

## Chemical Abstracts Coverage of the Preclinical Sciences Journal Literature

ROBERT K. POYER

Coordinator of Public Services, Library, Medical University of South Carolina, Charleston,  
South Carolina 29425

Received July 6, 1981

This study examines Chemical Abstracts (CA) coverage of the journal literature cited by researchers from preclinical science departments affiliated with medical schools. Using references from 70 dissertations written between 1973 and 1977 in the fields of anatomy, biochemistry, immunology, microbiology, pathology, pharmacology, and physiology, coverage of journal articles was studied. Approximately 57% of the cited literature was covered. Biochemistry (83.5%), pharmacology (80.4%), and microbiology (76.2%) were covered the best while pathology (27.8%) was covered the worst. Coverage of anatomy (55.4%), immunology (40.1%), and physiology (42.4%) was dependent upon the research being conducted. With the advent of multidata base searching, Chemical Abstracts should be considered as a source of references for preclinical science researchers.

### INTRODUCTION

In a recent study of substituting Chemical Abstracts (CA) online for the printed copy, Baldinger et al. questioned whether such an expensive tool (\$5000 in 1981) should be purchased when it is used by only a small fraction of its patrons.<sup>1</sup> Although pharmacology and biochemistry faculty, staff, and graduate students at the Medical University of South Carolina are the primary users of CA in our library, we had not looked at CA's potential for covering the literature used by other preclinical science researchers. With the advent of multidata base computer searching and the recent development of automatic cross data base searching capabilities by BRS, Lockheed, and SDC, the feasibility of using CA as a potential source of information has increased. Two questions arise. First, does CA cover references cited by other preclinical science researchers? If so, should CA be considered as a source of references? Second, how well does CA cover the literature cited by biochemists and pharmacologists who work in preclinical science departments associated with medical schools?

We examined these questions by looking at CA's coverage of journal articles cited in dissertations written in the preclinical sciences which would also provide references to literature cited in clinical medicine. Dissertations were chosen for a number of reasons. First, dissertations require a rigorous and thorough review of the literature.<sup>2</sup> Second, the research presented in the dissertation is overseen by a committee of the graduate faculty which should insure that the relevant literature be known and cited. Third, a majority of the dissertation authors in this study had published or had in press journal articles which were part of or were reporting the results of their dissertations. Most of these articles were published in what Garfield identifies as "significant journals of science".<sup>3</sup> This indicates quality research was done and was reported and accepted (via publication) by the scientific community. We hoped the comprehensive and thorough coverage of the literature would provide a sound basis for evaluating coverage by CA.

### MATERIALS AND METHODS

Seventy dissertations, written between 1973 and 1977 by individuals from preclinical science departments affiliated with medical schools, were chosen. The dissertations were obtained from SUNY Upstate Medical Center (33), Medical University of South Carolina (22), University of Mississippi, Jackson (9), Duke University (3), University of Florida (2), and University of Tennessee, Memphis (1). Ten dissertations in each of the following disciplines were studied: anatomy, biochemistry,

immunology, microbiology, pathology, pharmacology, and physiology.

A total of 7969 journal articles including references to review journals were identified and verified; abstracts, letters to the editor, and editorials were not included. Each reference was searched in the author index of CA the year it was published to see if the article was listed. The abstract was then consulted to determine if the reference cited was actually covered. In a few cases an author published a paper with the same title in two journals. If both papers were not readily available to check content, the author used the journal article cited in the dissertation to determine coverage by CA. If an article was not covered, we searched the following year up to 3 years after publication, because not all articles are indexed the year they are published.

### COVERAGE BY CHEMICAL ABSTRACTS

A total of 4549 or 57.1% of the journal references were covered by Chemical Abstracts; if the 35 articles published before 1907 are excluded, the percentage rises to 57.3%. (Table I shows coverage of those journal titles cited a minimum of six times). Coverage by discipline which is shown in Table II varies from a high of 83.5% for biochemistry to a low of 27.8% for pathology. Coverage by decade which is shown in Table III does not vary too much from decade to decade; CA did cover the 1970-1977 years the best.

### DISCUSSION

Before looking at CA's coverage we must look at the criteria used to choose an article for inclusion. CA attempts to cover the worldwide literature of chemistry by scanning over 14 000 journals, looking for articles that contain new chemical information. If an article cites a known chemical method, technique, or fact, it would not necessarily be indexed. For instance, thousands of researchers cited Lowry's article on protein measurement<sup>4</sup> but not all of those articles are indexed by CA. One should consult "Subject Coverage and Arrangement of Abstracts by Sections in Chemical Abstracts"<sup>5</sup> because it describes what is covered in CA as well as what is not covered.

When one examines the journal titles cited and covered, a number of observations are evident. First, CA appears not to cover every article from the biochemical journals. It is difficult to ascertain whether they were missed inadvertently or because of indexing policy. Second, the journals cited not only cover the preclinical sciences but also clinical medicine titles. Although there are a multitude of journals cited, CA indexes and abstracts at least a couple of articles from most

Table I. CA Coverage of Journals Cited at Least Six Times

title	cited	covered	title	cited	covered
<i>Acta Endocrinol.</i>	17	15	<i>Curr. Top. Microbiol. immunol.</i>	9	3
<i>Acta Med. Scand.</i>	9	5	<i>Diabetes</i>	13	9
<i>Acta Pathol. Microbiol. Scand.</i>	20	5	<i>Diabetologica</i>	16	13
<i>Acta Physiol. Scand.</i>	52	18	<i>Endocrinology</i>	184	154
<i>Adv. Cancer Res.</i>	8	7	<i>Eur. J. Biochem.</i>	23	23
<i>Adv. Exp. Med. Biol.</i>	9	6	<i>Eur. J. Immunol.</i>	31	8
<i>Adv. Immunol.</i>	27	12	<i>Experientia</i>	13	12
<i>Am. Heart J.</i>	11	1	<i>Exp. Brain Res.</i>	6	3
<i>Am. J. Anat.</i>	30	4	<i>Exp. Cell Res.</i>	24	10
<i>Am. J. Cardiol.</i>	23	15	<i>Exp. Neurol.</i>	16	1
<i>Am. J. Clin. Nutr.</i>	8	3	<i>Exp. Parasitol.</i>	10	3
<i>Am. J. Clin. Pathol.</i>	16	3	<i>FEBS Lett.</i>	15	14
<i>Am. J. Dis. Child.</i>	7	0	<i>Fed. Proc., Fed. Am. Soc. Exp. Biol.</i>	39	24
<i>Am. J. Med. Sci.</i>	8	4	<i>Gastroenterology</i>	15	8
<i>Am. J. Med.</i>	20	4	<i>Gen. Comp. Endocrinol.</i>	9	8
<i>Am. J. Obstet. Gynecol.</i>	28	6	<i>Genetics</i>	12	7
<i>Am. J. Ophthalmol.</i>	21	3	<i>Histochemie</i>	7	7
<i>Am. J. Pathol.</i>	60	11	<i>Hoppe-Seylers Z.</i>		
<i>Am. J. Physiol.</i>	291	189 (1) <sup>a</sup>	<i>Physiol. Chem.</i>	7	7
<i>Am. J. Roent.</i>	8	1	<i>Hum. Genet.</i>	12	1
<i>Am. Rev. Respir. Dis.</i>	23	7	<i>Hum. Hered.</i>	6	1
<i>Anal. Biochem.</i>	21	21	<i>Immunochemistry</i>	13	13
<i>Anal. Chem.</i>	12	12	<i>Immunology</i>	50	26
<i>Anat. Rec.</i>	26	3	<i>Infect. Immunol.</i>	21	7
<i>Ann. Intern. Med.</i>	25	7	<i>Int. Arch. Allergy Appl. Immunol.</i>	10	6
<i>Ann. N. Y. Acad. Sci.</i>	87	48	<i>Int. J. Cancer</i>	55	8
<i>Ann. Rheum. Dis.</i>	15	5	<i>Int. Rev. Cytol.</i>	6	3
<i>Ann. Surg.</i>	9	3	<i>Invest. Ophthalmol. Vis. Sci.</i>	8	3
<i>Annu. Rev. Biochem.</i>	17	17	<i>JAMA</i>	19	2
<i>Annu. Rev. Microbiol.</i>	8	6	<i>Jpn. J. Pharmacol.</i>	7	7
<i>Annu. Rev. Physiol.</i>	8	3	<i>Johns Hopkins Med. J.</i>	17	2 (1)
<i>Arch. Biochem. Biophys.</i>	45	45	<i>J. Am. Chem. Soc.</i>	22	22
<i>Arch. Ges. Virusforsch.</i>	7	3	<i>J. Am. Oil Chem. Soc.</i>	18	18
<i>Arch. Intern. Med.</i>	6	1	<i>J. Anat.</i>	22	2 (1)
<i>Arch. Int. Pharmacodyn. Ther.</i>	13	12	<i>J. Appl. Physiol.</i>	17	2
<i>Arch. Ophthalmol.</i>	20	3	<i>J. Bacteriol.</i>	79	58
<i>Arch. Pathol.</i>	37	10	<i>J. Biochem.</i>	16	16
<i>Arch. Surg.</i>	8	1	<i>J. Biol. Chem.</i>	265	262
<i>Arthritis Rheum.</i>	10	3	<i>J. Cell Biol.</i>	87	53
<i>Atherosclerosis</i>	6	4	<i>J. Cell Physiol.</i>	20	17
<i>Aust. J. Exp. Biol. Med. Sci.</i>	9	1	<i>J. Clin. Endocrinol. Metab.</i>	19	14
<i>Bacteriol. Rev.</i>	9	6	<i>J. Clin. Invest.</i>	110	94
<i>Biochem. Biophys. Res. Commun.</i>	82	81	<i>J. Clin. Pathol.</i>	12	7
<i>Biochem. J.</i>	93	87	<i>J. Comp. Neurol.</i>	31	5 (1)
<i>Biochem. Pharmacol.</i>	28	27	<i>J. Comp. Physiol. Psychol.</i>	8	1
<i>Biochem. Z.</i>	10	9	<i>J. Embryol. Exp. Morphol.</i>	7	2
<i>Biochemistry</i>	81	78	<i>J. Endocrinol.</i>	41	32
<i>Biochim. Biophys. Acta</i>	133	128	<i>J. Exp. Med.</i>	313	115
<i>Biol. Reprod.</i>	6	3	<i>J. Exp. Zool.</i>	6	1
<i>Blood</i>	13	7	<i>J. Gen. Microbiol.</i>	8	4
<i>Brain Behav. Evol.</i>	6	1	<i>J. Gen. Physiol.</i>	46	26
<i>Brain Res.</i>	65	30	<i>J. Gen. Virol.</i>	21	16
<i>Br. J. Cancer</i>	8	1	<i>J. Histochem. Cytochem.</i>	33	31
<i>Br. J. Exp. Pathol.</i>	31	17	<i>J. Immunol. Meth.</i>	11	5
<i>Br. J. Haematol.</i>	6	1	<i>J. Immunol.</i>	188	93
<i>Br. J. Ophthalmol.</i>	13	0	<i>J. Infect. Dis.</i>	17	3
<i>Br. J. Pharmacol.</i>	18	15	<i>J. Lab. Clin. Med.</i>	43	30
<i>Br. Med. Bull.</i>	6	2	<i>J. Lipid Res.</i>	25	25
<i>Br. Med. J.</i>	16	3	<i>J. Med. Chem.</i>	11	11
<i>Can. J. Biochem.</i>	16	13	<i>J. Mol. Biol.</i>	115	106
<i>Cancer</i>	30	4	<i>J. Mol. Cell Cardiol.</i>	14	6
<i>Cancer Chemother. Rep.</i>	7	3	<i>J. Natl. Cancer Instit.</i>	97	28
<i>Cancer Res.</i>	78	46	<i>J. Neurochem.</i>	35	34
<i>Cardiovascular Res.</i>	6	4	<i>J. Neurol. Neurosurg. Psychiatr.</i>	7	0
<i>Cell</i>	8	7	<i>J. Neuropathol. Exp. Neurol.</i>	6	0
<i>Cell Immunol.</i>	62	21	<i>J. Neurophysiol.</i>	16	2
<i>Circulation</i>	34	5	<i>J. Nucl. Med.</i>	6	5
<i>Circ. Res.</i>	80	50	<i>J. Nutr.</i>	18	8
<i>Clin. Chim. Acta</i>	16	16	<i>J. Parasitol.</i>	6	2
<i>Clin. Chem.</i>	9	9	<i>J. Pathol.</i>	13	6
<i>Clin. Exp. Immunol.</i>	32	12	<i>J. Pediatr.</i>	6	0
<i>Clin. Immunol. Immunopathol.</i>	9	4	<i>J. Pharm. Sci.</i>	7	7
<i>Clin. Pharmacol. Ther.</i>	6	4	<i>J. Pharmacol. Exp. Ther.</i>	63	59
<i>Clin. Sci.</i>	18	8	<i>J. Pharm. Pharmacol.</i>	9	9
<i>Cold Spr. Harbor Sym. Quant. Biol.</i>	39	29	<i>J. Physiol.</i>	147	57 (12)
<i>Comp. Biochem. Physiol.</i>	13	12	<i>J. Protozool.</i>	15	8
<i>C. R. Hebd. Seances Acad. Sci. D</i>	23	10 (1)	<i>J. Reprod. Fertil.</i>	11	8
<i>C. R. Seances Soc. Biol. Ses. Fl.</i>	12	6	<i>J. Reticulendothel. Soc.</i>	16	2

Table I (Continued)

title	cited	covered	title	cited	covered
<i>J. Theor. Biol.</i>	7	6	<i>Prog. Brain Res.</i>	9	2
<i>J. Ultrastruc. Res.</i>	17	7	<i>Prog. Cardiovasc. Dis.</i>	6	2
<i>J. Virol.</i>	99	81	<i>Prog. Med. Virol.</i>	7	0
<i>Kidney Int.</i>	9	5	<i>Prog. Nucleic Acid Res. Mol. Biol.</i>	8	7
<i>Klin. Wochenschr.</i>	7	3 (1)	<i>Psychopharmacology</i>	10	8
<i>Lab. Invest.</i>	52	16	<i>Q. J. Exp. Physiol.</i>	7	0
<i>Lancet</i>	69	49 (1)	<i>Radiat. Res.</i>	7	3
<i>Life Sci.</i>	26	21	<i>Rec. Prog. Horm. Res.</i>	12	11
<i>Lipids</i>	13	13	<i>Scand. J. Clin. Lab. Invest.</i>	18	14
<i>Lymphology</i>	6	0	<i>Scand. J. Immunol.</i>	8	3
<i>Metabolism</i>	12	12	<i>Science (Washington, D.C.)</i>	196	113
<i>Microvas. Res.</i>	10	1	<i>Steroids</i>	7	6
<i>Mol. Gen. Genet.</i>	16	14	<i>Surgery</i>	20	1
<i>Mol. Pharmacol.</i>	6	6	<i>Surg. Gynecol. Obstet.</i>	18	2
<i>Natl. Cancer Instit. Mongr.</i>	30	6	<i>Surg. Forum.</i>	6	0
<i>Nature (London)</i>	281	185	<i>Trans. Assoc. Am. Physic.</i>	6	2
<i>Nauyn-Schiedebergs Arch. Pharmacol.</i>	6	5	<i>Trans. N. Y. Acad. Sci.</i>	8	1
<i>Neuroendocrinology</i>	13	11	<i>Transplantation</i>	96	15
<i>Neuropharmacology</i>	6	5	<i>Transplant Proc.</i>	38	6
<i>N. Eng. J. Med.</i>	65	14	<i>Transplant Rev.</i>	35	12
<i>Pediatrics</i>	7	0	<i>Virchows (Pathol. Anat.)</i>	9	5
<i>Pfluegers Arch.</i>	33	10 (1)	<i>Virology</i>	264	150
<i>Pharmacol. Rev.</i>	20	19	<i>Z. Naturforsch.</i>	6	3
<i>Physiol. Rev.</i>	30	16	<i>Z. Zellforsch.</i>	24	8
<i>Proc. Natl. Acad. Sci. U.S.A</i>	264	213			
<i>Proc. R. Soc. London., Ser. B</i>	26	5 (1)			
<i>Proc. Soc. Exp. Biol. Med.</i>	178	116			
<i>Prog. Allergy</i>	8	5			

<sup>a</sup> Number of articles from the journal published prior to 1907, the year Chemical Abstracts began publication.

Table II. CA Coverage by Discipline

discipline	ref cited	ref covered (%)
anatomy	863	478 (55.4%)
biochemistry	989	826 (83.5%)
immunology	1715	688 (40.1%)
microbiology	1436	1094 (76.2%)
pathology	1042	290 (27.8%)
pharmacology	940	756 (80.4%)
physiology	984	417 (42.4%)
total	7969	4549 (57.1%)

Table III. CA Coverage by Year

year	ref cited	ref covered (%)
1871-1906 <sup>a</sup>	35	0 (0%)
1907-1909	12	6 (50.0%)
1910-1919	26	13 (50.0%)
1920-1929	56	34 (60.7%)
1930-1939	152	54 (33.3%)
1940-1949	260	120 (46.2%)
1950-1959	830	447 (53.9%)
1960-1969	3060	1741 (56.9%)
1970-1977	3528	2134 (60.5%)
total	7969	4549 (57.1%)

<sup>a</sup> Chemical Abstracts began publishing in 1907.

titles. This demonstrates not only the interdisciplinary nature of scientific research but also the wide journal coverage by CA. It also points out the selectivity in choosing articles for inclusion.

An examination of coverage by discipline shows CA covering 83.5% of the biochemistry and 80.4% of pharmacology journal literature cited. One of the dissertations written under auspices of a biochemistry department on the characterization of feline leukemia virus accounted for almost 50% of the biochemistry references not covered by CA; many of these references were to cancer and virology journals. Many of the remaining references not covered come from physiology, endocrinology, and medical journals, which is not surprising since the authors are affiliated with medical schools. As for pharmacology, a majority of the journal articles not covered by CA are from

medical titles, especially cardiology; also 32 references not covered were from physiology titles.

Over 76% of the references cited in the microbiology dissertations were covered by CA. At first glance this was surprising, but when the journals cited in this study are compared, with those of Sengupta's,<sup>6</sup> this result should be expected. Of the 30 highest cited journals in this study, 23 are in Sengupta's top 30 which contain a number of biochemical and interdisciplinary titles. Perhaps the reason CA covers microbiology so well can be explained by Sengupta's analysis of microbiology. Microbiology is going in new directions, particularly in immunochemistry, biochemical virology, and molecular genetics, with a shift away from the medical aspects. It appears these shifts in microbiological research are major areas covered by CA.

Chemical Abstracts covers between 40 and 55% of the references cited in immunology, physiology, and anatomy. In these disciplines coverage by CA is determined by the particular research being conducted. For example, in physiology, 90% of the references cited on PAH and octanoate transport in necturus kidney slices were covered while only 4% were covered in a dissertation on fusimotor innervation in primate muscle spindle.

As expected CA did not cover the journals cited in the pathology dissertations. This is due primarily to dependence on the medical literature. Most clinical medicine articles do not normally contain new chemical information and thus would not be indexed by CA.

## SUMMARY

As one might expect, Chemical Abstracts is a prime source of references in the areas of biochemistry and pharmacology and, as this study shows, microbiology. Use of CA for locating immunology, physiology, and anatomy references depends on the research a scientist is doing or planning to do. CA appears to be a poor choice for locating references for pathology.

This study indicates Chemical Abstracts does contain many references cited by scientists in the preclinical sciences. With the advent of multidata base searching on the dictionary files of BRS (CROS), DIALOG (DIALINDEX), and SDC (DBI),

one will be able to determine relatively quickly whether CA covers the literature needed by these researchers.

#### ACKNOWLEDGMENT

I would like to thank Lilian Peng and Kathy Brown who looked up a number of references in Chemical Abstracts to determine coverage.

#### REFERENCES AND NOTES

- (1) Baldinger, E. L.; Nakeff-Plaats, J. P. S.; Cummings, M. S. "An Experimental Study of the Feasibility of Substituting Chemical Abstracts

- On-Line for the Printed Copy in a Medium-Sized Library". *Bull. Med. Libr. Assoc.* **1981**, *69*, 247-251.
- (2) Fox, D. J. "The Research Process in Education"; Holt, Rinehart and Winston: New York, 1969.
  - (3) Garfield, E. "Significant Journals of Science" *Nature (London)* **1976**, *264*, 609-615.
  - (4) Lowry, O. H.; Rosebrough, N. J.; Farr, A. L.; Randall, R. J. "Protein Measurement with the Folin Phenol Reagent". *J. Biol. Chem.* **1951**, *193*, 265.
  - (5) "Subject Coverage and Arrangement of Abstracts by Sections in Chemical Abstracts", 1975 ed.; Chemical Abstracts Service: Columbus, OH, 1974.
  - (6) Sengupta, I. N. "The Literature of Microbiology". *Int. Libr. Rev.* **1974**, *6*, 353-369.

## Molecular Substructure Searching: Computer Graphics and Query Entry Methodology

W. J. HOWE\* and T. R. HAGADONE\*

The Upjohn Company, Kalamazoo, Michigan 49001

Received July 27, 1981

The increased availability of interactive graphics hardware has enabled the construction of true end-user oriented substructure search (SS) systems. However, because of the complexities inherent in the structural definition of substructure search queries, the design of a graphical query module should be accompanied by a careful consideration of SS query methodology. This paper discusses some of the important design considerations and then describes features and capabilities of the SS query definition module of Upjohn's COUSIN system. Techniques are presented for graphical entry of the core substructure of interest, for indicating variable atom, group, or fragment attachments, for specifying indefinite positioning of groups, and for tightening or relaxing search constraints on individual bonds. The R-group notational method for variable attachments and indefinite positioning also appears to have utility outside the context of a computerized retrieval system.

#### INTRODUCTION

Interactive graphics is a tool which has been used quite effectively as a full structure input medium in a variety of chemistry-based computer systems.<sup>1</sup> Even more effectively, graphics can be used to permit direct end-user specification of substructure search queries; yet relatively little work has been done in this area.<sup>2</sup>

The nature of substructure searching is such that the design of a graphical front-end should recognize and be guided by two important factors: (1) the data to be input is usually very complex and (2) the end users of an SS facility may run searches as often as daily or as infrequently as once a year. The graphical SS query module must be designed, therefore, with the relatively computer-naïve user in mind. This is not easy to do since, on the one hand, the graphical controls must be sophisticated enough to permit reasonably complex queries to be entered and executed with a single search, eliminating (as much as possible) the need for post-search manipulations of result files. Among other things, the graphical controls should therefore provide a capability for indicating variable structural units that are allowed at a particular location in the substructure, indefinite positioning of groups, and variable constraints on bond-type specificity. Yet on the other hand, characteristics of the user population require that the system be as simple to use and error tolerant as possible. This can be accomplished through a variety of means, the most effective of which is continuous system monitoring of the "chemical reasonableness" of the query as it is being constructed; the degree of handholding and the level of detail in messages suggesting ways to fix a problem should increase automatically as the user gets into more complex regions of the query.

With the preceding goals in mind we began several years ago to develop a graphical substructure search module for Upjohn's COUSIN system.<sup>4</sup>

#### OVERVIEW OF COUSIN'S GRAPHICAL SS QUERY MODULE

COUSIN is an interactive graphics-based chemical and biological information system which currently operates on a data base of 65 000 compounds. It is end-user oriented and provides capabilities for not only substructure searching but also structure retrieval and display, compound registry, full structure searching, retrieval and display of biological screening data, file manipulation, report generation, and searches keyed on a variety of textual and numeric data types associated with each compound. The following discussion will deal only with the portion of COUSIN that is used to define a substructure search query. Other parts of the system, including the part that actually carries out a substructure search once the query has been entered, will be described elsewhere.<sup>5</sup>

**Method of Interaction.** To interact with the COUSIN system, the user may either type commands on a keyboard (the typed characters appear on the graphics screen) or draw structure diagrams and associated symbols using a graphics tablet and stylus (the drawing appears on the screen). Examples of the former are commands to display particular compounds, combine files, activate the drawing controls, and so on. On the other hand, the tablet is used when the data to be entered is pictorial, such as structures to be registered or SS queries to be defined. Either the tablet or the keyboard is active at a given time.

When the user indicates via a typed command that he wishes to enter a substructure search query, the system responds by displaying the drawing controls pictured in Figure 1 (details of the drawing controls are shown in Figures 2-4). The stylus and tablet are also activated. The stylus is held like a pen, with the tip touching the horizontal surface of the tablet which sits in front of the display. As the user moves the stylus across the tablet, a cursor or "tracking cross" follows the motion on