

Current Comments

Excerpta Medica—Abstracting the Biomedical Literature for the Medical Specialist

Number 28

July 14, 1980

Few entrepreneurs have been able to successfully combine an academic with an industrial or commercial career. Individuals like Carl Djerassi come to mind. In the world of scientific publishing an outstanding example is Pierre Vinken. He will be known to many *Current Contents*[®] (CC[®]) readers as the co-editor (with G.W. Bruyn) of the *Handbook of Clinical Neurology*,¹ a monumental compilation in 45 volumes.

Outside the world of science Vinken has kept a relatively low profile. Few of my library colleagues recognize his name as the driving force behind *Excerpta Medica* (EM), the subject of this essay. However, he is today the chief executive officer of Elsevier-NDU, the giant Dutch publishing conglomerate. Scientists and scholars who read CC will recognize that name. ISI[®] indexes over 150 of their journals, not to mention their books, in our various information services.

I first became acquainted with Vinken in the late 1950s. Since then we have been professional co-admirers, even though in the strictest sense our companies are competitors. Unfortunately, too often these days libraries have to make choices between information services that complement one another.

While I was visiting Portugal last year when we were both on vacation, I talked to Vinken about *Excerpta Medica*. EM is the largest English-language abstracting service in biomedicine. Of course, Information Retrieval Ltd.

(IRL), BIOSIS, and others include the biomedical literature in their coverage of the life sciences literature. I'll have more to say about IRL and BIOSIS publications in the future.

At the time I talked to Vinken, I'd just finished an essay on *Index Medicus* (IM),² the first in a series on various competitive information services I hope to write about in the future. EM has a number of features that distinguish it from other information systems, so I decided to let CC readers know more about it.

EM is similar to IM in that both are information retrieval services concentrating on the biomedical, and especially clinical, literature. Of course, the major difference between the two is that EM is also an abstracting service while the printed IM is an indexing service. Indexing services provide the article's "address"—a citation giving journal and article title, volume, issue, year of publication, and pagination. However, IM's computerized counterpart, MEDLINE, includes abstracts for more than 40% of the citations in its data base. Clifford Bachrach, editor of IM, says, "These are all but a few of the substantive English language abstracts published with the original articles."³

Abstracting services offer a brief summary of the article's content in addition to the citation. Abstracts save researchers time lost in tracking down articles that aren't as relevant as the titles alone may suggest. There is a substantial literature on the advantages and dis-

advantages of abstracts that I will not review here. Briefly, abstracts provide added value to the user but they also increase subscription costs and lengthen production time.

The 18th edition of *Ulrich's International Periodicals Directory* (1979-80)⁴ lists about 75 abstract serials in the biological and medical sciences. Most of them are limited to covering the literature of a particular field—immunology, cancer therapy, or psychopharmacology, for example. Only a few abstract serials attempt to cover the biomedical and clinical literature comprehensively. Among these are *Bulletin Sinaletique* based in Paris, *Referativnyi Zhurnal* in Prague, *Zentralblatt* in Berlin, *Referativnyi Zhurnal* in Moscow, and *Excerpta Medica* in Amsterdam.

EM was started in 1946 by a group of physicians as an independent non-profit foundation. The purpose of the Excerpta Medica Foundation was to help "the progress of medical knowledge by making available to medical and related professions information on all significant basic research and clinical findings reported in any language throughout the world."⁵ In those early days, *EM* appeared in 13 sections divided along the general lines of the medical school curriculum. The sections covered anatomy, pathology, internal medicine, physiology, and other basic science and clinical specialties.

However, the rapid growth of medical literature made such broad divisions impractical. Individual sections became too large for one group of editors to manage. Also, subscribers had a hard time locating their specialized interests within the general categories. As a result, large sections were split into smaller, separate sections.⁶ By the end of the 1960s, *EM* published over 30 sections—more than double the original number. Vinken was responsible for creating *EM's* computerized data base, from which these publications were generated. He was supported in this by Bob Blanken, a Philadelphia emigrant.

The increase in coverage and number of sections obviously added considerable costs for publishing and staffing. Occasionally, the Dutch government contributed funds to support *EM's* publications. In 1972, *EM* was acquired by Elsevier-NDU, one of the world's largest publishing houses for scientific books and journals. The Excerpta Medica Foundation still exists as a separate entity with no publishing responsibilities. Instead, the foundation organizes medical congresses, sponsors development programs in the medical information field, and grants a travel award to young biomedical researchers every two years.⁶

As a division of Elsevier-NDU, *EM* today publishes 43 abstract journals and two drug literature indexes for profit (see Table). These publications, appearing between ten and 30 times a year, are geared to the medical specialist's interests. No matter how narrow those interests are, *EM* claims to cover the most significant biomedical research in all fields and in all languages.⁵ A large part of this claim rests on *EM's* organization of editors and abstractors.

The editors who screen, index, and select articles to be abstracted in *EM* are practicing medical specialists and researchers living in the Amsterdam area. These executive and index editors are salaried by *EM* on a part-time basis. In addition, there is a volunteer International Editorial Board made up of biomedical specialists who aid in journal selection, indexing, and classification problems. They also recruit volunteer abstractors, who are also physicians and biomedical researchers, from around the world. By using an international network of paid and volunteer specialists as indexers and abstractors, *EM* is confident its publications represent the most important biomedical research published in all nations.⁷

EM's editors screen about 400,000 articles each year from more than 4,000 biomedical, chemical, and other scientific journals. *EM* also screens con-

Table. *Excerpta Medica* abstract journals and literature indexes. Numbers are not consecutive as a result of section reorganization or deletions.

Section Number	Title
1.	Anatomy, Anthropology, Embryology & Histology
2.	Physiology
3.	Endocrinology
4.	Microbiology: Bacteriology, Mycology & Parasitology
5.	General Pathology and Pathological Anatomy
6.	Internal Medicine
7.	Pediatrics & Pediatric Surgery
8.	Neurology & Neurosurgery
9.	Surgery
10.	Obstetrics & Gynecology
11.	Otorhinolaryngology
12.	Ophthalmology
13.	Dermatology & Venereology
14.	Radiology
15.	Chest Diseases, Thoracic Surgery & Tuberculosis
16.	Cancer
17.	Public Health, Social Medicine & Hygiene
18.	Cardiovascular Diseases & Cardiovascular Surgery
19.	Rehabilitation & Physical Medicine
20.	Gerontology & Geriatrics
21.	Developmental Biology & Teratology
22.	Human Genetics
23.	Nuclear Medicine
24.	Anesthesiology
25.	Hematology
26.	Immunology, Serology & Transplantation
27.	Biophysics, Bioengineering & Medical Instrumentation
28.	Urology & Nephrology
29.	Clinical Biochemistry
30.	Pharmacology & Toxicology
31.	Arthritis & Rheumatism
32.	Psychiatry
33.	Orthopedic Surgery
34.	Plastic Surgery
35.	Occupational Health & Industrial Medicine
36.	Health Economics & Hospital Management
37.	Drug Literature Index
38.	Adverse Reactions Titles
40.	Drug Dependence
46.	Environmental Health & Pollution Control
47.	Virology
48.	Gastroenterology
49.	Forensic Science
50.	Epilepsy
51.	Leprosy & Related Subjects

ference proceedings, books, monographs, and doctoral dissertations. The editors select about 240,000 articles each year for entry in *EM's* computerized data base (EMBASE) as citations. Of these, about 150,000 are abstracted

for *EM's* various sections, and an additional 60,000 are published as citations in the *Drug Literature Index* and *Adverse Reactions Titles*. Abstracts and citations are available to *EM* subscribers in print or on magnetic tape within six weeks to six months after receipt of the original document.⁷

EM claims that papers on even the most specialized subject can be rapidly located and retrieved from its large data base and various printed publications. *EM* relies on thorough and detailed subject indexing to make sure that all relevant papers on a given subject are accessible to its users. There are several reasons why *EM's* subject indexing is more specific than most other abstracting or indexing services.

First, *EM's* editors index on the basis of the entire article. In this way, *EM* believes many important concepts, not mentioned in the article's title or summary, can be used as headings precisely identifying the subject of the article.⁶ Most other services index on the basis of title and/or summary alone. *IM*, however, is a notable exception which, like *EM*, selects headings from the article's text, title, and summary.³

Also, *EM* allows its specialist medical editors considerable freedom when indexing articles. Most abstracting and indexing services require their editors to work from a controlled list of pre-selected subject headings when screening articles. But *EM* editors simply write down important concepts in their own words or in the words of the author.⁶ These terms are then fed into a computer which selects preferred terms and synonyms from an extensive thesaurus, called the Master List of Medical Indexing Terms (MALIMET). MALIMET consists of about 180,000 preferred terms and 250,000 synonyms.⁵

But the really distinctive feature of *EM's* subject indexing is that articles are indexed on two levels in *EM's* printed publications—primary and secondary terms.⁷ Primary terms refer to the paper's main concepts, including drug

and chemical names, diseases and syndromes, diagnostic procedures, and anatomical or physiological concepts. Secondary terms concentrate on the more quantitative information reported in a paper—size of test sample or duration of test period, for example. Since 1978, primary terms were themselves divided into "A" and "B" categories. "A" primary terms describe main concepts discussed in the paper, and "B" terms refer to related concepts not discussed in detail or mentioned briefly in the text.⁵ Editors can assign as many as 50 primary terms of each type (A or B) per abstract!⁸

The subject index entries in any of *EM's* printed abstract journals themselves represent a kind of "mini-abstract" giving specific information on the article's content.⁷ For example, if someone is interested in abdominal radiotherapy, the subject index gives the following information:

Abdominal radiotherapy, intestine trauma, radiation injury, radiotherapy, surgical management, radiation injured gut, 5.1% out of 1824 patients. [abstract #] 414.⁵

With such a detailed description of the article, the researcher can easily decide whether it's worth the time and effort to locate the full abstract, retrieve the article, or write the author for a reprint.

Abstracts can also be located through *EM's* classification system (EMCLAS). EMCLAS is made up of about 6,500 "pigeonholes" which determine where an abstract will appear in *EM's* printed journals. For example, *EM's* physiology section is divided into 26 "chapters," such as cell physiology, smooth muscle, digestion, and respiration. Each chapter is further divided according to functional properties. For example, "respiration" is divided into mechanics of breathing, gas exchange, and regulation. Each *EM* abstract journal and literature index has an average of 140 subdivisions.⁹

The subject indexing for abstracts and citations in EMBASE is even more specific. In addition to primary and secondary terms and EMCLAS headings, abstracts can be retrieved from EMBASE using the Item Index System (EMTAGS). EMTAGS consists of about 200 general terms describing, for example, the type of article (review, editorial, note), age of test group (infant, teenage, adult), kind of experimental animal (cat, pig, rat), and organ system studied (digestive tract, nervous system, respiratory tract).⁸

EMTAGS gives the kind of information not included in the article's citation, abstract, or the primary and secondary index terms assigned to it. Thus, searching EMBASE on-line will often produce more relevant articles than looking through the subject indexes of *EM's* printed abstract journals. A 1977 evaluation of drug information retrieval services by Mary Madden and Ann MacDonald, Australian Department of Health, is pertinent here.¹⁰ Comparing *IM* and *EM's Drug Literature Index*, they found that the manual, printed indexes produced only 70% of the relevant references retrieved from the computerized data bases, MEDLARS and *EM's DRUGDOC (EMD)*. (*EMD* should not be confused with RINGDOC, an abstracting service covering the pharmaceutical literature. RINGDOC is part of Derwent Publications Ltd., England.)

When they compared just the data bases, Madden and MacDonald found that *EMD* retrieved about twice as many references as MEDLARS for the ten drugs selected for the study. In fact, they concluded that "for drug-oriented searches of the literature, *Excerpta Medica DRUGDOC* was found to provide the most comprehensive and complete bibliography [of all nine information systems studied]. However... MEDLARS complements this system and they should be used in combination."¹⁰

Last February, the National Library of Medicine announced that articles on any chemicals (not only drugs) can be retrieved from MEDLARS by using the chemical's registry number, assigned by Chemical Abstracts Service, or its name and all synonyms for it.¹¹ By increasing its on-line indexing specificity, MEDLARS now offers direct access to information on about 20,000 chemicals.

Other studies indicate that *EM* provides more complete bibliographies than most other information services, and that the proportion of newer references is higher with *EM*. A study sponsored by the German pharmaceutical company E. Merck compared on-line search results for five data bases—BIOSIS, *Chemical Abstracts*, MEDLARS, RINGDOC, and *EM*.¹² It found that *EM* produced the best results, in terms of the number of relevant citations retrieved. RINGDOC, a data base built for the pharmaceutical industry, came in second. However, although RINGDOC only covers one-tenth the number of journals as *EM*, it still produced only 30% less relevant citations than *EM*! This is another illustration of the Bradford law of scattering.¹³

James Powell, Jr., Upjohn Company, Kalamazoo, Michigan, evaluated *EM*'s performance on-line and found that its coverage of the biomedical literature is complete and worldwide.¹⁴ He noted that *EM*'s in-depth indexing and extensive detailed subject heading thesaurus gives *EM* an edge over *IM*. He concludes, "Assuming that a given article is included in both data bases, the probability is much better that the article will be retrieved through *EM*'s indexing."¹⁴

But being the largest abstracting service in biomedicine has its disadvantages, too. It costs more to produce an abstracting service than an indexing service. Printing abstracts, hiring people to write abstracts and enter them in the data base, two-level indexing of the entire article—all this adds to the already high cost of publishing. Ob-

viously, the extra cost must be reflected in the subscription price. A complete set of *EM*'s 43 abstract journals today costs about \$8,000 per year. This may be too high a price for smaller research libraries faced with reduced budgets. In comparison, a yearly subscription to *IM* costs \$150, *Chemical Abstracts* is \$5,000, and *Science Citation Index*® \$4,500.

However, the average price for an individual *EM* abstract journal is about \$200 per year, about the same as a yearly subscription to a biomedical research journal. For example, *EM*'s psychiatry section, appearing in two volumes and 20 issues, provides about 8,000 abstracts per year at a cost of about \$260. *EM*'s cancer section, published in three volumes and 30 issues, annually provides about 12,000 abstracts for about \$375.

Remember that *EM* is geared to the specialist physician or researcher—it is divided into 43 sections so the specialist can subscribe directly to his or her field of interest. Moreover, there is some duplication between sections to ensure comprehensive coverage of the biomedical literature. An abstract is published in an average of 1.8 sections, and an abstract's citation may also appear in the two drug literature indexes. Thus any one section will report on research published in other related sections, saving the individual the expense of subscribing to more than one section.

Of course, overlapping coverage is not attractive to librarians. August La Rocco and Cyril Feng, University of Miami Medical School, performed a cost comparison study of *EM*'s abstracting journals in 1977.¹⁵ In that year, for the sections studied, *EM* published almost 53,000 abstracts at a unit cost to the library of 2.14 cents per abstract. Corresponding abstract journals—*Chemical Abstracts*, *Biological Abstracts*, and *Psychological Abstracts*, among others—charged an average of 0.9 cents per abstract. Thus, *EM* was more expensive by 138% overall.

The same study concluded that *EM's* practice of splitting sections into narrower subject fields increased the total subscription cost to libraries without increasing the number of abstracts published. The additional costs only covered the expense of separate covers, bindings, and indexes between split sections. Also, splitting sections sometimes made it more difficult to research a given topic. Instead of referring to one *EM* section, researchers would have to use an average of 2.41 sections to compile a complete bibliography on any of 100 subjects studied by La Rocco and Feng.¹⁵

There seems to be a substantial group of *EM* subscribers who feel that it is too expensive. This is not an unusual reaction to any commercial service that must compete with government subsidized services like *IM*. It is my opinion, however, that *EM's* major drawback as

a general abstracting service is that it does not produce a unified printed index. This isn't an easy task considering that its 45 separate publications are designed to serve the specialist. But I believe a unified index would assuage the feelings of librarians who must authorize *EM's* \$8,000 subscription cost. As time goes by and on-line use of *EM* increases, this objection may fade away! Whatever problems some librarians may have with it, *EM* provides a valuable service to its primary market of practicing physicians and clinical researchers.

* * * * *

My thanks to Linda Cooper and Alfred Welljams-Dorof for their help in the preparation of this essay. ©1980 ISL

REFERENCES

1. Vinken P J & Bruyn G W, eds. *Handbook of clinical neurology*. Amsterdam: North Holland. 45 vols.
2. Garfield E. Some reflections on *Index Medicus*. *Current Contents* (51):5-11, 17 December 1979.
3. Bachrach C. Personal communication. 16 June 1980.
4. Ulrich's *international periodicals directory*. New York: Bowker, 1979. 2156 p.
5. *Excerpta Medica user manual*. Amsterdam: Excerpta Medica, 1979. 77 p.
6. Blanken R R. *The Excerpta Medica abstract journals*. Amsterdam: Excerpta Medica, 1977. 11 p.
7. Blanken R R & Stern B T. *Excerpta Medica's system for the automated storage and retrieval of biomedical information*. *Fed. Proc.* 33:1719-21, 1974.
8. *The Excerpta Medica Mark II biomedical information system*. Amsterdam: Excerpta Medica, 1979. 28 p.
9. *Guide to the Excerpta Medica classification and indexing system*. Amsterdam: Excerpta Medica, 1978. 140 p.
10. Madden M & MacDonald A. An evaluation and comparison of nine drug information retrieval systems. *Drug Info. J.* 11:47-59, 1977.
11. Schuyler P L & Stander R L. Chemical term records in the MESH vocabulary file: searching chemicals in MESH and MEDLINE. *NLM Tech. Bull.* February 1980, p. 7-12.
12. Gruch W. Search corner. *Online Rev.* 2:122-3, 1978.
13. Garfield E. Bradford's law and related statistical patterns. *Current Contents* (19):5-12, 12 May 1980.
14. Powell J R, Jr. Evaluation of *Excerpta Medica* on-line. *Spec. Libr.* 67:153-7, 1976.
15. La Rocco A & Feng C. *Excerpta Medica* abstracting journals: a case study of costs to medical school libraries. *Bull. Med. Libr. Ass.* 65:255-60, 1977.