# Computerized Drug Information Services*

DANIEL R. SMITH, R. O. BEAUCHAMP, Jr.,**
JOSEPHINE L. GARBER, and MARVEL A. DAUGHERTY
A. H. Robins Company Research Laboratories, Richmond, Va. 23220

Received November 24, 1971

**To compare computerized services in chemistry, pharmacology, toxicology, and clinical medicine of pharmaceutical interest, equivalent profiles were run on magnetic tape files of *CA-Condensates, CBAC, Excerpta Medica, MEDLARS* and *Ringdoc.*** The results of **these searches, which covered individual chemical compounds used in pharmacology and medicine, are tabulated for overlap of services, relative speed of citing references, and unique areas of journal coverage.**

In recent years, retrieval of published information of interest to the pharmaceutical industry has been augmented by the availability of a variety of computer bases.[1-5] The references contained in these computer bases range over several disciplines, including organic chemistry, biochemistry, drug metabolism, pharmacology, toxicology, and clinical and veterinary medicine. In some services, emphasis is primarily on chemistry and pharmacology, excluding clinical results; others stress clinical medicine and superficially treat chemistry and pharmacology. The computer bases offer a method of retrieving technical data on either a current awareness or retrospective basis and expand techniques presently limited primarily to printed indexes.

A study was initiated to determine the feasibility of utilizing external computer services in conjunction with an internal computer base and accessible library facilities. The external services were selected for their potential value related to research queries on product information and biological activities in the broadest sense. In addition, the procedures and methods for posing computer questions were evaluated. Chemical retrieval by means of line notations, fragments, words, or word fragments were among factors considered as well as subject areas encompassing diverse types of pharmacological and chemical activity. The forms of the computer output were compared—i.e., a printout with only keywords and literature citations *vs.* a more detailed reference including an abbreviated abstract.

From the foregoing general approach, several computer bases were evaluated to ascertain if present research information functions could supplement or fullfill the needs of research and management, and also cover present and anticipated product information. Cost of each service as well as its scope of journals influenced selection of the computer base.

* Presented at the Division Chemical Literature, Symposium-Critique of Chemical Information, 161st Meeting, ACS, Los Angeles, Calif., March 29, 1971.

** To whom correspondence should be addressed.

## DESCRIPTION OF COMPUTER BASES

Data bases analyzed for coverage were: (1) *Chemical Abstracts Condensates (CAC),* (2) *Chemical-Biological Activities (CBAC),* (3) *Excerpta Medica,* (4) *MEDLARS (Index Medicus),* and (5) *Ringdoc.* Results are illustrated in Figure 1.

The data bases were compared as to number and type of periodicals abstracted, rapidity of receipt of abstract, form of abstract, and methods of submitting computer questions. The number of journals abstracted and approximate elapsed time for publication of abstracts are listed in the following table.

| Computer Base | No. Jnls. Abstr. | Approx. Time Elapsed for Printed Abstr. to Appear |
|---|---|---|
| *CAC* | 12,000 | 2–4 mos. |
| *CBAC* | 580 | 2–4 mos. |
| *Excerpta Medica* | 3,500 | 3–6 mos. |
| *MEDLARS* | 2,300 | 2–6 mos. |
| *Ringdoc* | 330 | 2–4 mos. |

The sources of the computer bases utilized in this study and the number of years of reference material on tapes for possible retrospective searches are listed as follows:

| Computer Base | Abstract Backlog, Yrs. | Source (or Location) |
|---|---|---|
| *CAC* | Sept. 1968 on | University of Georgia |
| *CBAC* | 1965 on | University of Georgia and in-house |
| *Excerpta Medica* | 1968 on | 3i (Philadelphia, Pa.) |
| *MEDLARS* | 1964 on | Pharmaceutical Manufacturers Assoc. (PMA) |
| *Ringdoc* | July 1964 on | In-house only |

Although the backlog was not pertinent to the comparison study under consideration, it is an important factor in retrospective searches which are required in the routine functions of research information.

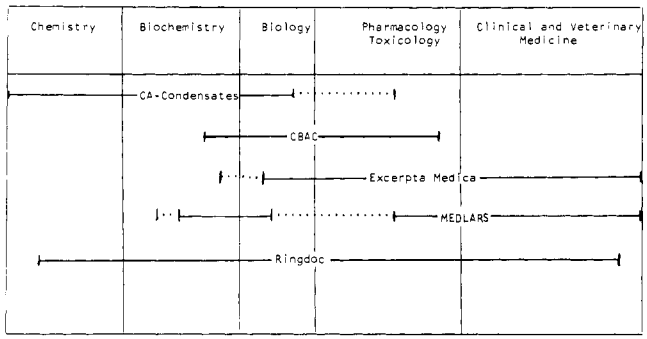


Figure 1. Pharmaceutical information spectrum

——Broad coverage .... Partial coverage

## SELECTION AND PROCEDURES

The five abstracting services chosen for this study were available on magnetic tapes either through purchase or by sending individual questions to service bureaus. As stated in the previous section, *CA-Condensates (CAC)*, the magnetic tape version of *Chemical Abstracts*, covers completely the chemical and medicinal chemistry areas. Although a very large data base of more than 12,000 journals is available, the subjects of pharmaceutical interest comprise only a portion of the total. The keyword approach to questioning somewhat limits its usefulness.

*CBAC* which may be considered a subset of the *CA-Condensates* data base, contains abstracts from 580 journals and has certain advantages in that the entire text of the abstracts may be searched. In a pharmaceutical search, emphasis on biological activity (not clinical) often provides fewer false drops than *CA-Condensates*.

Of the services evaluated, *Ringdoc* covers the widest range of information, but, like *CBAC*, has a limited data base. Information areas extend from synthetic chemistry to clinical medicine, but only 330 journals are abstracted by *Ringdoc*. Access to these computer files is possible through several methods such as chemical fragmentation code and/or an extensive biological code or a keyword retrieval system.

*MEDLARS* and *Excerpta Medica* provide the most comprehensive coverage of the medical field from 2300 and 3500 journals, respectively. *MEDLARS*, like *CA-Condensates*, is searched primarily through keywords in a controlled vocabulary. *Excerpta Medica* also uses a controlled vocabulary, but is more accessible for chemical searches. This is made possible by an internal thesaurus and an associated Wiswesser Line Notation chemical tape.

Table I Compounds Searched in Computer Bases

Group I. Compound names are MeSH Terms
1. Brompheniramine
2. Butaperzaine
3. Doxapram
4. Fenfluramine
5. Metaxalone
6. Glycopyrrolate
7. Methamphetamine
8. Methenamine
9. Methocarbamol
10. Phenazopyridine
11. Pheniramine
12. Tybamate

Group II. Compound names are not MeSH Terms
1. Flupenthixol
2. INPEA
3. Sulfameter

A total of fifteen compounds (Table I) was submitted to each data base to determine potential value to the research information user by establishing time lapse in receiving computer output, ease of submitting computer questions, and speed of abstracting and overlap. These compounds are arranged in two groups according to whether or not the terms were available in MeSH (*Me*dical *S*ubject *H*eadings index terms, used in *Index Medicus*).

Questions submitted to *MEDLARS*, *CAC*, and *CBAC* include chemical fragments, complete words, word fragments, combinations thereof, etc.; utilizing AND, OR, and NOT logic.

Approximate time intervals for receiving computer output are indicated in the table below.

| Computer Base | Output Received |
|---|---|
| *CAC* | Biweekly |
| *CBAC* | Biweekly |
| *Excerpta Medica* | Weekly |
| *MEDLARS* | Monthly |
| *Ringdoc* | Weekly |

Each of the external computer centers maintains a competent staff and has cooperated fully in these experimental trials runs.

Retrospective searches have also been submitted to each data base, using the same ground rules for submitting questions as for current awareness searches.

## RESULTS

References were collected over an eight-month period from the five computer bases, and results were tabulated at six and eight months for comparative purposes. Though this may not be sufficient time to draw any far reaching conclusions, certain trends have been noted and may be considered in planning search requirements for research information.

Table IIA shows the total number of reference citations from the six-month tabulation and lists the number of these citations retrieved by each tape service. To derive the largest possible number of relevant hits from the *MEDLARS* tapes, the compounds chosen for this group were in the MeSH term list (See Table I, Group I). This approach was expected to produce maximum overlap between *MEDLARS* and the other services. The "total number of references cited" in the first column indicates the number of different articles cited and not the total number of references received from all printouts. For example, a citation from *Ringdoc* to an article previously cited in a *CA-Condensates* printout would add one reference to the *Ringdoc* column but not to the total column on the left. Therefore, the sum of the individual columns usually exceeds the total reference column. The figures in parenthesis represent the per cent retrieval by each service based on the totals as shown in the first column.

Totals in Table IIA show that the maximum yield from any one service was 63% of the total references in the file at that time. Although *MEDLARS* and *Excerpta Medica*, because of their large journal coverage in the medical field, produced far more hits than the other services, neither can be called a "complete service" at this point, at least for fulfilling the total reference needs of the pharmaceutical industry.

Table IIB, a similar tabulation taken two months later, shows the changes in the percentages of total citations retrieved by the individual services. The bar graphs in Figure 2 illustrate clearly the comparison of the results of the totals for this compound group in Table II. The per-

Table II. Reference Citations Retrieved from Tape Services

| Compounds Searched | Total No. of References Cited Excluding Duplications | No. of References (and % of Total) Cited by | | | | |
|---|---|---|---|---|---|---|
| | | *CA-Condensates* | *CBAC* | *Excerpta Medica* | *Medlars* | *Ringdoc* |
| **A. SIX-MONTH TABULATION** | | | | | | |
| Brompheniramine | 7 | | 2 (29) | 5 (71) | 1 (14) | |
| Butaperazine | 15 | | 1 (7) | 12 (80) | 4 (27) | 2 (13) |
| Doxapram | 28 | | | 23 (82) | 7 (25) | 1 (3,5) |
| Fenfluramine | 45 | 7 (16) | 1 (2) | 32 (71) | 11 (24) | 16 (36) |
| Glycopyrrolate | 13 | 1 (8) | 3 (23) | 11 (85) | 3 (23) | 2 (15) |
| Metaxalone | 2 | | | 2 (100) | | |
| Methamphetamine | 135 | 15 (11) | | 75 (56) | 53 (39) | 25 (19) |
| Methenamine | 29 | 5 (17) | 2 (7) | 11 (38) | 12 (41) | 7 (24) |
| Methocarbamol | 11 | 1 (9) | 2 (18) | 7 (64) | 3 (27) | 2 (18) |
| Phenazopyridine | 16 | 1 (6) | 1 (6) | 10 (63) | 3 (19) | 1 (6) |
| Pheniramine | 12 | | | 8 (67) | 4 (33) | 1 (8) |
| Tybamate | 20 | 1 (5) | 3 (15) | 15 (75) | 2 (10) | 3 (15) |
| Total of all compounds | 333 | 31 (9) | 15 (5) | 211 (63) | 103 (31) | 60 (18) |
| **B. EIGHT-MONTH TABULATION** | | | | | | |
| Brompheniramine | 9 | | 2 (22) | 5 (55) | 3 (33) | 2 (22) |
| Butaperazine | 16 | | 1 (6) | 13 (81) | 5 (31) | 3 (18) |
| Doxapram | 33 | | | 26 (78) | 15 (45) | 1 (3) |
| Fenfluramine | 52 | 7 (13) | 1 (2) | 42 (80) | 27 (51) | 20 (38) |
| Glycopyrrolate | 17 | 1 (5) | 3 (17) | 14 (82) | 8 (47) | 3 (17) |
| Metaxalone | 2 | | | 2 (100) | | |
| Methamphetamine | 161 | 17 (10) | | 101 (62) | 71 (44) | 33 (20) |
| Methenamine | 41 | 6 (14) | 3 (7) | 22 (53) | 18 (43) | 10 (24) |
| Methocarbamol | 13 | 1 (7) | 2 (15) | 8 (61) | 9 (69) | 2 (15) |
| Phenazopyridine | 22 | 1 (4) | 1 (4) | 16 (72) | 10 (45) | 2 (9) |
| Pheniramine | 19 | 1 (5) | | 12 (63) | 6 (31) | 5 (26) |
| Tybamate | 21 | 1 (4) | 3 (14) | 16 (76) | 10 (47) | 4 (19) |
| Total of all compounds | 406 | 35 (8) | 16 (3) | 277 (68) | 182 (45) | 85 (21) |

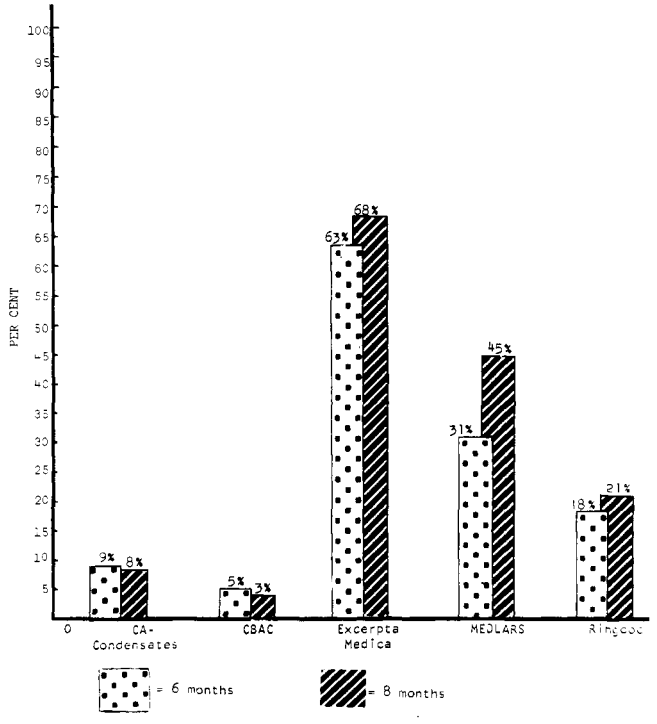


Figure 2. References cited—per cent of total

centages illustrated for *CA-Condensates* and *CBAC* may appear quite low relative to the other services, but this is a reflection of the type of materials selected in this study. *CA-Condensates* and *CBAC* do not abstract clinical articles which accounted for the major portion of cited references.

Unique or nonduplicated references from each source are shown in Table IIIA along with the total number of citations from all services. The very high percentages (in parentheses) of single-service retrieval citations in some searches such as those for fenfluramine, methamphetamine, doxapram, and tybamate are the result of difference in journal coverage, retrieval ability, and currency. The data here, like that in Table II, were tabulated at six months. Table IIIB extends the nonduplicated reference listings from six to eight months, and the totals of the two tabulations are compared with bar graphs (Figure 3). The decrease in per cent of unique references cited decreases here as expected, but should gradually level off.

For the chemical compound search where *MEDLARS* has neither a specific MeSH term nor a means of designating specific chemicals (see Table I, Group II), far greater differences were observed in the number of cited references than in the preceding tables. Table IV, A and C, shows the six-month results from three such compound searches. While *MEDLARS* output for these searches was usually very heavy, tabulated number of hits was very low because of the difficulty in determining the relevancy of hits. Unless the compound name appears in the title, a copy of the original article is needed to determine relevancy, and this is frequently difficult to obtain. As expected, the eight-

Table III. Total Number of Citations Found in Only One Service

| Compounds Searched | Total No. of References Cited Excluding Duplications | No. of references (and % of total) cited *only* by | | | | |
|---|---|---|---|---|---|---|
| | | *CA-Condensates* | *CBAC*[1] | *Excerpta Medica* | *MEDLARS* | *Ringdoc* |
| **A. SIX-MONTH TABULATION** | | | | | | |
| Brompheniramine | 7 | | 1 (14) | 4 (57) | 1 (14) | |
| Butaperazine | 15 | | 1 (7) | 8 (53) | 1 (7) | 1 (7) |
| Doxapram | 28 | | | 20 (71) | 4 (14) | 1 (4) |
| Fenfluramine | 45 | 1 (2) | 1 (2) | 16 (36) | 2 (4) | 8 (18) |
| Glycopyrrolate | 13 | | | 7 (54) | | 1 (8) |
| Metaxalone | 2 | | | 2 (100) | | |
| Methamphetamine | 135 | 7 (5) | | 56 (41) | 31 (23) | 11 (8) |
| Methenamine | 29 | 1 (3) | | 8 (28) | 9 (31) | 6 (21) |
| Methocarbamol | 11 | | | 6 (55) | 2 (18) | |
| Phenazopyridine | 16 | 1 (6) | 1 (6) | 10 (63) | 3 (19) | 1 (6) |
| Pheniramine | 12 | | | 7 (58) | 3 (25) | 1 (8) |
| Tybamate | 20 | | 2 (10) | 13 (65) | | 2 (10) |
| Total of all compounds | 333 | 10 (3) | 6 (2) | 157 (47) | 56 (17) | 32 (10) |
| **B. EIGHT-MONTH TABULATION** | | | | | | |
| Brompheniramine | 9 | | | 4 (44) | 2 (22) | 1 (11) |
| Butaperazine | 16 | | | 8 (50) | 1 (6) | 1 (6) |
| Doxapram | 33 | | | 18 (54) | 6 (18) | |
| Fenfluramine | 52 | 1 (2) | | 18 (34) | 2 (38) | 1 (2) |
| Glycopyrrolate | 17 | | | 8 (47) | 1 (5) | |
| Metaxalone | 2 | | | 2 (100) | | |
| Methamphetamine | 161 | 7 (4) | | 67 (41) | 32 (20) | 9 (5) |
| Methenamine | 41 | 2 (4) | | 16 (39) | 7 (17) | 4 (9) |
| Methocarbamol | 13 | | | 4 (30) | 3 (23) | |
| Phenazopyridine | 22 | | | 10 (45) | 3 (13) | 1 (4) |
| Pheniramine | 19 | 1 (5) | | 10 (52) | 1 (5) | 2 (10) |
| Tybamate | 21 | | | 9 (42) | | 1 (4) |
| Total of all compounds | 406 | 11 (2) | | 174 (43) | 58 (14) | 20 (5) |

[1]All references from CBAC had duplicates from other services.

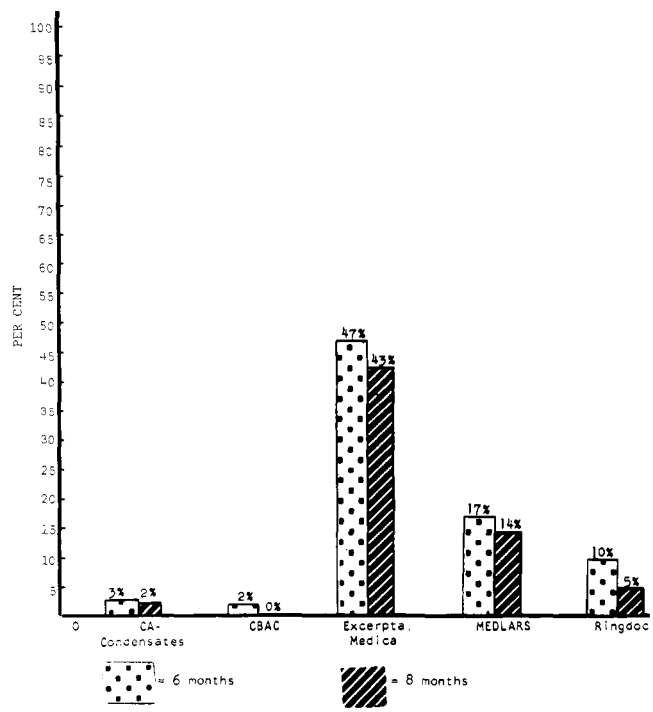


Figure 3. Unique references cited—per cent of total

month tabulations for these compounds shown in Table IV, B and D, were reasonably close to the equivalent data for the Group I compounds from all services except *MEDLARS* which was considerably lower in every case.

An additional set of records was maintained concurrently containing citations which were reported to the library by research personnel in the course of scanning journals of personal interest. Since these journals are major publications in their respective fields and are covered by one or more of the tape services, abstracts from them would be expected to appear in future tape printouts. Approximately 70–80% of the articles reported in 1970 have been received also from computerized tape services.

Probably the most difficult area in which to make any comparison is that of cost. Not only is the pricing schedule different for each service, but the outputs, depth of coverage, speed of abstracting, number of terms allowed per profile, etc., vary from one service to another. The initial selection of compounds is also an important controlling factor.

As a first approximation, a cost-per-hit basis might be one of the more realistic approaches, but one must realize that these results do not take into account items such as currency, ease of processing the particular computer output form, and completeness of coverage. An alternate method of comparison could be based on the search cost per product per month regardless of the output expected. The calculated values should not be considered absolute, but a relative assessment of the cost for employing computerized retrieval services.

Table V illustrates these methods of calculations.

Table IV. Total Number of Citations

| Compounds Searched | Total No. of References Cited Excluding Duplications | No. of References (and % of Total) Cited by: *CA-Condensates* | *CBAC* | *Excerpta Medica* | *MEDLARS* | *Ringdoc* |
|---|---|---|---|---|---|---|
| **A. SIX-MONTH TABULATION** | | | | | | |
| Flupenthixol | 20 | | 2 (10) | 14 (70) | 3 (15) | 3 (15) |
| INPEA | 27 | 5 (18) | 2 (7) | 10 (37) | 3 (11) | 9 (33) |
| Sulfametin | 26 | | 2 (7) | 24 (92) | 2 (7) | |
| Total of all compounds | 73 | 5 (7) | 6 (8) | 48 (65) | 8 (11) | 12 (16) |
| **B. EIGHT-MONTH TABULATION** | | | | | | |
| Flupenthixol | 23 | | 2 (8) | 17 (74) | 4 (17) | 2 (8) |
| INPEA | 32 | 5 (15) | 2 (6) | 18 (56) | 10 (31) | 9 (28) |
| Sulfametin | 37 | | 2 (5) | 36 (97) | 2 (5) | |
| Total of all compounds | 92 | 5 (5) | 6 (6) | 71 (77) | 16 (17) | 11 (12) |
| **C. SIX-MONTH TABULATION, REFS. CITED *ONLY* BY THESE SERVICES** | | | | | | |
| Flupenthixol | 20 | | 2 (10) | 12 (60) | 2 (10) | 2 (10) |
| INPEA | 27 | 4 (14) | 2 (7) | 10 (37) | 1 (3) | 8 (29) |
| Sulfametin | 26 | | 1 (3) | 22 (84) | 1 (3) | |
| Total of all compounds | 73 | 4 (5) | 5 (7) | 44 (60) | 4 (5) | 10 (13) |
| **D. EIGHT-MONTH TABULATION, REFS. CITED *ONLY* BY THESE SERVICES** | | | | | | |
| Flupenthixol | 23 | | 1 (4) | 14 (61) | 2 (8) | 2 (8) |
| INPEA | 32 | 4 (12) | 1 (3) | 13 (40) | | 3 (9) |
| Sulfametin | 37 | | | 33 (89) | 1 (2) | |
| Total of all compounds | 92 | 4 (4) | 2 (2) | 60 (65) | 3 (3) | 5 (5) |

Table V. Calculated Values for Employing Computerized Retrieval Services

| Service | Total No. of Hits | No. of Compounds per Profile | No. of Profiles | Cost for 6 months, $ | Cost per Hit, $ | Cost per Product per Month, $ |
|---|---|---|---|---|---|---|
| *CA-Condensates* (U. of Ga.) | 39 | 15 | 1 | 91 (odd issues only) | 2.34 | 1.00 |
| *CBAC* (U. of Ga.) | 15 | 15 | 1 | 65 | 4.34 | 0.73 |
| *Excerpta Medica* (3i) | 281 | 20 | 1 | 780 | 2.77 | 6.50 |
| *MEDLARS* (PMA) | 103 | 3 | 5 | 150 (400)[a] | 1.46 (3.88) | 1.67 |
| *Ringaoc* (A. H. Robins) | 60 | (15) | - | | 1.40 | 5.60 |

[a] A coding fee of $250 would have to be added to the price of the 5 profiles on the first months printout.

[b] On approximately 10% total use of $500.

## ACKNOWLEDGMENT

The authors gratefully acknowledge the valuable assistance provided by Levonne Young and Shirlee Farrar in compilation of data and preparation of manuscript.

## LITERATURE CITED

(1) Angstadt, Howard P., "Searching the Current Chemical Literature by Computer," *J. Chem. Doc.* **10,** 277–81 (1970).
(2) Hamilton, H. J., and Park, M. K., "Correlative Searching of Bibliographic and Chemical Structure Data Bases of Chemical-Biological Activities," *J. Chem. Doc.* **10,** 268–72 (1970).
(3) Park, M. K., Carmon, J. L., and Stearns, R. E., "The Development of a General Model for Estimating Computer Search Time for CA-Condensates," *J. Chem. Doc.* **10,** 282–4 (1970).
(4) Williams, Martha E., and Schipma, Peter B., "Design and Operation of a Computer Search Center for Chemical Information," *J. Chem. Doc.* **10,** 158–62 (1970).
(5) Zabriskie, Kenneth, H., Jr., and Lynch, Michael F., "Processing, Publishing, Storing, Correlating, and Retrieving Biochemical Information at Chemical Abstracts Service," *J. Chem. Doc.* **6,** 30–3 (1966).