JAMIA
The Practice of Informatics

Technical Milestone ■

Medical Subject Headings Used to Search the Biomedical Literature

MARGARET H. COLETTI, MLS, HOWARD L. BLEICH, MD

Abstract The National Library of Medicine's MEDLINE (MEDLARS Online) database was the first database to be searched nationwide via value-added telecommunication networks. Now available on the World Wide Web free of charge from the National Library of Medicine and from many other sources, it is the world's most heavily used medical database. MEDLINE is unique in that each reference to the medical literature is indexed under a controlled vocabulary called Medical Subject Headings (MeSH). These headings are the keys that unlock the medical literature. MeSH multiplies the usefulness of the MEDLINE database and makes it possible to search the medical literature as we do today. This paper commemorates the 40th anniversary of the introduction of MeSH and salutes some of the farsighted persons who conceived and developed the MEDLINE database.

■ **J Am Med Inform Assoc.** 2001;8:317–323.

> The impact of MEDLARS on the medical library world is not that of the familiar metaphor—the pebble dropped into the pond, casting concentric circles that reach many points on the shore. The impact is no pebble for sure. It is a mighty rock. The waves it will cause will surge and splash for a long time to come. MEDLARS is the great bibliographical breakthrough of our generation.—RALPH T. ESTERQUEST

Affiliations of the authors: Beth Israel Deaconess Medical Center, Boston, Massachusetts (MHC, HLB); Harvard Medical School, Boston (HLB).

Correspondence and reprints: Howard L. Bleich, MD, Beth Israel Deaconess Medical Center, 330 Brookline Avenue, Boston, MA 02215; e-mail: <bleich@caregroup.harvard.edu>.

Received for publication: 8/23/00; accepted for publication: 3/16/01.

Spoken by Ralph T. Esterquest at a symposium in 1963,[1] these words proved prescient as well as poignant. The National Library of Medicine's MEDLINE (MEDLARS Online) database was the first database to be searched nationwide via value-added telecommunication networks. Now available on the World Wide Web free of charge from the National Library of Medicine and from many other sources, it is the world's most heavily used medical database.

MEDLINE is unique in that each reference to the medical literature is indexed under a controlled vocabulary called Medical Subject Headings (MeSH). These headings are the keys that unlock the medical literature. MeSH multiplies the usefulness of the MEDLINE database and makes it possible to search the medical

literature as we do today. This paper commemorates the 40th anniversary of the introduction of MeSH, and it salutes some of the farsighted individuals who conceived and developed the MEDLINE database.

John Shaw Billings and the Origin of Index Medicus

At the dawn of the Civil War, a student at the Medical College of Ohio needed information about epilepsy for his graduation thesis. The student, John Shaw Billings, spent six months in libraries in Cincinnati, New York, and Philadelphia gathering the necessary literature. Medical literature was then fragmented into thousands of lists, bibliographies, and indexes; to search the medical literature was drudgery that Billings would not forget.[2]

After graduation, Billings distinguished himself as a field surgeon in the Union army. When the war was over, he got a job as an assistant in the Surgeon General's office, where he took charge of a collection of books that originated on a shelf behind the desk of Joseph Lovell, a former Surgeon General. Lovell had died in 1836, but his collection continued to grow, and when Billings arrived it numbered about 2,000 volumes. In 1866, the Surgeon General's Office and the collection moved to space in Ford's Theatre,[3] in Washington, the site of President Lincoln's assassination.

Billings built up the collection, and by 1876 it had 52,000 books and pamphlets. In time, it would become the largest medical library in the country.[4] Recalling his earlier experience in his college library, Billings decided to index the collection by author and subject. In 1874, he began the work of indexing by subject the journal articles, books, theses, reports, and pamphlets, to produce the comprehensive *Index-Catalogue of the Library of the Surgeon-General*. The *Index-Catalogue* would eventually be published in several series of revolving alphabetic volumes; new entries for anatomy received just after the release of "A" had a long publication delay.

Accordingly, starting in 1879, Billings published the new items from medical journals as a separate, monthly current awareness service called *Index Medicus*. In those days, all issues of *Index Medicus* for an entire year, including both author and subject lists, easily fit into a single bound volume; they require 16 bound volumes today. The idea that each article in the medical literature should be indexed under the subjects discussed in the article would lead to the development of the MeSH that we use today.

After Billings retired from the United States Army in 1895, he became director of the New York Public Library, where he served with distinction for several years. The directorship of the Library of the Surgeon General's Office and the responsibility for *Index Medicus* were passed from one senior medical officer to another in a succession of four-year Army posts. In 1927, *Index Medicus* was merged with the American Medical Association's competing bibliography and renamed *Quarterly Cumulative Index Medicus*.

World War II brought urgency for current medical information, and a new director, Harold Wellington Jones, started publishing a weekly index called *Current List of Medical Literature*. Jones also arranged for a 1943 survey of the Library by professional librarians. The highly critical report of this survey[5] would serve as a blueprint for major improvements to the Library's collection, staffing, operations, and facilities over the next two decades. In response to one of its recommendations, Jones developed "Principles to be Used in the Selection of Subject Headings," based on guidelines originally formulated by Billings.[6]

Frank Bradway Rogers and Medical Subject Headings

In 1949, Colonel Frank Bradway "Brad" Rogers became the director of the Army Medical Library—then the name of what had been the Library of the Office of the Surgeon General of the Army. Like each of his predecessors, Rogers was a physician. Unlike them, he was sent by the Army to obtain a master's degree in library science as preparation for the job.

The first director to be formally trained in librarianship, Rogers used the principles developed by Jones and work done under contract at the Welch Medical Library of Johns Hopkins University to produce, in 1951, a standardized list of subject headings for the *Current List of Medical Literature*.[7] In 1956, the U.S. Congress bestowed on the Armed Forces Medical Library (as it had been renamed in 1952) statutory authority as the National Library of Medicine, and made it a separate institution within the United States Public Health Service. Later, the National Library of Medicine would become part of the National Institutes of Health.

In 1960, Rogers guided the publication of the newly revived monthly *Index Medicus* along with a freshly revised and expanded list of standardized subject headings. The list—called Medical Subject Headings, or MeSH—consists of single- and multi-word terms

that are used to index and catalog the medical literature. This paper commemorates the 40th anniversary of MeSH—the controlled vocabulary, or thesaurus, that makes it possible to search the biomedical literature in the way we do today.

Dedicated to widespread, timely dissemination of medical information, Rogers investigated mechanical ways to facilitate the production of the monthly *Index Medicus*. In 1958, he and Seymour Taine, then head of National Library of Medicine's indexing operation, initiated the Index Mechanization Project, which used punched paper tape and punched cards, card-sorting machines, and a high-speed camera. This project speeded production of the printed *Index Medicus*, but it did not support automated retrieval.

This early experience with automated techniques, and Rogers' interest in the concept of coordinate indexing developed by Mortimer Taube,[8] influenced the structure and philosophy of the first edition of MeSH.

> Subject cataloging and periodical indexing are, in their major dimensions, identical processes, and a single authority should be used for both. Headings should be considered as pointers, not descriptive labels. Topical subheadings, as substitutes for phrase headings, can be standardized, and should be, in contrast to specific main headings that generally represent broad concepts.[9]

Main headings became more specific in subsequent editions, but MeSH continues to be used both in subject cataloging and indexing (a novel notion in 1960), and standardized topical subheadings remain a key feature of the MeSH vocabulary.

Each year, *Index Medicus* gets bigger. In 1960, the annual cumulation occupied three volumes. In 1999, the author entries filled six volumes and the subject entries another ten. To look up articles indexed under a particular topic over a span of years requires consulting many hefty volumes.

Index Medicus provides a way, albeit an awkward way, to search for articles written by a particular author or articles that pertain to a particular subject, but the bound volumes provide no practical way to find articles that discuss two (or more) topics together; one can find articles about timolol or articles about myocardial infarction, but there is no easy way to focus a search on articles that discuss both. Also, under each MeSH term, the published volumes of *Index Medicus* list only those articles for which that term has been designated a Major Focus—one of the (usually two to four) principal topics discussed in the article; to list the remaining articles would render the printed volumes too thick and too expensive.

MEDLARS

In 1960, with experience from the Index Mechanization Project and with funds from the National Heart Institute, the National Library of Medicine began development of the Medical Literature Analysis and Retrieval System (MEDLARS). As part of this effort, a MeSH database was developed, and a revised and restructured MeSH was produced in 1963 under the direction of Winifred Sewell.[10] MeSH has been revised and updated annually ever since. The 1963 edition contained the first version of the MeSH hierarchies, or "Tree Structures"—soon to become a powerful aid to automated retrieval. From the beginning, MeSH tree structures have been polyhierarchic, meaning that a main heading can appear in more than one subcategory. For example, **Tuberculosis, Hepatic** appears in the Infectious Diseases tree under **Tuberculosis** as well as in the Digestive System Diseases tree under **Liver Diseases**.

In 1964, the Library introduced MEDLARS—which stored bibliographic references in a computerized, searchable database and performed computerized phototypesetting for *Index Medicus* and other, more specialized bibliographies. Beginning in 1965, searches could be submitted to the National Library of Medicine or to one of the decentralized processing centers that were established in the United States and overseas. Specially trained librarians, who had attended courses that lasted as long as 3 months, then formulated each search and submitted it to a MEDLARS Search Center, where punched cards were fed into a computer, and the resulting printout was shipped back by parcel post. In the United States, turnaround time averaged 4 to 6 weeks.[9]

Searching offline in batch mode delayed gratification; not infrequently, the results of a search came back after the clinical question was no longer relevant. In addition, it was hard to anticipate how many hits a particular search would yield. To help circumvent this problem, MEDLARS allowed three search statements, with each statement processed against the previous ones. Thus, the experienced searcher could aim for recall with the first statement, for specificity with the third, and for an acceptable balance with the second. Still, to expand a search that yielded too few hits or to focus one that yielded too many required resubmission and further delay; searching needed to be done online.

In December 1968, funded by the National Library of Medicine, the first real-time, online bibliographic retrieval system was inaugurated at the SUNY

Biomedical Communication Network, headquartered at the SUNY Upstate Medical Center Library in Syracuse, New York.[9] Nine medical libraries used teletypewriter terminals to search 90,000 references taken from the MEDLARS database. Searches could be done for authors' names or medical subject headings, and they could be limited by date or language; there was no provision for searching text words in the title or abstract. Searches were submitted online, but until the results were printed offline and mailed back, the searcher did not know how many hits the search would produce, or whether they were on target. Still, the nine libraries did more searches via SUNY than the rest of the country did via MEDLARS.[10]

In 1970, the National Library of Medicine, which received requests for interlibrary loan via Teletypewriter Exchange, or TWX, experimented with nationwide online access to about 100 journals indexed for the *Abridged Index Medicus*, or AIM.[11] Short-lived but popular, the project, called AIM-TWX, provided further evidence that people wanted to search online.

MEDLINE

In 1971, the National Library of Medicine introduced MEDLARS Online, or MEDLINE, on more advanced retrieval software, and made it available via value-added telecommunication networks. Named ELHILL, this software permitted trained searchers—primarily librarians—to search by words in the title as well as by fixed fields, such as medical subject heading, journal title, year of publication, and the like. ELHILL was named for (Joseph) Lister Hill, the senator from Alabama and strong advocate of health care for whom the Lister Hill National Center for Biomedical Communications, which he helped create, was also named.

Before they were allowed to use the ELHILL software to search the MEDLINE database, searchers were required to take a two-week course in online searching and in the use of MeSH. Research workers and clinicians discussed their need for information with one of these specially trained searchers, who translated the request into the appropriate search strategy and conducted the search. This was known as "mediated searching"; with rare exceptions, "end-user searching" did not exist.

In the early days of MEDLINE, librarians typically took time to prepare each search before going online. Connect time was too expensive to waste, and the initial system required the searcher to type each term in an exactly correct format. Printed volumes, such as the Permuted MeSH, were used to guide the searcher to unfamiliar MeSH terms. Searchers had to know (or look up) which MeSH terms were singular (**Eye, Hand,** and **Foot**), and which were plural (**Fingers** and **Toes**). They had to know when the noun preceded the adjective (**Diet, Reducing**) and when the adjective preceded the noun (**Diabetic Diet**), as well as when to apply a subheading to a main heading (**Hypertension/Diet Therapy**) and when to use the Boolean AND with a separate medical subject heading (**Hypertension** AND **Diabetic Diet**). They had to use specific abbreviations for journal names (New Engl J Med for *The New England Journal of Medicine*, but JAMA for the *Journal of the American Medical Association*), and they had to know that authors' names were stored in the form Smith JA.

As text word searching and more MeSH cross-references became available online, and as connect time became less expensive, searchers began to use immediate feedback from the system to prepare and refine their searches. The searcher who was unsure of a MeSH term could, instead of looking it up in a hefty book, search by title or abstract words, find a few on-target references, and see how they were indexed. Instead of the searcher's spending time to prepare the search in a form that would be acceptable to the computer—instead of the searcher's working for the computer—the computer began to work for the searcher.

Production of the MEDLINE Database

Journals selected by the National Library of Medicine for indexing are sent to a team of specially trained indexers who analyze each article and assign medical subject headings to it. They assign an average of eight to ten headings to each article. As previously mentioned, an average of two to four of these are designated the "major focus" of the article. These headings are the principal subjects discussed in the article and are the headings under which the article appears in the printed *Index Medicus*. The remaining headings refer to topics that are discussed but are not the main points.

Most headings can also be assigned subheadings; in this way, an article that discusses the side effects of Aldomet can be indexed under the main heading **Methyldopa** with the subheading **Adverse Effects**. The indexer also examines each article for other types of descriptors, such as publication types (**Historical Article, Review, Letter to the Editor, Guideline,** etc.) and check tags (**Human, Animal, Male, Female**, etc.).

The indexer uses the most specific term available in the hierarchic MeSH tree structure. Thus, an article

that discusses anorexia and bulimia, but no other type of eating disorder, will be indexed under **Anorexia** and under **Bulimia**, but not under **Eating Disorders, Mental Disorders,** or **Psychiatry**. Because MeSH is organized hierarchically, software allows searches that are limited to articles indexed under the general term **Mental Disorders** as well as searches that "explode" **Mental Disorders** to include all articles indexed with any of the more specific MeSH terms that appear under it in the MeSH hierarchy. Under **Mental Disorders** we find **Anxiety Disorders, Personality Disorders, Eating Disorders,** and the like, and under **Eating Disorders** we find **Bulimia, Anorexia, Hyperphagia,** and the like. An explosion of **Mental Disorders** would retrieve references indexed under any of these topics and more. In some sections of MeSH each child term is a subset of its parent, whereas in other sections groupings are more pragmatic than logical—that is, useful for retrieval when a parent term is "exploded."

Originally, the practice of indexing under the most specific term possible served to limit the number of entries listed under any one term in the printed volumes of *Index Medicus*.[11] More recently, it has made it possible to limit a search to a specific area of medical practice or research. For example, in 1988, the term **HIV** was added to the MeSH vocabulary. One year later, to accommodate research into subspecies of the HIV virus, the terms **HIV1** and **HIV2** were added. The addition of these more specific terms allows the clinician or researcher to further focus a search.

During the past 40 years, the number of MeSH main headings has grown from 4,400 to close to 20,000. The cross-reference structure has also been greatly expanded. Today, there are more than 120,000 cross-references or "entry terms," including word-order variations (e.g., **Mental Disorders, Organic** see **Organic Mental Disorders**), which help point users from their own words to MeSH terms.

One problem with searching by MeSH terms has always been that articles indexed before a term is introduced are not indexed under that term. Indeed, the original description—the seminal paper—about any new topic is almost never indexed under that topic. Thus, when Gruntzig, Senning, and Siegenthaler described percutaneous transluminal coronary angioplasty in *The New England Journal of Medicine* in July 1979,[12] the indexers at the National Library of Medicine had no way to know that this new technique would become sufficiently important to justify a new medical subject heading. This article (as well as a few briefer, earlier descriptions) was indexed under terms such as **Angiography, Catheterization, Heart Catheterization,** and **Coronary Vessels**. As more articles were published, the importance of the technique became clearer, and in 1981, the medical subject heading **Angioplasty, Balloon** was introduced, followed, in 1989, by the term **Angioplasty, Transluminal, Percutaneous Coronary**.

To avoid missing references indexed prior to the introduction of a MeSH term, searchers can use text words in titles and abstracts to search citations indexed during prior years. One program[13] used to search the MEDLINE database uses algorithms that attempt to analyze the meaning of the title and abstract. Other MEDLINE interfaces search each subject term both as a MeSH term and as text words.

Accuracy of the MEDLINE Database

In the MEDLINE database, text words and authors' names are occasionally misspelled. Some of the misspellings originate in the printed manuscript. End users, too, misspell entries. If an end user and the MEDLINE database misspell a title word in the same way, and if the search software automatically maps from (misspelled) title words to MeSH, the user who misspells an entry can still find the relevant MeSH term and get a good result.[13]

Since searching by MeSH is so important, the question arises as to how accurately MeSH terms are applied. Funk and Reid studied the MEDLINE database for consistency of indexing,[14] but to our knowledge there are no published studies of accuracy. On the other hand, one of us has done searches to try "to catch" the indexers. Many of these searches uncovered apparent omissions, but examination of the original article revealed, with rare exception, why it was indexed as it was. Particularly in the area that is most important to searching—the assigning of MeSH terms—the MEDLINE database seems to be remarkably accurate.

MeSH, MEDLINE, the World Wide Web, and the Full-text Search Engine

In recent years, inspired in part by the need to find documents on the Web, full-text search engines have become popular. When using a full-text search engine, the user types a query in a designated area and clicks GO or SUBMIT to retrieve the result. Full-text search engines vary in how they select and present the "hits," but most try to present first those documents that contain as many of the query words as often as possible. Thus, a "hit" that contains all query

words many times and close together will be presented before a "hit" that contains only one of the query words once. Often the algorithms ignore prepositions and other "stop words," and they give more weight to rare terms, to terms that appear close to the beginning of the document, and to terms that appear in the text in the same order that they appear in the query.

Full-text search engines routinely deal with lexical variants, such as *anesthesia* and *anaesthesia* or *color* and *colour*, and with plural forms. To some extent they deal with synonyms. In addition, they often provide advanced functions that give the user some control over the search strategy, such as the ability to limit retrieval to items in which all the query words appear or to require that certain words be adjacent or near one another. In addition, sophisticated full-text searching systems exploit structured information about concepts and their relationships within controlled vocabularies. The Unified Medical Language System (UMLS) Metathesaurus,[15] which provides access to many controlled vocabularies, including MeSH, is used in natural language processing and full-text searching applications, including a number of Web search engines.

Full-text searching techniques enable retrieval across disparate collections of information, of which the Web is the largest and best example. Still-better retrieval would result if material on the Web were more uniformly structured and tagged and if search engines took better advantage of the structure and the tagging. The desire for better retrieval of information on the Web explains some of the current interest in "metadata" standards and mark-up languages, such as SGML and XML.

The Web is big, Web browsers are common, and full-text search engines are popular; as a result, users are familiar with typing in a designated area, pressing GO, and retrieving the result. The ability to search MEDLINE with the same look and feel as the rest of the Web has a lot of appeal. The authors, however, wish to inject a note of caution. In contrast to most other items on the Web, documents in the MEDLINE database are indexed uniformly, hierarchically, and accurately. Searching with MeSH terms is often the major contributor to the completeness and precision of a MEDLINE search. The way a full-text search engine maps a user's query to MeSH has a fundamental limitation; namely, that it does not dialog with the user to resolve ambiguity or to offer alternatives. Thus, the user who misspells a MeSH term, who does not know that a more precise MeSH term exists, who expresses a concept in words that map to several different MeSH terms, or who uses words that the full-text search engine cannot map to MeSH, is likely to get an inferior retrieval.

Not only are today's full-text search engines limited in their ability to map users' queries to MeSH, but in the case of MEDLINE they do not present hits in order of descending utility. When a search produces a large number of hits that cannot be cut down, we think it important that the software present items in the best medical journals first. An article published in *The New England Journal of Medicine* is far more likely to be helpful than is an article that satisfies the same search criteria but is published in a language other than English, in a priority 3 journal, and that lacks an abstract online.

Conclusion

This article has focused on MeSH in the context of the development of *Index Medicus* and MEDLINE, with which MeSH is inextricably connected. Certainly MeSH has had its primary influence in medical bibliographic retrieval. Yet, as the first major thesaurus designed for use in an automated system, MeSH is also a landmark in the representation of medical vocabulary and concept data in machine-readable form. In the 1980s, when the UMLS project began,[15] many medical informatics researchers began to examine the characteristics of machine-readable medical vocabularies and classifications closely. At that time, MeSH stood essentially alone as an example of a vocabulary with an explicitly tagged database format, context-free identifiers, polyhierarchy, an extensive cross-reference structure, text definitions, and a reliable update cycle. Today many of these characteristics are seen as critical requirements for controlled vocabularies that will be used in clinical systems.[16]

In the bibliographic arena, we think serious MEDLINE searching requires the ability to make lists of references by MeSH term, title word, abstract word, author, year, institution, source, and the like. It requires the ability to combine these lists with Boolean AND, OR, and occasionally AND NOT, and to be guided by the number of hits along the way. End users often need help finding medical subject headings and subheadings that they do not know or cannot spell, help navigating the MeSH hierarchy, and help with the names of authors and journals.

MeSH has had a glorious history, of which the National Library of Medicine can be justly proud.

MeSH is so powerful that, with few exceptions, any search that can be done with MeSH should be done with it. Because full-text search engines are currently limited in their ability to map a user's query to medical subject headings, there is a risk that MeSH will be de-emphasized. We hope this will not happen. May the first 40 years be just the beginning.

The authors thank Ms. Betsy Humphreys, Ms. Lucretia McClure, and others for their many thoughtful suggestions.

References ■

1. Taine SI. Bibliographic data processing at the National Library of Medicine: discussion. In: Goldhor H (ed). Proceedings of the 1963 Clinic on Library Applications of Data Processing, Apr 28–May 1, 1963. Champaign, Ill.: University of Illinois at Urbana-Champaign, 1964:124–5.
2. Kunz J. *Index Medicus*: a century of medical citation. JAMA. 1979;241:387–90.
3. Blake JB. From Surgeon General's bookshelf to National Library of Medicine: a brief history. Bull Med Libr Assoc. 1986; 74:318–24.
4. DeBakey ME. The National Library of Medicine: evolution of a premier information center. JAMA. 1991;266:1252–8.
5. Metcalf KD. The National Medical Library: Report of a Survey of the Army Medical Library. Chicago, Ill.: American Library Association, 1944.
6. Oswald JF. The Development of the Medical Subject Heading. Philadelphia, Pa.: Drexel Institute of Technology School of Library Science, 1955:56.
7. Adams S. The way of the innovator: notes toward a prehistory of MEDLARS. Bull Med Libr Assoc 1972 Oct; 60(4): 523-33.
8. Rogers FB. Review of *Emerging Solutions for Mechanizing the Storage and Retrieval of Information* (vol. 5 of Studies in Coordinate Indexing, by M. Taube and associates). Coll Res Libr. 1960;21:489–92.
9. Pizer IH. A regional medical library network. Bull Med Libr Assoc. 1969;57(2):101–15.
10. Sewell W. Medical Subject Headings in MEDLARS. Bull Med Libr Assoc. 1964;52:164–70.
11. McCarn DB. Planning for online bibliographic access by the Lister Hill National Center for Biomedical Communications. Bull Med Libr Assoc. 1970;58(3):303–10.
12. Gruntiz AR, Senning A, Siegenthaler WE. Nonoperative dilation of coronary-artery stenosis: percutaneous transluminal angioplasty. N Engl J Med. 1979;301:61–8.
13. Horowitz GL, Bleich HL. PaperChase: a computer program to search the medical literature. N Engl J Med. 1981;305(16): 924–30.
14. Funk ME, Reid CA. Indexing consistency in MEDLINE. Bull Med Libr Assoc. 1983;71:176–83.
15. Lindberg DA, Humphreys BL, McCray AT. The Unified Medical Language System. Methods Inf Med. 1993;32(4): 281–91.
16. Cimino JJ. Desiderata for controlled medical vocabularies in the 21st century. Methods Inf Med. 1998;37(4–5):394–403.