

Computerunterstützte Identifizierung basischer Substanzen*

Günter Schmidt und Ernst Klug

Institut für Rechtsmedizin der Freien Universität Berlin, Hittorfstr. 18, D-1000 Berlin 33

Computer-Supported Identification of Basic Substances

Summary. The potential applications of a personal computer with integrated database software in toxicologic analysis are discussed. A toxicologic database including 370 basic compounds, is presented that was built up without knowledge of a higher programming language and was used as a base for drug identification in normal analysis by PC, DC and GC.

Key words: Computer, analytic toxicology – Data management, identification – Toxicology, computer-supported

Zusammenfassung. Die Möglichkeiten eines Personal-Computers mit installierter Datenbanksoftware in der chemisch-toxikologischen Analytik werden aufgezeigt. Es wird eine toxikologische Datenbank mit 370 basischen Wirkstoffen vorgestellt, die ohne Kenntnis einer höheren Programmiersprache erstellt wurde und als Grundlage für die Identifizierung dieser Wirkstoffe in der routinemäßigen Analytik mittels PC, DC and GC dient.

Schlüsselwörter: EDV, analytische Toxikologie – Data-Management, Identifizierung unbekannter Fremdstoffe – Toxikologie, Computereinsatz

Die Identifizierung einer oder mehrerer unbekannter Substanzen in einem biologischen Material gehört zu den grundlegenden Tätigkeiten in einem forensisch-toxikologischen Laboratorium. Bei den fast immer vorliegenden niedrigen Konzentrationen, wie sie nach parenteraler Aufnahme die Regel sind, wird man in erster Linie chromatographische Verfahren heranziehen, wobei die Kombination GC-MS unter bestimmten Umständen die sichersten Aussagen erlaubt.

Todesfälle durch saure bzw. neutrale Verbindungen sind verhältnismäßig einfach aufzuklären. Umfangreiche Statistiken aus England und Wales [1]

* Auszugsweiser Vortrag Arbeitstagung Nord- und Westdeutscher Rechtsmediziner, Aachen, Mai 1986

Sonderdruckanfragen an: G. Schmidt (Adresse siehe oben)

haben ergeben, daß von 2122 Todesfällen durch solche Verbindungen 98% auf 8 leicht zu identifizierende Wirkstoffe zurückzuführen waren (Seco-, Amo-, Pento-, Buto- und Phenobarbital, Acetylsalicylsäure, Paracetamol und Glutethimid). Dies entspricht in etwa auch unseren Erfahrungen. Ein anderes Bild geben die sogenannten basischen Arzneistoffe: an unserem Institut wurden zwischen 1981 und 1986 insgesamt 218 Vergiftungen durch solche Mittel bearbeitet. Neben 142 Vergiftungen durch Morphin/Heroin wurden 76 Intoxikationen durch 25 weitere basische Substanzen aufgeklärt. Dazu kam noch eine Vielzahl, meist nur im subtherapeutischen Bereich vorliegender Wirkstoffe, die alle jedoch eine genauere Identifizierung erforderlich machten.

Voraussetzung für eine erfolgreiche Identifizierung ist die Entwicklung einer genügend großen Anzahl nur schwach korrelierender Chromatographie-Systeme. Connors [2] hat 1974 gezeigt, daß im Idealfall zur Beurteilung von 200 Substanzen 5 DC-Systeme erforderlich erscheinen.

In unserem Laboratorium stehen aus eigenen Untersuchungen hRf-Werte aus einem PC-, 5 DC- sowie Retentionsindizes aus einem GC-System zur Verfügung. Die korrigierten Rf-Werte wurden durch Mehrfachbestimmung unter gleichzeitigem Einsatz von über das ganze Feld verlaufenden Testreihen nach Vidic und Klug [3] ermittelt, die Retentionszeiten im wesentlichen nach den von Berninger und Möller 1977 [4] angegebenen Bedingungen.

Bei der Analyse gehen wir so vor, daß neben einem gaschromatographischen Screening auf leicht flüchtige Stoffe zunächst dünnschichtchromatographisch überprüft wird, ob exogene Substanzen vorhanden sind. Eventuelle Flecke werden dann eluiert, in weiteren Systemen chromatographiert sowie gaschromatographisch untersucht. Falsche Zuordnungen beim Vorliegen mehrerer Verbindungen können so weitgehend vermieden werden.

Die Auswertung wurde früher manuell aus Tabellen durchgeführt, danach führte Vidic [3] ein Kennnummernsystem ein, das wiederum durch die Verwendung von Schlitzlochkarten verbessert wurde. Unter Berücksichtigung der Fehlerbreite und Toleranzen konnte so jede eingegebene Substanz unter beliebigem Einsatz ermittelter analytischer Daten identifiziert werden. Die Auswertung gestaltete sich allerdings umständlich und zeitaufwendig.

Hier bietet sich als Alternative der Computer an, der ja seit einigen Jahren zur Information und Dokumentation auch in der toxikologischen Chemie verwendet wird. 1983 ersetzte Schepers et al. [5] das bisher übliche starre „Suchfenster“, das lediglich eine Auflistung der möglichen Verbindungen erlaubt, durch ein statistisches Suchkonzept (MLL = mean list length), das von Franke et al. 1985 [6] zu einem Retrieval-System weiterentwickelt wurde. 1984 haben Maier und Derksen [7] ein etwa 600 Verbindungen umfassendes System vorgestellt, das eine schnelle Zuordnung und somit sichere Identifizierung einer unbekanntenen Substanz durch den Computer erlaubt.

Verwirklicht wurde dies durch ein selbstentwickeltes Programm in der Programmiersprache Fortran, das über ein Institutsterminal per Modem mit dem Großrechner der Klinik abrufbar war. Dies war notwendig wegen der erforderlichen Speicherkapazität. Ein ähnliches System wurde von Gill et al. 1985 [8] in Großbritannien entwickelt. Nachteilig ist hier die verhältnismäßig hohe Rechenzeit, als deren unmittelbare Folge die CPU-Zeit des Großrechners stark

Suchen Vorw. Rueckw. Korrigieren Loeschen Neuaufnahmen Tabelle Format
 Master Detail Aktuell PRINT HELP END (Tabelle 1: t1)

< 125> INN:<DIPHENHYDRAMIN > Triv.:< >
 hRf: J1<70> J2< 0> sDC< 0> D1<54> D3<61> D5<84> D6<53> D7<76>
 Färbung: Drag<+ > p-Na<- > Nin<- > BrKG<++ > H2SO4< ? >
 Fluoresz.: PC<- > DC<- > Elektroph.: Ret<100>
 Flüchtigkeit: <- > Extration < b >
 GC1 <1860> GC2 < 0> HPLC < 0>
 weitere Informationen: <j> --> Taste d drücken

Durch Eingabe des Anfangsbuchstabens einen Befehl auswahlen
 Gefundene Saeetze: 1

Abb. 1a. Hauptformatmaske

Suchen Vorw. Rueckw. Korrigieren Loeschen Neuaufnahmen Tabelle Format
 Master Detail Aktuell PRINT HELP END (Tabelle 2: t2)

Codenummer <125>
 Kommentar: <Abbauprodukte siehe J-Ordner IV, Anfärbung des Abbauproduktes >
 < mit Pervitin R. blau. >
 < >
 < >
 Literatur: <Vycudilik, W., Pollak, St.: Nachweis von Diphenhydramin. Z Rechts- >
 <medizin 95, 129 (1985). Literaturübersicht. >
 Substanzinformationen: Taste m drücken
 Leichenbefunde: Taste d drücken

Durch Eingabe des Anfangsbuchstabens einen Befehl auswahlen
 Gefundene Saeetze: 1

Abb. 1b. Zusatzformatmaske

beansprucht wird. Ein sehr umfangreiches Retrievalsystem für die toxikologische Analytik, das neben Substanzdaten auch Literatur- und Falldaten beinhaltet, wurde von Battista et al. [9] 1985 auf der Basis des Informations-Vermittlungs- und Verarbeitungs-Systems (IV + V) vorgestellt, ein kompiliertes Pascal-Programm, das auf einem Personal-Computer ablauffähig ist. Kommerzielle Datenbanksoftware (dBaseII), auf einem PC installiert, wurde erstmals von Wolf und Reiter [10] 1986 eingesetzt, jedoch gab es durch die Kapazität der Datenträger starke Einschränkungen.

In unserem Institut steht als Hardware ein 32-Bit-Mehrplatz-PC Siemens MX2 mit Gleitkomma-Prozessor zur Verfügung, dessen Betriebssystem SINIX V 2.0 (XENIX, eine Weiterentwicklung von UNIX III) auf der Programmiersprache „C“ beruht. Als Softwarepaket kommt das Datenbank-Entwicklungs-

ACEGO SQL/ACE Steuerprogramm
 Copyright (C) 1981-1985 Relational Database Systems, Inc.
 SIEMENS-IND0031246

INFORMIX Version 2.00

```

Bitte eingeben:   hRf J1 (**, unbekannt = >return<)          >>
                  hRf D1 (**, unbekannt = >return<)          >>
                  hRf D3 (**, unbekannt = >return<)          >># 60
                  hRf D5 (**, unbekannt = >return<)          >>
                  hRf D6 (**, unbekannt = >return<)          >>
                  hRf D7 (**, unbekannt = >return<)          >>
GC1 (***, nicht messbar = -1, unbekannt = >return<)        >>#1870
  
```

Abb. 2a. Eingabeparameter des Suchprogramms

!Gesuchte Verbindung:		00	00	60	00	00	00	1870
!Code	INN-Name	J1	D1	D3	D5	D6	D7	GC1
0,1596!	3 DOXYLAMIN	34	23	56	77	43	74	1915
0,2700!	108 ISOAMINIL	74	51	55	80	54	75	1830
0,3024!	13 2-(N-(2-DIETHYLAMINOETHYL))	32	24	52	71	30	66	1890
0,4641!	266 ETIFELMIN	83	54	53	78	60	80	1885
0,5376!	245 MELPERON	80	63	64	85	61	86	1900
0,9504!	125 DIPHENHYDRAMIN	70	54	61	84	53	76	1860

Programm beendet.

Abb. 2b. Suchprogramm-Ausdruck

system INFORMIX [11] zum Einsatz, das es gestattet, ohne Kenntnisse einer höheren Programmiersprache eine Datenbank aufzubauen, zu verwalten und abzufragen.

Der Aufbau der Datenbank erfolgt in einer leicht erlernbaren Kommandosprache, indem zunächst die benötigten Eingabevariablen definiert werden. Der Rechner erzeugt daraus eine Standard-Eingabemaske für die Datentabellen, ein Format, mit dem bereits wie mit einer Karteikarte gearbeitet werden kann. Neben Datensatz-Eingabe und -Änderung können wie bei einer Karteikarte Suchvorgänge durchgeführt werden, jedoch können hier bereits wesentlich komplexere Sachverhalte bearbeitet werden, da gleichzeitig nach mehreren Parametern in beliebig vorgegebenen Intervallen gesucht werden kann, wobei auch Daten aus mehreren Tabellen miteinander verknüpft werden können. Über das Standard-Format hinaus haben wir die in den Abb. 1a und 1b am Beispiel des Diphenhydramins wiedergegebenen Bildschirmformate erstellt, wobei die Datentabelle bereits Eingabemöglichkeiten für weitere Untersuchungssysteme vorsieht.

Unsere Datensammlung enthält derzeit mehr als 350 relevante basische Verbindungen. Neben den hRf-Werten und der Anfärbbarkeit mit verschiedenen gebräuchlichen Reagentien wurde das Extraktionsverhalten in dem bei uns durchgeführten Trennungsgang, Fluoreszenzfarben und Flüchtigkeit als bei jeder Analyse anfallenden Grunddaten ermittelt, wobei besonderer Wert auf

die Vollständigkeit der Datensätze gelegt wurde. Physikalische Werte, wie UV- und IR-Maxima, Schmelzpunkte usw. wurden nicht aufgenommen, da diese Parameter wegen der meist recht geringen Konzentration der nachzuweisenden Verbindungen bei uns keine besondere Rolle bei der Identifizierung spielen und sie bei Bedarf leicht in einschlägigen Tabellen nachzuschlagen sind. Literaturangaben zur Analytik, zum Metabolismus und zur Kasuistik sind jedoch vorgesehen. Diese Angaben erfolgen in Klartext, sind also sehr speicherplatzintensiv. Da sie nicht für alle Verbindungen notwendig sind, wurde dafür eine separate Tabelle angelegt.

Die weitaus wichtigste Einsatzmöglichkeit der Datenbank besteht jedoch in der Mithilfe bei der Substanzidentifizierung. Nach der Eingabe von einem bis zu sieben analytischen Meßwerten liefert der Rechner eine Liste der in Frage kommenden Substanzen, geordnet nach zunehmender Wahrscheinlichkeit, die zusammen mit den eingegebenen Werten eine rasche Identifizierung ermöglicht (Abb. 2a, 2b). Voraussetzung ist – wie bei jedem Retrievalsystem – daß die gesuchte Verbindung überhaupt gespeichert ist. Unvollständige Datensätze werden hingegen berücksichtigt. Kommen mehrere Verbindungen mit gleicher Wahrscheinlichkeit in Frage, kann aus der Angabe der Parameter der noch nicht durchgeführten Trennsysteme leicht das optimale, die endgültige Identifizierung ermöglichende System ersehen werden.

Die Erstellung des Listenprogramms wird durch das Datenbanksoftwarepaket (Listengenerator) erleichtert. Das Programm ist so strukturiert, daß es nach Eingabe der gewünschten Parameter die möglichen Verbindungen unter Berücksichtigung der analytischen Streubreite heraus sucht. Definiert man für die Abweichung vom Sollwert des Einzelsystems den Diskrepanzindex DI_i mit

$$DI_i = 1 - \frac{(X_i - x_i)^2}{\sigma^2}$$

wobei $(X_i - x_i)^2$ das Quadrat der Differenz des ermittelten Sollwertes X_i und x_i der aktuelle Meßwert im System i bedeutet sowie σ^2 die Standardabweichung des Systems, dann resultiert die wahrscheinlichste Verbindung aus dem Maximum des Produktes der einzelnen Diskrepanzindizes. Existiert kein aktueller Meßwert (das System ist nicht untersucht worden) oder ist kein Sollwert bekannt, so wird DI_i des betreffenden Systems i nicht berücksichtigt, indem auf 1 gesetzt wird. Letzterer Fall tritt in unserer Datensammlung nur sehr selten auf.

Die Standardabweichung bei der chromatographischen Untersuchung biologischer Proben ist jedoch nicht mit der Standardabweichung des Meßwertes der Reinsubstanz im jeweiligen System identisch. Der Matrixeffekt ist vielmehr zusätzlich zu berücksichtigen [12]. Daher ist die Standardabweichung nicht exakt zu ermitteln, so daß auf empirische Werte zurückgegriffen werden muß. Für unsere PC- und DC-Systeme haben wir die Standardabweichung mit ± 10 hRf-Einheiten vorgegeben, für die Gaschromatographie mit ± 50 für den Indexwert. Diese verhältnismäßig groß gehaltenen Fenster ermöglichen es, bei vertretbaren Rechenzeiten alle in Frage kommenden Verbindungen zu erfassen. Der Einsatz der verschiedenen, nur wenig korrelierenden Laufmittel eliminiert den größten Teil der in einem System ausgeworfenen Verbindungen. Die Wertung der verbliebenen in Betracht kommenden Verbindungen gestattet dann die eindeutige Identifizierung.

Gegenwärtig sind Untersuchungen im Gange, anhand der gespeicherten Daten die Leistungsfähigkeit und Korrelation der einzelnen Systeme statistisch zu untersuchen. Hier leistet der seitens der Datenbanksoftware her ermöglichte direkte Dialog Hilfestellung, da hier beliebige Fragestellungen on line untersucht werden können. Die Anwendungen sind so zahlreich, daß der weitere Ausbau der Datenbank gewissermaßen vorprogrammiert ist.

Die gesamte Datenbank benötigt lediglich 250 kB Speicherplatz (davon 41 kB für Programme und Datenbankverwaltung), ist also auf einer nicht einmal zur Hälfte gefüllten Diskette unterzubringen. Pro Datensatz werden lediglich 117 Bytes benötigt, für zusätzliche Angaben zu Analytik, Kasuistik und Literatur, die bei ca. 50% der Datensätze angefallen sind, stehen in einer separaten Tabelle jeweils 490 Byte zur Verfügung. Das Datenbank-Softwarepaket belegt 900 kB auf der Festplatte. Die Geschwindigkeit einer Suchabfrage läßt sich (einschließlich des Datenbankprogramm-Aufrufs) an den CPU-Zeiten von 1,2 $\frac{1}{60}$ s für die Benutzerphase und 3,9 $\frac{1}{60}$ s für die Systemphase ersehen. Die Laufzeit bis zum Ergebnisausdrucks liegt bei ca. 30 s.

Literatur

1. Moffat AC, Owen P, Brown C (1978) Evaluation of weighted discriminating power calculations as an aid to the selection of chromatographic systems for the analyses of drugs. *J Chromatogr* 161: 179–185
2. Connors KA (1974) Use of multiple R_f values for identification by paper and thin-layer chromatography. *Anal Chem* 46:53–58
3. Vidic E, Klug E (1975) Papier- und dünn-schichtchromatographische Systematik für basische Arzneistoffe. *Z Rechtsmed* 26:283–291
4. Berninger H, Möller MR (1977) Retentionsindices zur gaschromatographischen Identifizierung von Arzneimitteln. *Arch Toxicol* 37:295–305
5. Schepers PGAM, Franke JP, de Zeeuw RA (1983) System evaluation and substance identification in systematic toxicological analysis by the mean list length approach. *J Anal Toxicol* 7: 272–277
6. Franke JP, de Zeeuw RA, Schepers PGAM (1985) Retrieval of analytical data and substance identification in systematic toxicological analysis by the mean list length approach. *J Forensic Sci* 30: 1074–1080
7. Maier RD, Derksen J (1984) Computereinsatz in der praxisbezogenen forensisch-toxikologischen Analytik. *Z Rechtsmed* 92: 159–168
8. Gill R, Law B, Brown C, Moffat AC (1985) A computer search system for the identification of drugs using a combination of thin-layer chromatographic, gas-liquid chromatographic and ultraviolet spectroscopic data. *Analyst* 10: 1059–1065
9. Battista HJ, Bertha E, Kerschhackel G, Mistrik I, Koch W, Wegscheider W (1985) Ein modernes Informationssystem für die toxikologische Analytik auf der Basis des (IV + V)-Systems *Z Rechtsmed* 95: 235–253
10. Wolf W, Reiter A (1986) Der Einsatz des Bürocomputers A 5120 in der toxikologisch-analytischen Routinepraxis unter Verwendung des Datenbanksystems dBase II. *Krim for Wiss* 63: 163–166
11. Manual INFORMIX V2.0 (1986) Siemens AG, Nr U 2605-J-Z95-1
12. Bogusz M, Klys M, Wijsbeek J, Franke JP, de Zeeuw RA (1984) Impact of biological matrix and isolation methods on detectability and interlaboratory variations of TLC R_f -values in systematic toxicological analysis. *J Anal Toxicol* 8: 149–154