

Medical Information Methods

MARCUS ROSENBLUM

FEW GRADUATES appear to have more than a casual acquaintance with technical information sources and searching procedures . . . the United States has no shortage of information, but a real shortage of people who are talented and trained and motivated to make use of this information—WILLIAM T. KNOX, *May 28, 1959, Hearings before the Committee on Science and Astronautics, U.S. House of Representatives.*

An irony of contemporary science is that the business of producing information is out of proportion to the business of retrieving it. Science is like a salmon which deposits thousands of eggs with the prospect that a mere few may live. Dismayed by this uneconomical process, friends have suggested that I report a few elementary possibilities for saving time and energy in what Theodore Melnechuk calls "the 'helical' process of information exchange." (Melnechuk, director of communications of the Neurosciences Research Program, Massachusetts Institute of Technology, is also author of the appeal: "Please do not pollute the literature: It is the water you drink!")

At the risk of repeating information which may be embarrassingly familiar, the following notes touch briefly on the sources of information, the handling of information (literally, passing from one hand to another), and the packaging of information in the tradition not of Madison Avenue but of Madison, James, for

whom the newest addition to the Library of Congress is named.

This paper is addressed mainly to the needs of those who write or read medical literature, typically the teacher, student, or scholar, rather than the busy practitioner. Medical information methods needed by clinical practitioners have special criteria such as accessibility, immediacy, authority, brevity, specificity, and ubiquity. In particular, such criteria apply to methods of managing drug information. We are concerned here with the intellectual and theoretical organization of medical knowledge more than with the techniques of applying such knowledge.

Information Sources

The sources of information may be classed as primary and secondary. Direct experience is the primary source. In science, primary sources of information are sometimes divided into clinical experiments, studies in controlled laboratory environments, genetic studies, and statistical analyses of large numbers to overcome the variations which perplex the investigator (1). These four primary sources of scientific information are not exactly distinct or independent: in combination they may be thought of as composing a fifth category. Not to be excluded as factors in these studies are the basic biological experiences: DNA, imprinting, revelation, imagination, logic, intuition, insight, or instinct, all the unfathomable permutations of the information process.

The secondary sources of information are

Mr. Rosenblum is an adviser to the Public Health Service for handling scientific information.

divided into two categories which may be labeled "original" and "derivative." (It is a convention in library circles to speak of original materials as primary and of derivative materials as secondary publications. For present purposes, publications are regarded as secondary to the eye of the primary observer. Therefore, it seems advisable to use the terms "original" and "derivative.")

Original sources include journals, data books, notebooks, monographs, conference proceedings, letters, and talks or conversations.

Derivative sources are devised to save time for the student or investigator, to provide some perspective on the accumulated information, or to satisfy intuitive longings for system and order. They include manuals or handbooks, textbooks, reviews, abstracts, indexes, bibliographies, atlases, directories, and standard reference works. A descriptive catalog of these resources in the medical field alone would constitute a substantial document in itself, and would likely have to be revised annually. Considering the advantages to the student and investigator, such a catalog might be worthwhile especially in circumstances where the services of a sophisticated librarian will not substitute.

Gathering Methods

Contrary to expectations, the usual impulse of a person in search of knowledge is to ask not a librarian but a friend. At the specialized nose cone of scientific exploration, the scientist's colleague is likely to have more suggestions than a librarian. (In highly specialized fields, the invisible colleges or information-exchange groups provide prompt access to the latest findings by those in the field of inquiry.) On the other hand, the modern library system offers vast opportunities for inquiry by students and generalists. Among the library resources are the following.

The National Referral Center for Science and Technology, Library of Congress, maintains an inventory of specialized information centers, in and out of Government (2).

The Science Information Exchange, Smithsonian Institution, maintains an inventory of research in progress (3).

The Clearinghouse for Federal Scientific and Technical Information, U.S. Department of

Commerce, catalogs, announces, and sells unclassified technical documents issued by Government agencies and contractors, including medical translations (4).

The National Library of Medicine can search, find, and print extensive bibliographies of the medical literature in less than 2 weeks from receipt of the request to delivery of the product. Since 1964, the Library has used magnetic tapes and a computer to record, store, search, and retrieve titles of medical reports. These titles are recovered by the system called MEDLARS (Medical Literature Analysis and Retrieval System) and they are set in type by the system called GRACE, which uses a photoelectric process (5), for *Index Medicus* (discussed later in this section).

The National Clearinghouse for Mental Health Information is building a comprehensive collection of literature in its field (6).

In addition to these national resources, many institutions and associations, reference centers, clearinghouses, research centers, and educational units may answer specialized questions.

Among the growing number of special information centers of importance to the health professions, some of the better known are operated by the Public Health Service, and others are financed by Public Health Service grants. They include the National Clearinghouse for Poison Control Centers, the Cancer Chemotherapy Service Center, and the National Center for Health Statistics.

The number of information centers varies according to the definition. Russell Shank defines an information center broadly as "any agency, publication, or other unit, large or small, which gathers data about a topic, wide or narrow, with the express purpose of serving it in various formats and according to variously worded demands to any interested party, professional or lay, who can qualify, with money, interest, or understanding or otherwise, to use the service."

The Department of Defense definition reads, "Any functional element is performing as an information analysis center if it collects, reviews, digests, analyzes, appraises, summarizes and provides advisory and other user services concerning the available scientific and technical information and data in a well-defined, special-



Computer room at the National Library of Medicine

ized field. A center exclusively concerned with review or analysis of scientific or engineering data shall be considered an information analysis center. Such centers are distinguished from documentation centers and libraries, whose functions are primarily concerned with the handling of documents rather than the technical information contained in the documents.”

One of the virtues of the medical information system as it stands today is that almost any port of entry may connect with the resources of the system as a whole. Though the present system is voluntary and largely informal, rather than monolithic, it is unusually cooperative. For example, through an interlibrary loan system, medical libraries are capable of locating a book anywhere in the United States and either borrowing a copy or obtaining a print of specified pages. The library locates the work through a union catalog, if the publication is not contained in the library's own catalog of accessions. One may ask a library for a specific work, for a bibliography on a specific subject, or for suggestions about how to look for answers to a given question. The library

in turn may request the National Library of Medicine to print out an appropriate list of citations from recent periodical literature. (The National Library of Medicine has been working successfully on developing regional centers to use the MEDLARS tapes. These tapes will also be used abroad.) Any reader may use the library system by consulting the library catalog, by specific inquiries at the reference desk, by drawing upon the loan system, or by obtaining a bibliography as a lead to further research.

Before a reader begins a thorough search of the literature, he usually builds up a background by looking at selected works. At this stage a talent for skimming, rather than scanning, is a prime asset. This talent is assisted by authors and editors who provide meaningful titles, informative chapter headings and subtitles in a table of contents, abstracts or summaries, accurate citations or bibliographies, and a reasonably detailed and careful index.

For elementary inquiries, the index for a single book may point to the specific passages that satisfy the reader. For a quick survey of publications on a given topic, *Index Medicus* con-

tains the most comprehensive list of literature citations available in English (7). A hand search of these volumes may help to refine questions which would be appropriate to a bibliographic search by the National Library of Medicine, or it may turn up enough citations to make further search unnecessary. In any event, *Index Medicus* is indispensable to a search of medical literature.

A second important aid published by the Public Health Service is the Research Grants Index (8). This index is not only an indicator of current research; it also contains lists of publications of investigators associated with specified subjects.

Access to specialized literature is offered also by such indexes as the Hospital Literature Index, published by the American Hospital Association; the Index to Dental Literature, published by the American Dental Association; and the Industrial Hygiene Index, published by the Mellon Institute, Industrial Hygiene Foundation.

A citation index may be available to some readers. This lists the works which have cited a specific author. For example, under Fox, Theodore, "Crisis in Communication," the index might list every subsequent document which cited this work. The evolution of an idea may be followed in either direction by means of citations.

Another resource for spotting literature is *Current Contents*, a weekly which reprints only the table of contents of each of a variety of scientific journals (9). This publication announces new papers as rapidly as they become available.

Finally, new information is heralded by announcements issued by various publishing units and libraries. (Thanks to its computer, the National Library of Medicine today is able to publish new catalog cards and its *Current Catalog* of new accessions every 2 weeks. The accumulated titles are published annually.)

For those who screen a number of publications by title and yet find that the titles are not sufficiently informative, even if annotated, a variety of abstract services are available. Usually abstract services are so expensive that small libraries cannot afford them, or they are so highly specialized that they serve few readers.

Abstracts of publications of interest to medical readers are likely to appear in *Biological Abstracts* (10), *Chemical Abstracts* (11), or *Excerpta Medica* (12). For convenience, however, the librarian may be able to direct certain readers to such specialized services as those which abstract literature on arthritis (13), cancer chemotherapy (14), or gastroenterology (15).

An especially economical entrance to a subject, if available, is a review. The National Library of Medicine periodically publishes a Bibliography of Medical Reviews (16). Reviews are separately listed also in the *Index Medicus* monthly. As well as listing in *Index Medicus*, the bibliography is scheduled to be published monthly as a separate periodical.

Generally, reviews are either critical or historical, sometimes both. The critical reviewer may be inclined to settle for a less comprehensive treatment than the historical reviewer. Some of the latter, on the other hand, may content themselves with little more than an annotated bibliography. Ideally, a review should be open-ended so that the state of the art may be brought up to date annually by a more or less invisible college dedicated to the field. Textbooks perform this service in a limited fashion.

Information Handling

The classic trinity of information handling wears the labels of "storage," "search," and "retrieval."

The brain is not always the best place to store information; frequently information belongs in a box, on a shelf, in a desk, in files, or on wall charts. Space for the storage of information has become so much a social challenge today that it has called up new technology. One dramatic response has been the microform. Photography reduces the size of a printed page by as much as 300 to 1, and still the image is sharp enough to be magnified to readable size where it may also be printed afresh. Such extreme reductions are rarely necessary: a reduction of 18 to 1 on a microfiche, a transparent sheet, has been approved as a desirable standard for most Government documents. A 60-page document can be printed on a standard single sheet, about 4 inches by 6 inches. Trailer

sheets can be used for longer reports. Since 1964, all documents in the collection of the Clearinghouse for Federal Scientific and Technical Information have been reproduced on microfiche and are available in that form (17).

Other forms which assist in the compact storage of information are the familiar notebook, punchcards, and tapes. Where a computer is available, tapes or electromagnetic disks are capable of prodigious feats of storing information under conditions which permit prompt and accurate retrieval without loss, distortion, or deterioration. For small collections of literature, some students find it convenient to fix abstracts on 3- by 5-inch cards. A few journals print abstracts in a size appropriate to a card file.

The main objective of storage is not merely to preserve information or to reduce the space it occupies. Primarily, the intention is to facilitate search and retrieval. Search implies not merely the recovery of a specific piece of information, identified beforehand; sometimes it implies also the capacity to uncover previously unsuspected information. For this reason, the system, as well as the facility, of storage is sometimes designed deliberately to provide many cross-references or to place in juxtaposition a variety of loosely associated documents. When the objective is rapid retrieval of a specific item, of course, the system, such as a numerical file of insurance policies, will be tight.

As a guide to retrieval of information, various indexing methods apply (18,19). For a small collection of books, it may be convenient to arrange them by size or color, and trust memory to locate "the long thin book" or the "yellow book." For larger collections, the usual tendency is to set up a system of classes, usually in a hierarchy: *eyes* would be classed under *face*, and *retina* under *eyes*.

Such hierarchies may be built according to the order of a discipline or according to a mission. A disciplinary structure for microbial studies, for example, might be subdivided into parasites, bacteria, rickettsia, and viruses, with further subdivisions in each class. A mission-oriented hierarchy for control of tuberculosis might arrange its file according to the neighborhoods or families under surveillance, with

subdivisions according to location or family contacts.

In contrast to the hierarchical plan, another philosophy of indexing is that each term is unique in itself. This system uses key words to identify each document. If the reader is unable to think of a necessary key word, cross-references may clue him in. Computers can be programed to substitute the key word for an associated word that is presented: for example, *cavity* may be referred to *caries*, or vice versa. The order of the alphabet or a numerical code for each document then finds the items sought. Combining the unit plan and the tree plan, the National Library of Medicine's Medical Subject Headings (MeSH) provides standard terms for searching the biomedical literature.

Conceivably, *caries* can cue in more documents than the reader wants. He then adds terms to reduce the number eligible: *caries*, *lower*, *left*, *bicuspid*, *female*, *Sudanese*, *underweight*, and so on. Such refinement implies a depth of indexing which permits the use of as many as 30 terms to an item. Large collections can be indexed to that depth and selected by modern technology. Small collections are indexed with tab cards.

Intermediate collections often use a so-called "peekaboo" technique, which identifies documents associated with key terms by holes punched in index cards. Various other ingenious techniques have been marketed to facilitate search and retrieval.

(In the marketing process, there is a tendency for the salesman to refer to the technique as a "system." The technical "system" is not to be confused with the intellectual system of organizing the information. The technical device is the servant of the intelligence which organizes the information: it cannot organize the information for the user any more than a 3- by 5-inch card can assign an index term.)

Distributing Information

Anyone who is linked to an information system of moderate size may find that much of the work of search and retrieval is performed for him automatically. Many centers conduct a system of selective dissemination of information which keeps a file of the reader's interests, referred to as a "profile," so that documents

which concern the reader are identified as they arrive. The next step may be simply to notify the reader of the title, or perhaps to send an abstract, before he orders the complete document. Or the complete document may be sent out at once. In any event, the responses of the reader are fed into the system with the result that he subsequently receives more (or less) of the particular kinds of documents that he commented on.

On a much less elaborate scale of information management, cooperative groups of readers arrange to share responsibility for reviewing a selected body of publications: each screens the papers in the journals assigned to him and notifies the others of any items which may be of special interest to them individually or to the group as a whole.

The comparative costs of various methods of retrieving and distributing information are a decisive factor in operating such services. Of equal importance is the ability to meet the needs of the user. For example, a user may prefer to receive information rather than a document.

The distinction is that a document may oblige the user to form judgments he is not qualified to make, when it should be possible for him to obtain reliable conclusions from eminent authorities. Or the user may simply wish to save time. For example, if he wishes to know the approved Federal standard size for microfiche, he will be satisfied to learn that it is 105 mm. x 148 mm. (about 4 by 6 inches), rather than scan a document to find this detail.

At the point where there is a choice between information and a document, the helical process of information handling executes a full turn. The consumer of information also operates as a generator of information, and his thoughts turn to the shape of the package as well as to the contents.

Information Packaging

Fortunate is he who has access to competent editorial services for the packaging of information. Few institutions offer literary assistance to authors whose verbal skills, no matter how strong, may be less important than their scholarly contributions. Such a service impinges on the originality of the author no more than the audiovisual and translation services. Literary

and editorial services are technical aids: they do not share professional responsibility for the final message. Nevertheless, many authors are reluctant to use them when they are available and even less likely to demand them in their absence.

It may not be germane at this point to mention the special obligations or talents of the medical writer. The attitude and the psychology of the writer (20) may be more critical to his performance than the amount of technical advice applied. In any event, the reader who aims to write will learn more from the excellent works on this subject (21-23) than from these few remarks.

For those aiming at journal publication, there is no "Blue Book" to guide them in selecting an audience. The selection is determined by the character of the report if it is appropriate to a specific professional society, but the interdisciplinary nature of many reports today often creates uncertainty as to the best medium. For the time being, writers must live with this uncertainty.

Fox (24) has suggested that journals have two functions and should be separated and identified accordingly: it is their duty to record or to stimulate. Others doubt that journals still serve effectively as a medium of exchange (25) and assert that 90 percent of them consist of "background noise." Still others propose that Government writings at least ought to be issued first as administrative documents, to be announced, cataloged, and indexed, and released for journal publication only if the primary readers feel they deserve a wider audience.

Modern methods of reproducing small quantities of a report directly from the original typescript have made distribution preliminary to formal publication both practical and economical.

Following publication the feedback begins, provided the selection of an audience has been effective. Both the marketing and feedback stages are beyond the scope of this essay, as are certain other critical elements in the information cycle, such as the art of listening.

Perhaps this description of information methods seems more like a model system than a de facto system. In actuality, an author may be an egocentric who confines his sources to his

own data and publishes as he pleases where he pleases without regard to the pros and cons, and still becomes a Nobel prize winner. We have yet to acquire case studies that tell how to succeed in handling information. Meanwhile, some may wish to experiment with the facilities noted above (26).

Summary

Improved library resources available may substantially ease the task of acquiring information for scientific exchange. Management of the store of information acquired varies according to the size of the collection and the mission of the collector. A variety of inventions has become available to improve such management. In the final process of packaging information, the chief deficiency appears to be lack of technical assistance (or indifference to the use of technical aids). New methods of publication may be required to accommodate the growing volume of scientific information.

REFERENCES

- (1) McKeown, T.: Social medicine. British encyclopedia of medical practice. Butterworth & Co. (Publishers), Ltd., London, 1960, pp. 267-282.
- (2) National Referral Center for Science and Technology: A directory of information resource in the United States. Library of Congress, Washington, D.C., 1965.
- (3) Smithsonian Institution: Science information exchange. Washington, D.C., 1965.
- (4) United States Department of Commerce: Clearinghouse for Federal scientific and technical information. Springfield, Va., undated.
- (5) U.S. Public Health Service: The MEDLARS story. National Library of Medicine, Bethesda, Md., 1963.
- (6) U.S. Public Health Service: National Institute of Mental Health. PHS Publication No. 20. U.S. Government Printing Office, Washington, D.C., 1960, revised.
- (7) U.S. Public Health Service: Index medicus. National Library of Medicine, Washington, D.C., monthly, cumulated annually.
- (8) U.S. Public Health Service: Research grants index. Division of Research Grants, Bethesda, Md., two volumes, annually.
- (9) Institute for Scientific Information: Current contents. Philadelphia, Pa., weekly.
- (10) Biological Abstracts. Philadelphia, Pa., semi-monthly. Also B.A.S.I.C. a keyword index to Biological Abstracts, semimonthly, and Bioresearch Titles, a monthly index to documents not covered by Biological Abstracts.
- (11) American Chemical Society: Chemical abstracts. Chemical Abstracts Service, Ohio State University, Columbus.
- (12) Excerpta Medica Foundation: Excerpta medica (abstracting publications in 23 categories, usually monthly and annually, covering more than 2,700 journals), Amsterdam.
- (13) U.S. Public Health Service: Arthritis and rheumatic diseases abstracts. National Institute of Arthritis and Metabolic Diseases, Bethesda, Md., monthly.
- (14) U.S. Public Health Service: Cancer chemotherapy abstracts. National Cancer Institute, Bethesda, Md., bimonthly.
- (15) U.S. Public Health Service: Gastroenterology abstracts and citations. National Institute of Arthritis and Metabolic Diseases, Bethesda, Md., serial.
- (16) U.S. Public Health Service: Bibliography of medical reviews. National Library of Medicine, Bethesda, Md., 1966.
- (17) Executive Office of the President: Federal microfiche standards. Clearinghouse for Federal Scientific and Technical Information, Office of Science and Technology, Springfield, Va., 1965.
- (18) International Business Machines: Reference manual: Index organization for information retrieval. Data Processing Division, White Plains, N.Y., 1962.
- (19) Shank, R.: New concepts in indexing. Bull Med Lib Assoc 53: 385-396, July 1965.
- (20) Kubie, L. S.: Blocks to creativity. Int Sci Techn No. 42: 69-85, June 1965.
- (21) Trelease, S. P.: How to write scientific and technical papers. The Williams & Wilkins Company, Baltimore, Md., 1958.
- (22) Strunk, W., Jr., and White, E. B.: The elements of style. The Macmillan Company, New York, 1959.
- (23) Peterson, M. S.: Scientific thinking and scientific writing. Reinhold Publishing Corporation, New York, 1961.
- (24) Fox, T.: Crisis in communication. Athlone Press, London, 1965.
- (25) Price, D. J. deS.: Networks of scientific papers. Science 149: 510-515, July 1965.
- (26) Fleming, T. P.: A guide to the medical literature. Columbia University Press, New York, 1961.