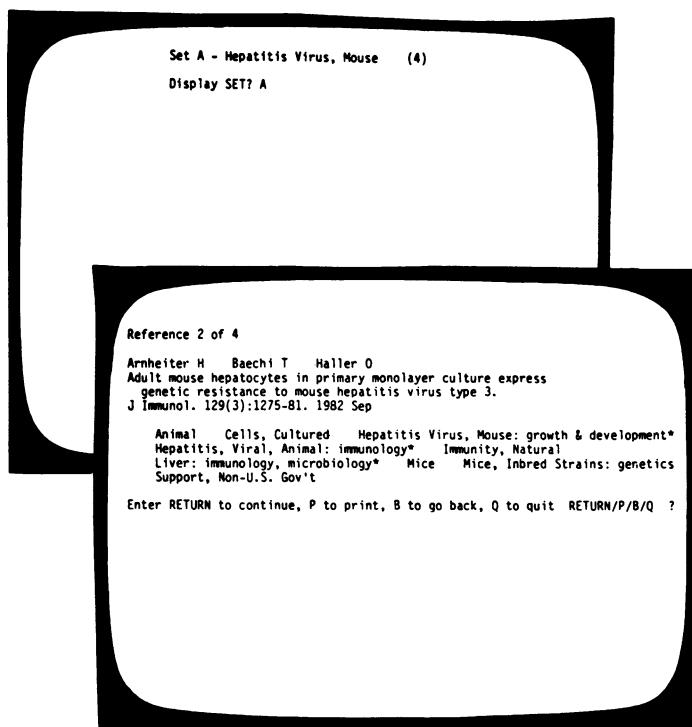


FIG. 4.—Users select from similar screen display generate a printed bibliography.

5, 6, and 7. Citations in a set are displayed one per screen and are saved for later printing or recall, as desired by the user. The display feature allows citation review and printing of selected citations in alphabetical order by journal title (Figure 4). The subject headings assist users in deciding whether to read the article or to search other topics as well. The last two options allow the user to begin a new search or to stop [3].

CREATION OF THE DATABASE

For creation of a miniMEDLINE SYSTEM database the library drew on its experience in compiling special bibliographies and the annual faculty publications bulletin of the medical center. That required formulating author searches to retrieve appropriate references and then transferring the data to the library's computer so they could be reformulated to produce the desired bibliographic publication. While technically similar, the miniMEDLINE SYSTEM was an immense and more complex project because the volume of records, updating steps, indexing procedures, and interactive requirements required a more sophisticated approach.

Options: Tapes or Data Capture?

There were two options for creation of the miniMEDLINE file: (1) the data could be loaded from the MEDLARS tapes that are available for purchase from the NLM through NTIS (National Technical Information Service), or (2) the data could be captured directly from a MEDLINE terminal in a fashion analogous to our method for special bibliographies or to that used by other libraries to capture cataloging records from an OCLC terminal by means of a printer port interface.

The library chose the data capture approach, better known as "downloading," for several reasons. The miniMEDLINE idea was originally an experimental pilot project; if it succeeded, it would be refined during the first year. The costs of purchasing the full MEDLARS tapes at \$50,000, or even \$20,000 for one year's tapes, seemed prohibitive in 1982. Because the required file was small compared to the entire MEDLARS file, this expense also seemed unreasonable for an in-house system. Although the NLM was considering approaches to meet these small-user demands, policies and operational procedures were not yet developed. (NLM announced availability of MED-

LARS subset tapes in October 1984. Georgetown plans to subscribe to this tape service and, therefore, to change its data-gathering approach.)

Besides the expense involved in acquiring the MEDLARS tapes in 1981, there were several technical problems with the tape approach that seemed difficult to overcome at the time. The tape format available from NTIS consisted of "the ELHILL unit record format, which is considerably more complex and more difficult to work with than even the MARC format because it contains two levels of indexing in binary format rather than in character format as in the MARC tapes" [4]. There was no easy way to extract a subset of references from the NLM tapes to meet the library's criteria for a simple journal title list and MeSH tree numbers with a related group of subject terms. Because of computer storage restrictions, the library wanted to use a standard MEDLINE search strategy and select only specific fields from each record (such as author, title, source journal, and subject headings) to develop the subset file. The tapes also required preprocessing on a mainframe computer for conversion to the library's minicomputer-compatible format. This further complicated the process and introduced additional costs.

Nevertheless, in 1982 Georgetown wanted to mount the system quickly so that the incoming fall class of students could use self-service searching and so that the library could begin testing the system. Although the data-capture approach had inherent error problems, these were considered minor compared to the costs and technical problems associated with the MEDLARS tape option. Since downloading was not costly, the library would be able to subsidize the project, providing free searching to Georgetown registered users. Consequently, the decision was made to proceed with the "downloading" approach.

Software Design Plans

Once the decision was made to launch the project, plans to design and program the system were implemented. Criteria were established for searching capabilities, and priority was given to developing a user-friendly approach to information access. It was decided to include truncated keyword searching for author and title words and (in a more specific manner) for subject headings and journal titles. The system was designed so that users could create sets of citations by searching any of the four index fields. These sets could then be combined through simple Boolean "and/or" operations for

retrieval of desired references. Other design features included display of selected citations, one per screen, using the MeSH terms as a pseudo abstract. Citations would be saved for later printing, with the printout reformatted and sorted in alphabetical order by journal title, because this is how journals are shelved in a typical scientific library [5].

High priority was given to creating a push-button approach that users could learn easily. Therefore, the end product, a one-point-of-entry design that doesn't require reentry of terms, eliminates problems for users who have difficulty typing. Simple instructions and hints are provided in the screen displays. Clear, uncluttered, easy-to-read screens are maintained throughout the system. With a few simple keystrokes users are able to search the system.

Data Capture Through Downloading

Design of the database using the downloading approach required utility programs with a number of subfunctions. The major routines included downloading, magtape, edit, and miscellaneous operations. These subfunctions are the first step in creation of the database. They are used to capture citations from a search conducted through the NLM's MEDLINE file and to transfer them immediately to the library's computer. Operator intervention is not required for continuous downloading. However, the system indicates if the downloading has stopped or if an error has occurred. Once checked for accuracy, the set of citations is generated. If errors occur in downloading or if citations need to be deleted, a number of routines are conducted. The final step integrates references with the master file and index, thereby making the references available for searching. Monitoring can take place while the process is in operation.

A number of subfunctions, such as reading references from a tape, replicating a tape (to guard against database degradation), writing update and master tapes, and indexing the entire master tape, are included in the system procedures. The software programs include statistical compilation for administrative reports and editing functions for changes in the miniMEDLINE journal list [6].

When the MEDLINE database is updated each month, those citations from the 160 selected journals are searched on the library's MEDLINE terminal in a subject search and are then transferred to the library's dedicated minicomputer through an interface cable connection from the MEDLINE terminal. Between 5,000 to 6,000 new references from the 160 journals are added each

THE miniMEDLINE SYSTEM

month to the miniMEDLINE database. Plans are to continue building the database, to add abstracts, and to keep at least three years' worth of references (or as many as disk storage will allow). The programs described above convert the database into a user-friendly mode that allows access through natural language. Index files are generated and each update is then merged into the existing miniMEDLINE database.

SYSTEM HARDWARE AND SOFTWARE

The miniMEDLINE SYSTEM operates on the library's PDP 11/44 minicomputer which has 2 megabytes of memory and three disk drives (which have two 254 megabytes each and a Winchester with 508 megabytes). The miniMEDLINE file currently requires approximately 80 megabytes of storage per year. Because the hardware configuration includes forty ports, the library has been able to dedicate six public access terminals to miniMEDLINE—three in the library and three at remote sites (in the University Hospital, in the School of Nursing, in a Basic Science department) that have direct access to the library's online catalog as well as miniMEDLINE. Lear Siegler ADM5 CRT terminals and OKIDATA Microline 82A printers are used by library patrons.

The system software is programmed in the ANS MUMPS language, using an Intersystems ISM-11 operating system. The nature of ANS MUMPS language allows programming structure with high speed that gives users direct access to references in the database.

COSTS OF MAINTAINING THE SYSTEM

Database development, equipment, computer storage, and online access are underwritten by the library. Because the major computer costs are absorbed by the library's internal automated system, it is difficult to separate some cost factors. Only the cost estimates on database creation and maintenance are included in Table 1 as a guide. These figures do not include the cost of the miniMEDLINE software applications programs, nor the cost of computer hardware required to replicate the system.

SYSTEM USE

Demonstrated broadly since January 1982 to librarians and physicians, miniMEDLINE became fully operational in August 1982 with the beginning of the academic year. It has been well received by students and faculty, gaining popularity almost

TABLE 1
COST OF DATABASE CREATION AND MAINTENANCE

Downloading Approach	Two-Year File	Current Year
Direct Cost for Data Capture	\$ 3,710	\$ 3,010
Disk Storage*	xxx	xxx
Supplies	600	300
Staff Time, approx. 10%*	5,200	2,600
Ongoing Programming	2,500	2,500
Total	\$12,010	\$ 8,410
NLM Subset Tapes Approach		
	Future Years	
Annual NLM Fee for Unlimited Use	\$10,000	
Ongoing Programming	2,500	
Preloading Tapes	600	
Disk Storage*	xxx	
Supplies	300	
Staff Time, approx. 5%*	1,200	
Total	\$14,600	

*Computer disk storage costs and staff salaries are dependent on an institution's financial structure.

overnight. Terminals available throughout the library are in high demand. Although miniMEDLINE is largely self-instructional, students and faculty taught by reference librarians quickly develop sound information-seeking skills and good comprehension of computer storage and retrieval capabilities such as Boolean logic.

The degree of sophistication exhibited in use of miniMEDLINE varies enormously. Analysis of data collected on number of searches, types of users, options selected, time spent, and searches printed reveals user behavior.

TABLE 2
STATISTICAL DATA

System Use	8/82- 6/83	7/83- 6/84	2-Yr. Total
1. Searches	11,274	16,897	28,171
2. Users (students, residents, faculty)	1,608	2,311	3,919
3. Time Spent (hours)	554	3,518	4,072
Total Time (minutes)	33,283	211,098	244,381
Average Time (minutes)	3	12	9
4. References Printed	47,207	99,662	146,869
5. Average References Displayed	25	42	34
Average References Printed	4	6	5

TABLE 3
ANALYSIS OF RETRIEVAL

The actual options used by patrons consist of:			
Search Options	8/82- 6/83	7/83- 6/84	2-Yr. Total
Author	499	1,070	1,569
Title	356	602	958
Journal	562	1,275	1,837
Subject	13,996	21,556	35,552
Display Options			
Option 5			
Single sets	13,332	21,045	34,377
Option 6			
All references— two or more sets (Boolean "or")	6,045	11,105	17,150
Option 7			
Combination of two or more sets (Boolean "and")	1,116	1,454	2,570

The statistical data show over 28,000 searches in two years, which represents almost five times the number of MEDLINE searches conducted by the librarians in that period. The computer connect time of 4,072 hours at \$22 per hour would have cost \$89,584 on MEDLINE. Because the miniMEDLINE SYSTEM runs at 9,600 baud, the average time per search is less than ten minutes. As noted in item 5 above, users are able to refine their own searches quickly and print only those references they need.

These data support the theory that novice users feel more comfortable working with single-subject searches, as noted in display option 5. A few more sophisticated users (31%) have learned to manipulate sets using the Boolean "or" as shown in option 6. And only 5% have learned to apply the more complex strategy shown in option 7 of Boolean "and" logic.

EDUCATIONAL FACTORS

There are some essential aspects of miniMEDLINE important to an educational setting. For students and faculty it is a painless means of acquiring computer literacy. The library also serves as an initial instructional base for users who want to search other end-user products available through vendors. Their curiosity about library holdings is stimulated as items appear on the CRT screen,

which subsequently increases their use of the literature. Despite the fact that minimal instruction is necessary, some professional guidance from librarians is preferred by more reluctant users. Educational sessions are provided both informally at the reference desk or formally in seminars at the library or at medical center departments. Lively communication exchanges take place, and the librarians are enjoying recognition from students and faculty for their knowledge of computers.

FUTURE PROSPECTS

Georgetown University agreed to replicate the LIS and the miniMEDLINE SYSTEM at other libraries in November 1982. Since then eight medical center libraries have contracted with the university to implement the miniMEDLINE SYSTEM: University of Texas at San Antonio, Texas Tech University, Johns Hopkins University, Cornell University, University of South Carolina, Rush University, Texas College of Osteopathic Medicine, and the Medical University of South Carolina. In less than two years, the miniMEDLINE SYSTEM has proved itself transportable and easily managed by librarians. Because the system is modular, a library can install miniMEDLINE as a part of its initial automation, and later implement an integrated system when funding is available.

The recent announcement from the NLM about making subset tapes available will allow Georgetown to further develop the miniMEDLINE SYSTEM. A number of enhancements are planned for the miniMEDLINE SYSTEM, such as adding abstracts and expanding the database. With conversion programs currently under development by Georgetown, user libraries will be able to refine their own database by adding or eliminating journals. As soon as the NLM tapes are available, user libraries will acquire them from the NLM or NTIS and with the Georgetown conversion software programs in place, they will begin to process locally the monthly updates. Other features planned for 1985/1986 will further enhance searching capabilities of the system. The availability of the NLM subset tapes introduces a positive dimension to miniMEDLINE that will be beneficial to end users.

The future looks bright for libraries and end-users benefitting from the emerging trend of user-friendly information products. The PaperChase project was the first known system to demonstrate the feasibility of developing a user-friendly product by converting MEDLINE tapes [7]. The miniMEDLINE SYSTEM at Georgetown University

also demonstrates a friendly end-user product, but it contains a smaller file acquired through a unique approach of downloading data. Georgetown plans to enhance and develop it further by using the NLM tapes for database creation.

A natural progression toward even more user-friendly systems is inevitable. In the future we will see new products and many changes. Today we are experiencing only the beginning of an era that holds great promise.

ACKNOWLEDGMENTS

I wish to thank Walter Bechtold, D.D.S., for his encouragement and insistence that this project was technically feasible. I wish also to acknowledge Alan Zuckerman, M.D., for developing the initial programs and to thank Helen Bagdoyan for assisting me in system design and for critiquing this manuscript. To those who have subsequently enhanced and maintained the system, Susan Anderson, Patricia Breault, Robert Larsen, and Bonnie Cannard, I owe my deepest gratitude.

REFERENCES

1. Georgetown University Medical Center. Library information system manual. Washington, D.C.: Georgetown University, Dahlgren Memorial Library, 1984:G1.
2. Horowitz GL, Bleich HL. PaperChase: a computer program to search the medical literature. *N Engl J Med* Oct 19, 1981;305:924-30.
3. Broering NC. The Georgetown University Library Information System (LIS): a minicomputer-based integrated library system. *Bull Med Libr Assoc* 1983 July;71:317-23.
4. Zuckerman AE. Bibliographic systems. In: Broering NC, et al., eds. *Computers and libraries: a management seminar*. Washington, D.C.: Georgetown University Medical Center, 1982:52-5.
5. *Ibid.*, p. 58.
6. Georgetown University Medical Center. Library information system manual, G.
7. Horowitz GL, Bleich HL. PaperChase.

Received August, 1984; Accepted November, 1984.