J Forensic Sci, Jan. 1974, Vol. 19, No. 1

Leo Kazyak,[1] *B.S.*

Information Exchange and Computerized Data Retrieval for Toxicology

Computer applications in toxicology have been under development in our laboratory at the Walter Reed Army Institute of Research (WRAIR) during the past eight years, and our progress in this field has been reported periodically in two films and papers presented at scientific meetings [*1–5*]. Computerized retrieval of ultraviolet spectral data, infrared spectral data, and gas chromatographic data has been included in these presentations, and more recently the Registry of Human Toxicology, maintained by the toxicology section of the American Academy of Forensic Sciences (AAFS), has been programmed through our efforts so that computerized information files are now available for rapid access. As interest in computerization continues to grow, and computer systems are incorporated into laboratory operations (whether as dedicated instruments to control analytical equipment, or as information retrieval facilities provided by a data processing center on a time-sharing basis), the concept of a large information bank with rapid retrieval capability is most appealing to the researcher who must keep up with current developments, as well as to the toxicologist with an unusual case. The project described in this paper is an attempt to meet this requirement.

History

To coordinate computer applications in toxicology, a project intended to establish a computerized information bank was commenced at the Home Office Central Research Establishment (HOCRE) in Aldermaston, U.K. Under the sponsorship of Dr. Alan S. Curry, Director of HOCRE, this task was initiated by the author who had been granted a Secretary of Defense Research and Study Fellowship for that purpose. The object of a computerized information bank is to provide a simple and rapid means of information storage and retrieval that will be an inducement to data exchange. To achieve this capability at minimal cost, a small laboratory computer was used for the project. Thus, the system could be made available to any laboratory equipped with such a computer on a basis of participation in the information exchange network.

In order to reduce expensive duplication of effort, the project has been coordinated through the International Association of Forensic Toxicologists (TIAFT) with the cooperation of Dr. Curry and members of his staff. A combined effort will determine the basic format, computer programs, and operational procedures that will provide effective

Presented at the 25th Annual Meeting of the American Academy of Forensic Sciences, Las Vegas, Nev., 23 Feb. 1973. Received for publication 8 Feb. 1973; revised manuscript received 6 July 1973; accepted for publication 10 July 1973.

[1] Chief of Toxicology, Division of Biochemistry, Walter Reed Army Institute of Research, Walter Reed Army Medical Center, Washington, D.C. 20012.

information collection and exchange. Problems of computer equipment compatability are to be investigated after the system has become operational in Aldermaston and in Washington, with the information exchange between the Toxicology Section of WRAIR and Information Division of HOCRE to follow. However, a project of this magnitude requires considerable time to implement, especially in testing and improving computer programs that must operate efficiently and flawlessly. In this first phase of the project, file assignments and data formats are described that hopefully will merit further effort and consideration. Moreover, the computerized literature retrieval system, developed for HOCRE, is introduced to demonstrate what has been achieved in rapid search capability for the forensic sciences literature. Presumably that system, which now operates on a large computer, can be adapted to the small laboratory computer and become an important part of the information bank as well.

Computerized Literature Retrieval

To illustrate the advantage of data encoding when this approach can be utilized, the computerized literature retrieval system now in operation at the HOCRE laboratories is presented to familiarize anyone not acquainted with the many facets of automated data processing. The programs were developed on a large Burroughs 5500 time-sharing computer by James Lowe and Ian Clarke of the Home Office Scientific Research Branch of the Police National Computer Unit, and the system is now being emulated and transcribed for the minicomputer by Colin Brown at HOCRE.

The literature retrieval system is a simple and effective means of retrieving forensic sciences literature from a file of over 12,000 records. Initially, each record is distinguished by an accession number and descriptors or key words that indicate the nature and content of the record, that is, the published article. The accession number refers to the actual reproduction of the information on microfilm. The key words describe the record to the extent that it can be identified with topics, procedures, or data pertinent to the forensic sciences. Later, when these same key words are used in an inquiry, the record is retrieved (Fig. 1). The key words are selected from a thesaurus of descriptors that are common to the forensic sciences.

Prime number encoding is the basis of this system and has enabled the file search to be completed for each inquiry in a very short time, usually in only a few seconds. Briefly, prime number encoding means that each key word in the thesaurus has been assigned a different prime number, and each record in the file is in fact the product of prime numbers associated with those key words (Fig. 2). Identification with a record is effected when a prime number that represents a given key word can be divided evenly (that is, without a remainder) into the product of the primes that correspond to that particular record.

Conversion of key words to their assigned prime numbers is accomplished by the computer as soon as the key word has been entered. Additionally, another computer program converts key-worded records to a product of the corresponding primes, and assembles them into the file together with the accession number associated with the record.

With the system now in full operation, Home Counties Forensic Science Laboratories throughout the United Kingdom can be assisted quickly and easily whenever problems arise that require a search of the literature. Moreover, keeping the system updated requires very little effort beyond the task of microfilming and selecting key words for those publications relevant to the forensic sciences.

```
   2024  600
17.14                [redacted] CENTRAL RESEARCH ESTB.
                        LITERATURE RETRIEVAL SYSTEM

SNF 85500 15
SNF 85500 15FORENS

AND YOUR PASSWORD
[redacted]

02/02/72  5:19 PM.
GOOD AFTERNOON,  HELLO FORENS   YOU HAVE STATION 15

£
LIST OUTPUT


FILE:OUTPUT -02/02/72  5:19 PM.

 RUN CREPE/CRE
 RUNNING


DO YOU WANT GUIDANCE (YES/NO)
? NO GUIDANCE
0=
? TARGET IS 05144
? SPECTRO
? IR
? SPECTRA
? POLYMER
? /NEXT
1=
? TARGET IS 04950
? DIAZEPAM
? METHOD
? BLOOD
? HUMAN
? SPECTRO
? UV
? LEVEL
? THERAP
? /NEXT
2=
? TARGET IS 07820
? LEVEL
? IDENT
? POLYAMINES
? TLC
? SEMEN
? /SEEK
RECORDS SCANNED= 007891, TIME TAKEN= 00007598

000009 RESPONSES TO ENQUIRY 0:   TARGET IS 05144
   01758   05048   05054   05078   05098   05125   05144   05211
   10328

000001 RESPONSES TO ENQUIRY 1:   TARGET IS 04950
   04950

000001 RESPONSES TO ENQUIRY 2:   TARGET IS 07820
   07820

 END CREPE 2 MIN, 19.3 SEC.


BYE
 EXECUTE 2 MIN, 19.3 SEC.
 GOODBYE FORENS
02/02/72









£
BYE

 ON FOR  2 MIN, 44.1 SEC.
 EXECUTE .0 SEC.
 OFF AT  5:21 PM.
 GOODBYE FORENS
02/02/72
```

FIG. 1—*Example of an inquiry and the response of the computer (HOCRE Literature Retrieval System).*

KEYWORDS	PRIME NO. CODE
AMPHETAMINE	113
HUMAN	7919
PRELUDIN	22051
METHOD	17189
URINE	31991

113 X 17189 X 7919 X 22051 X 31991 = 10,850,643,705,281,008,903

8053

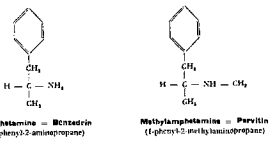
TOXICOLOGICAL ANALYSIS OF WECKAMINES (AMPHETAMINE, PERVITIN, PRELUDIN AND RITALIN) IN PHARMACEUTICAL COMPOUNDS AND URINE OF PERSONS SUSPECTED FROM DOPING

by

A. Heyndrickx and A. De Leenheer,
Department of Toxicology
State University of Ghent (Belgium)

INTRODUCTION

As in many laboratories we are often dealing with the problem of the identification of weckamines, as well in pharmaceutical compounds as in urine samples. These weckamines form a large group of compounds with a strong analogy in their chemical structure and properties. Only two of them are subcultted to the « opium law » in Belgium : Amphetamine and Methylamphetamine. These two compounds are considered by belgian court as doping agents. The toxicologist needs therefore, together with a method of group detection, one or several high specific identification procedures.



Amphetamine = Benzedrin (1-phenyl-2-aminopropane)

Methylamphetamine = Pervitin (1-phenyl-2-methylaminopropane)

FIG. 2—*Keyworded excerpt from the literature and conversion of encoded prime numbers to a product that serves as a record for storage in the computer.*

Toxicology Data Bank

Information sources, such as those maintained by HOCRE and WRAIR, were considered adequate to provide data for the initial files of the data bank. Nevertheless, the objective of this project was to promote exchange of data and information by as many laboratories as possible. Assuming that most forensic sciences laboratories will eventually acquire small, laboratory-type computers to operate automated equipment, they would then have direct access to the information bank as soon as programs and files could be incorporated into their systems. The incentive to supplement the files would be much greater since the benefits would be immediate. Thus, the usefulness of such information and the ease of file management, as well as the speed of record or data retrieval, would encourage cooperation among laboratories to the extent that the entire operation ought to be self-sustaining rather than dependent on the efforts and facilities of a single person or laboratory.

Perhaps the greatest problem in the design of an information bank is to determine what must go into it. Everyone is not in agreement as to what constitutes important or useful information. However, physical characteristics of compounds such as spectral data are now utilized extensively by most laboratories, and programs and files to serve these requirements have been modeled after the ultraviolet spectra identification program used in our laboratory for over five years. Separate files were designated for each collection of spectral data so that they could be searched independently as well as collectively within the contexture of the system. This concept is important because the function of most laboratory computers is to operate with other equipment, and for this reason they are frequently purchased as a component or an accessory interfaced to a specific analytical instrument such as a spectrophotometer or gas chromatograph. Hence, files that can be

used in conjunction with such apparatus would serve two functions and thereby increase their utility and importance.

From these preliminary considerations involving the spectral files emerged the design for the information bank. To keep the system flexible as well as useful, it had to be a composite of files that could be multifunctional. One major disadvantage was that such a system excluded any type of encoding that might improve file compression and reduce retrieval time, such as the prime-number encoded Literature Retrieval System of HOCRE. If the files were to be used individually for specific applications in dedicated equipment systems, encoding and decoding programs would have added another dimension to a very limited computer capacity, and would have unduly complicated the processing of data. Therefore, the system is a composite of eight individual files maintained on magnetic disk or tape, and designated as follows: (1) ultraviolet spectra, (2) mass spectra, (3) infrared spectra, (4) miscellaneous physical characteristics of compounds (for example, melting points, solubilities, gas chromatography retention indices, and normal dosage range and route of administration for drugs), (5) drug distribution tables (that is, distribution of drugs in body fluids and tissues compiled from postmortem reports and studies conducted with therapeutic doses), (6) selected literature references to procedures, (7) reported synonyms and trade names of drugs and compounds, and (8) accepted drug or compound names (based on the name ascribed to a compound and listed in the *Merck Index of Chemicals and Drugs*). Again, in the interest of flexibility, the system is by no means limited to eight files, and could conceivably be reduced or increased as use and requirements dictate. Also, programs can be written to suit individual needs, and the file data can be utilized for a particular application.

In general, the system is intended to provide a profile on every compound included. To retrieve this information, the name of a substance need only be entered into the computer together with a call for the general search program. If the name cannot be found in the accepted name index (File 8), the synonym file (File 7) is scanned in an attempt to locate the compound in question. Once the name has been found in either of the files it will be identified with the file number for that compound, and the remaining files will be searched for all data and information contained therein. The name and file number of the compound and all names synonymously associated with the compound will be printed (Figs. 3 and 4) as well as miscellaneous physical characteristics; ultraviolet, infrared, and mass spectra; drug distribution in body fluids and organs; and references to analytical methodology.

Examples of the ultraviolet spectral identification program (Fig. 5) and the mass spectral identification program (Fig. 6) have been included to demonstrate the use of specific files for individual applications. Other programs that will be available include the identification of compounds by gas chromatography retention indices and statistical evaluations of drug distribution data. Also, a ninth file will be added to correlate the *Chemical Abstracts Numerical Index of Chemical Compounds* with the system.

Programs, files, and the interpass program loader, which transfers programs on the tape to the computer, reside on a single magnetic tape. With provision for approximately 1000 records in each file, the eight files occupy half of a 2400-ft reel of tape. Therefore, the capacity of a single tape would be about 16,000 records. The main objection to the use of magnetic tape is the slow access and comparatively longer time of retrieval due to the serial search process. Computers with a magnetic disk peripheral can execute the programs in a fraction of the time, and the system has been adapted for disk operation.

```
:RUN,SEEK
AMOBARBITAL

NO.      1 AMOBARBITAL

SYN.:  5-ETHYL-5-ISOAMYLBARBITURIC ACID=BARBAMYL=BARBAMIL=AMYLOBAR-
       BITONE=5-ETHYL-5-ISOPENTYLBARBITURIC ACID=SOMNAL=DORMYTAL=
       5-ISOAMY-5-ETHYLBARBITURIC ACID=PENTYMAL=ISOMYTAL=EUNOCTAL=
       AMAL=MYLODORM=SEDNOTIC=AMYTAL=

     MELTING POINT: 156.8- 158.8 C.

                          SOLUBILITIES (MG./ML.)
 CNCL(3)     N(2)O     ETHANOL      ETHER       OTHER
   50.000      .873    200.000    160.000  9999.998 BENZENE
          (9999.99 INDICATES >10 G./ML.)

THERAPEUTIC DOSE:     30 TO   200 MG. ORALLY

           GLC RETENTION INDICES

   OV-1/SE-30   OV-17      QF-1
         1725      1985       2350          0          0          0

          0

           U.V. SPECTRA

        21.0 UG./ML.
AMOBARBITAL--IN 0.45 N KOH
 .789 P  254 .364 V  227

        21.0 UG./ML.
AMOBARBITAL--IN BUFFER PH 10.5
 .963 P  239

        14.0 UG./ML.
AMOBARBITAL--DIFFERENCE SPECTRUM
 .687 P  259 .000 V  237

           MASS SPECTRUM

AMOBARBITAL--GLC INLET
MASS = 226.27  BASE PEAK AT M/E  156
M/E          41    43    55    98   141   142   156 B 157   183   197
REL. INT.    18%   16%   18%   19%   62%   12%  100%   30%    9%   11%
     NO MOLECULAR ION

           I.R. SPECTRUM
 45                        STRONGEST BAND =  5.98 MICRONS
AMOBARBITAL--KBR DISC
 3.10 P  28%  3.15 V  35%  3.28 P  31%  3.38 V  96%  3.48 P  48%
 4.78 V  83%  5.78 P  16%  5.75 V  27%  5.88 P   9%  5.95 V  13%
 6.58 P   8%  6.48 V  88%  6.58 P  76%  6.65 V  88%  7.08 P  18%
 7.23 V  55%  7.38 P  34%  7.35 V  53%  7.48 P  34%  7.58 V  46%
 7.75 P  41%  7.68 V  46%  7.63 P  25%  7.98 V  73%  7.98 P  63%
 8.08 V  73%  8.18 P  42%  8.28 V  74%  8.38 P  54%  8.58 V  84%
 8.65 P  73%  9.08 V  89%  9.35 P  85%  9.45 V  88%  9.68 P  77%
10.08 V  92% 11.75 P  57% 12.18 V  66% 12.38 P  63% 13.18 V  84%
13.28 P  76% 13.48 V  88% 13.58 P  79% 14.78 V  96% 14.98 P  82%
```

FIG. 3—*First part of data profile on amobarbital.*

```
                    DRUG DISTRIBUTION DATA

                    A. POST MORTEM                    B. THERAPEUTIC
                    HIGH     MEAN     LOW      HIGH     MEAN     LOW
BLOOD               9.58     3.28      .54     2.58     2.58     2.58
URINE                .00      .00      .00     2.00     2.00     2.00
BRAIN                .00     4.60      .20      .00      .00      .00
LIVER                .00    12.58     1.00      .00      .00      .00
KIDNEY               .00      .00      .00      .00      .00      .00
SPLEEN               .00      .00      .00      .00      .00      .00
STOMACH CONTENT   159.00   541.00      .00      .00      .00      .00
BILE                 .00      .00      .00      .00      .00      .00
LUNG                 .00      .00      .00      .00      .00      .00
MUSCLE               .00      .00      .00      .00      .00      .00
TECHNIQUE              2        2        2        2        2        2
VICTIM                 2        0        2        2        0        0
                                                 DOSE:  100 MG.

    WITH                                              WITH
SECOBARBITAL                                      MEPROBAMATE
                    HIGH     MEAN     LOW      HIGH     MEAN     LOW
BLOOD               4.98     2.00     1.10     3.38     3.38     3.38
URINE                .00      .00      .00      .00      .00      .00
BRAIN                .00      .00      .00      .00      .00      .00
LIVER                .00      .00      .00     5.38     5.38     5.38
KIDNEY               .00      .00      .00      .00      .00      .00
SPLEEN               .00      .00      .00      .00      .00      .00
STOMACH CONTENT      .00      .00    71.00      .00      .00      .00
BILE                 .00      .00      .00      .00      .00      .00
LUNG                 .00      .00      .00      .00      .00      .00
MUSCLE               .00      .00      .00      .00      .00      .00
TECHNIQUE              7        7        7        7        7        7
VICTIM                 2        0        3        2        0        2
(STOMACH CONTENT IN MG.-TOTAL; OTHER DATA IN MG.-%)
VICTIM: CHILD(1), WOMAN(2), MAN(3)
TECHNIQUE: ATOM. ABS.(1)   U.V.(2)   I.R.(3)   FLUORES.(4)
           COLOR/COLORIMETRIC(5)   M.S.(6)   GLC(7)   TLC(8)   ENZYME(9)
           CRYSTAL(10)   MANOMET.(11)   EMISSION SPECT.(12)

              SELECTED LITERATURE REFERENCES FOR METHODS

GOLDBAUM, L.R., ANAL. CHEM.,24,1684,(1952)    (U.V. METHOD)

DE ZEEUW, PHARM. WEEKBLAD., 101,969, (1966)

WALDI, D., MED. MONATSSPIEGEL, NO.4,94, (1966)

COUTTS, R.T., J. PHARM. SCI., 57/12, 2096, (1968)

STEWART, J.T., ANAL. LETT., 2/8, 449, (1969)

GROVE, J., CLIN. CHEM. ACTA, 29, 253, (1970)

SINE, H.E., CLIN. CHEM., 16/7,587, (1970)
```

FIG. 4—*Second part of data profile on amobarbital.*

```
CASE NO. 1000  DEMONSTRATION EXAMPLE                    11/20/'72
                        SPECTRUM OF THE UNKNOWN
  297 P .533  277 V .200  250 P .735  246 V .730

PEAK RATIO =    .45
                        IDENTIFICATION

SPECTRUM NO.     8
MORPHINE--0.1 N KOH
  297 P   277 V   250 P   246 V

PEAK RATIO =    .45
     CONCENTRATION =   42.22 UG./ML. URINE

                    ***END OF SEARCH***


CASE NO. 1001  2ND DEMONSTRATION EXAMPLE                11/20/'72
                        SPECTRUM OF THE UNKNOWN
  254 P .709  227 V .364
                        IDENTIFICATION

SPECTRUM NO.     1
AMOBARBITAL--IN 0.45 N KOH
  254 P   227 V
     CONCENTRATION =   23.33 UG./ML. BLOOD

                    ***END OF SEARCH***


                    ***END OF SEARCH***

  UVIDT : STOP    0000
```

FIG. 5—*Computer output from program for ultraviolet spectra indentification.*

```
CASE NO. 1000  DEMONSTRATION OF MASS SPECTRA SEARCH     11/20/'72


                        SPECTRUM OF THE 'UNKNOWN'
M/E          41   43   55   98   141   142   156 B 157   183   197
REL. INT.    18%  16%  18%  10%  62%   12%  100%  30%   9%   11%
BASE PEAK AT M/E:  156

                        IDENTIFICATION

SPECTRUM NO.     1
AMOBARBITAL--GLC INLET
MASS=  226.27  BASE PEAK AT M/E: 156
M/E          41   43   55   98   141   142   156 B 157   183   197
REL. INT.    18%  16%  18%  12%  62%   12%  100%  30%   9%   11%

                    ***END OF SEARCH***


9 10
41 55 69 83 98 112 141 155 156 157
18 19 11 15 20 15 90 21 100 10
ENTER '0' FOR SEARCH ONLY OR '1' FOR CASE RECORD, FOLLOWED BY NO. OF
PEAKS TO BE MATCHED
0 9


                        SPECTRUM OF THE 'UNKNOWN'
M/E          41   55   69   83   98   112   141   155   156 B 157
REL. INT.    18%  19%  11%  15%  20%  19%   90%   21%  100%  10%
BASE PEAK AT M/E:  156

                        IDENTIFICATION

SPECTRUM NO.     4
BARBITAL
MASS=  184.19  BASE PEAK AT M/E: 156
M/E          41   55   69   83   98   112   141   155   156 B 157
REL. INT.    18%  19%  11%  15%  20%  19%   90%   23%  100%  12%

                    ***END OF SEARCH***


                    ***END OF SEARCH***
```

FIG. 6—*Computer output from program for mass spectra identification.*

Conclusion

The demand for information increases with the rapid advances in scientific technology, and with this advancement and the development of more sophisticated equipment data often accumulates faster than it can be evaluated and organized for reference. Obviously, computers are the answer to this problem, but program and systems development can be costly. Any duplication of effort can serve only to increase the expense. If the format and design of a general file for toxicology can be established soon enough, the problem of information exchange will be minimal and the cost of maintaining an information bank will be insignificant. With cooperation and a combined effort, much can be achieved for the mutual benefit of those who are willing to participate actively. Hopefully, this cooperation will come nationally from the membership of the American Academy of Forensic Sciences, and internationally from The International Association of Forensic Toxicologists.

References

[1] Kazyak, L., "Computerization in Analytical Toxicology," film presented at the Fourth International Meeting In Forensic Medicine, Copenhagen, Denmark, 17 Aug. 1966.
[2] Kazyak, L., "An Automation and EDP Approach to Analytical Toxicology," presented at the Eastern Analytical Symposium, New York, 14 Nov. 1968.
[3] Kazyak, L., "Laboratory Automation in Analytical Toxicology," film presented at the 21st Annual Meeting of the American Academy of Forensic Sciences, Chicago, Ill., 27 Feb. 1969.
[4] Kazyak, L., "Laboratory Automation in Analytical Toxicology," presented at the Symposium on Recent Developments in Research Methods and Instrumentation, National Institutes of Health, Bethesda, Md., 5 Oct. 1970.
[5] Curry, A. S., and Kazyak, L., "Computerized Information Retrieval for Toxicology," presented at the Sixth International Meeting of Forensic Sciences, Edinburgh, Scotland, 26 Sept. 1972.

Division of Biochemistry
Walter Reed Army Institute of Research
Walter Reed Army Medical Center
Washington, D.C. 20012