講座

商用データベース キーポイントと活用

〔第9回〕 医学・薬学・生物

HIRAKI Junko

平木 淳子*

〔9〕 Medical, pharmaceutical
and biological Information

Commercial databases : the Keypoints and practical use

Series

〔著者抄録〕 生命科学関連の研究は日進月歩であり、数多くの情報の中から必要なものを効率よく得るためにはデータベースは必要不可欠なものである。加えてその中に収められているデータの量とともに、その内容の質や検索機能が充分であることが要求される。本稿ではこの分野で利用価値の高いと思われるデータベースを文献データベース・化合物ディレクトリ型データベース・ニュース記事のデータベースとその内容によって3つに分け、その索引方法の特徴も含めて概観した。また異なるデータベースに採択された同一論文の比較や、複数ファイルを使った横断検索例、オンライン上での統計処理機能を利用した検索集合の解析例についてもあわせて紹介した。

〔著者付与キーワード〕 オンライン検索、データベース、医学、薬学、生物学、インデキシング、横断検索

〔Author abstracts〕 The study of life science is ever-progressing. So, databases are indispensable to select effectively the required item from an enormous amount of information. In addition, the quality of the data and the function for searching are very important factors as well as the quantity of the data. In this paper, the useful databases in this area are classified into three types according to contents namely references, compound-directory-type-data and news articles. The characteristics of the indexing system of these databases are explained. And presented here is a case study on the differences between those abstracts and indexing systems from different databases whose source article is the same. This paper also describes cross-file searching using multiple-file and an example of analysis for retrieved reference groups using the statistical function of a online system.

〔Keywords by Author〕 online retrieval, database, medical information, pharmaceutical information, biological information, indexing, cross-file searching

* 日本新薬(株) 研開発企画本部 情報システム部 情報資料課 (〒601 京都市南区西大路通八条下ル)
Tel.075 (321) 1111
NIPPON SHINYAKU CO.,LTD. Information & Documentation Sect. Information system Dept. R&D
Administration Division
(Hachijosagaru Nishiojiodori Minami-Ku Kyoto-shi, 601)

1. は じ め に

生命科学に関係した調査は古くから JICST 科学技術文献速報, 医学中央雑誌, Index Medicus, Excerpta Medica 等の冊子体や, 薬物の基礎から臨床までをカードに集積した de Haen のカードシステムを利用したり, あるいは1964年にスタートした RINGDOC のパンチカードや磁気テープを社内システムで運用して行なわれてきた。1978年に JOIS-I の公衆回線によるサービスが開始され, 1979年に LOCKHEED 社の商用オンラインシステムがスタートした時点で本格的な調査体制が整ったと思われる。このように生命科学分野の調査は長い実績を持っており, 現在ではすでに成熟期に達しているといつてよい。

一方この分野は単純に医学・薬学・生物学といった形に分割できない面もあり, 各々の分野が互いに複雑に関連しあつて成り立っている。このため, これらのデータベースの利用目的は多岐・多様にわたることが多い。またデータベースそのものが複雑な構造をとっているものもあり, 高度な索引体系をもつデータベースになるとその検索にかなりの熟練と経験を必要とする。加えて, 近年はホストシステムの機能自体が様々な面で充実・拡張され, 複雑な検索にも対応できるようになってきた。

そうした状況を踏まえ, 本稿においては第2章でこの分野のおもなデータベースをそのタイプに分けて概観した。また第3章では文献データベースの索引についての特徴を述べ, 第4章では様々な機能を使った検索例について紹介した。

2. 医学・薬学・生物学分野の主なデータベース¹⁾

2.1 文献情報のデータベース

医学・薬学・生物学領域の技術論文を収録する主なデータベースを表1にまとめた。まず医学関係のデータベースとして欠かすことのできない MEDLINE がある。これは米国国立医学図書館 (NLM) の作成する基礎・臨床医学から看護学, 環境・公衆衛生まで生物医学全般を対象とした安価

でポピュラーなデータベースである。また索引においても詳細な階層構造をなす統制語彙 Medical Subject Headings (MeSH) をもとに下位概念を含めた包括的な検索が可能であり, MeSH 語とそれらの論文中での役割を示す副件名 (Subheadings) との組み合わせによって概念的な限定を加えることができる。この MEDLINE とよく比較される EMBASE は1988年に大きく索引方針が変更され²⁾, 階層構造をもつシソーラス EMTREE や Subheadings に相当するリンク用語の導入によってデータ内容は MEDLINE とよく似たものになっている。しかし MEDLINE との収録対象雑誌の重複は, 両方でカバーする雑誌の約37%にすぎない。また医薬品に関するレコードが全文の約40%と重点がおかれ, ヨーロッパの情報が多く収録されているという特徴がある。国内のデータベースでは JICST 国内医学文献ファイル (JMEDICINE) があり, これは JICST ファイル (後述) から医学分野で国内発行雑誌分を抽出した JICST 分と医学中央雑誌のデータベースから JICST 分を補う形で抽出された医中誌分を合わせたもので, 両者の件数比率は約4対6である。また索引はそれぞれのシソーラスや索引方針に基づくものであるので, 包括的に検索するならば, 実際には二つのデータベースを一度に検索する感覚で注意を払わなければならない。JICST 作成分のレコードには, キーワード索引時に自動的に上位語が付加されているので下位概念まで含めた検索が容易にできる。各データベースを効率よく検索するための用語集であるシソーラスについては, MEDLINE の MeSH・EMBASE の EMTREE・JMEDICINE の JICST 科学技術用語シソーラスともにオンライン上で参照できるようになっており, 各語の階層関係や下位概念を含めるためのディスクリプトコードを見つけるために便利なものとなっている。

特徴的なデータベースでは, 雑誌発行日から2日以内にその書誌事項を提供するという速報性を重視した48HR が最近登場している。また書誌事項と抄録だけではなく図表・写真以外の全文を提

表1 主な文献データベース

データベース名称	収録範囲	収録源	収録件数	更新頻度	内容	オンライン提供システム	作成機関
医学 MEDLINE	1966-	3,900誌	730万 7,000/週	毎週	医学・薬学	DIALOG JOIS STN BRS DATA*STAR	National Library of Medicine (NLM) (7/4/93)
薬学 EMBASE	1974-	4,500誌	520万 6,500/週	毎週	医学・薬学	DIALOG STN BRS DATA*STAR	Elsevier Science Publishers B.V. Excerpta Medica Publishing Group (オランダ)
薬学 JMEDICINE	1981-	3,900誌	215万 20,000/月	毎月	医学 (国内)	JOIS	日本科学技術情報センター (JICST)
生物学 4BHR	最近64分	250誌	18,000	毎日	医学・心理学	BRS	Reliance Medical Information, Inc (7/4/93)
生物学 Comprehensive Core Medical Library	(1982-)	90誌+56books (全文)	29万4千	隔週	医学	BRS	BRS Online Products (7/4/93)
生物学 Derwent Drug File (RINGDOC)	1964-	900誌	120万 4,000/月	毎月	薬学	DIALOG ORBIT	Derwent Publications, Ltd. (イギリス)
生物学 JAPICDOC	1973-	270誌	19万 1,000/月	毎月	薬学 (国内)	JOIS JIP	日本医薬情報センター (JAPIC)
薬学 International Pharmaceutical Abstracts	1974-	650誌	21万	毎月	医薬品の開発・使用 製剤学	DIALOG BRS DATA*STAR	American Society of Hospital Pharmacists (7/4/93)
薬学 BIOSIS PREVIEWS	1969-	9,000誌, 学会抄録 モクワフ, ポート	840万 22,500/週	毎週	生物科学	DIALOG STN BRS DATA*STAR	Biological Abstracts Inc. (BIOSIS) (7/4/93)
遺伝学 Life Science Collection	1978-	5,000誌, 単行本, 特許 会議録	152万 10,000/月	毎月	生物学・医学・ 生化学・微生物学	DIALOG	Cambridge Scientific Abstracts (7/4/93)
遺伝学 CAB Abstracts	1972-	10,000誌, 単行本, 特許 技術報告書	280万	毎月	農業・獣医学・栄養学	DIALOG BRS STN DATA*STAR	CAB International (イギリス)
工学 AGRICOLA	1970-	2,000誌, 単行本, 会議録 研究報告書, 政府刊行物	386万	毎月	農業および関連分野	DIALOG	National Agricultural Library (NAL) (7/4/93)
工学 DERWENT Biotechnology Abstracts	1982-	1,100誌, 会議録, 特許	14万 1,100/月	毎月	遺伝子工学	DIALOG ORBIT	Derwent Publications, Ltd. (イギリス)
科学 JICST	1981-	15,000誌	623万 30,000/月	毎月	科学全般	JOIS STN ¹⁾	日本科学技術情報センター (JICST)
科学 C. A. -	1957-	9,000誌, 特許, 単行本, ポート	1,100万 17,000/2週	隔週	化学全般	DIALOG STN BRS DATA*STAR	Chemical Abstracts Service (CAS) (7/4/93)
科学 SCISEARCH	1974-	7,600誌	11.5百万 14,000/月	毎週	科学技術	DIALOG DATA*STAR	Institute for Scientific Information (ISI) (7/4/93)
科学 Current Contents	1990-	7,200誌	415万 14,000/週	毎週	学術雑誌の 目次と要誌	DIALOG BRS	"
一般 PASCAL	1973-	8,500誌, 会議録, 特許 技術報告書, 単行本	890万 40,000/月	毎月	自然科学	DIALOG	Centre National de la Recherche Scientifique Institut de l'Information et Technique (フランス)
特 ARIS	1987-	420誌	5,000 150/月	毎月	薬物の副作用 (国内中心)	DRUGBASE	榊小田島医薬情報センター
特 SEEDBASE	最近12年	Meller's Side Effects の全文	63,000	年4回 再装填	副作用と薬剤 の相互作用型	DIALOG BRS DATA*STAR	Elsevier Science Publishers B.V. Excerpta Medica Publishing Group (オランダ)
分野 TOXLINE	1965-	MEDLINE, BIOSISなどの 17冊, 77冊	182万 10,000/月	毎月	毒性	JOIS DIALOG BRS DATA*STAR	National Library of Medicine (NLM) (7/4/93)
分野 CANCERLIT	1963-	3,500誌, 会議録, 政府報告書, モクワフ	91万 7,500/月	毎月	癌関連	JOIS DIALOG BRS DATA*STAR	International Cancer Research DataBank Branch (ICRDB) (7/4/93)

収録件数は'93年1月を基本とした

¹⁾ STNでは1985年以降の国内発行雑誌分「JICST-EJ774」

供する Comprehensive Core Medical Journal (CCML)もある。CCMLは約90誌と56の単行本と対象が広いものであるが、このほかにも米国化学会発行の19誌に対応するSTNのCJACS、オランダのElsevier社発行の化学関連4誌に対応する同じくSTNのCJELSEVIERや単独誌に対応するものではDIALOGのTHE NEW ENGLAND JOURNAL OF MEDICINE ONLINEなどがあり、全文データベースは増加する傾向にある。

医学の中でも特定の領域や主題に的を絞ったデータベースもある。CANCERLITは癌関連の記事のみを収録したもので、米国国立ガン研究所の国際ガン研究データバンク部門(ICRDB)が作成している。1983年以降はMEDLINEの癌関係を主題としたレコードも追加され、1980年以降からはMeSHを使って索引されている。このCANCERLITとMEDLINEなどからAIDSに関する情報のみを収録したAIDSLINEというデータベースもNLMにより作成されている。

薬学を中心とするデータベースでは以前はRINGDOCと呼ばれていたDerwent Drug File (DDF)があり、雑誌論文のほかに学会抄録も全体のレコードの約3割を占めている。このデータベースは会員制でかなり価格の高いものであるが、その抄録と索引の詳細なことからきめ細かな検索に対応できるようになっている³⁾。これについては他のデータベースとの比較とともに後で述べる。国内では日本医薬情報センター(JAPIC)の作成するJAPICDOCがあり、これは医薬品の有効性と安全性に関する論文を収録しているが、同じく会員制で利用はJAPICの一部の会員に限られる。DDF, JAPICDOCどちらも対象誌・収録レコードの数が他の文献データベースに比べて少ないが、薬学特有の索引語や分類が検索に便利なデータベースである。このほかにもInternational Pharmaceutical Abstracts(IPA)では薬学分野全般ということで、医薬品の開発・歴史・製剤・法規制などの文献も含まれている。

薬学のなかでも特に副作用に的を絞ったデータベースとしてARIS, SEDBASE⁴⁾がある。ARIS

は医薬品卸売業である(株)小田島の医薬情報センターが作成したもので国内約400誌、海外20誌の論文のほか、厚生省医薬品副作用情報も収録対象となっている。索引に関しては副作用用語の辞書として厚生省監修の医薬品副作用用語集をもとにしたHYTARDという辞書があり、これは上下階層関係を整理してコード化されたものでオンラインでも参照できる。また薬剤に関してもDRUGファイルという医薬品辞書ファイルを別に持ち、同義名や薬効分類などのほか、そのコードによって投与経路などの区別もできるようになっている。また本文のMAINファイルでは副作用と原因薬剤にそれぞれ対応した番号が付与されて、その因果関係を規定した検索も可能である。SEDBASEはMeyler's Side Effects of DrugsとThe Side Effects of Drugs Annualsの全文を収録したものでEMBASEと同じExcerpta Medica社によって作成されたものである。これはEMBASEのような文献データベースと異なり、文献などに報告された副作用や相互作用を評価・濃縮したもので、医薬品とその副作用が1レコードの中に1対1で索引されそれぞれのディレクトリ型となっている。よって薬物名で検索するとその薬剤の副作用の一覧が、副作用名で検索するとその副作用を起こす原因薬剤の一覧が得られる。TOXLINEは薬物および化学物質の毒性関連文献を17のサブファイルから抽出したものでMEDLINEからのレコードが約40%を占めている。

生物学に関する文献情報としてはBIOSIS PREVIEWSがあげられ、生物学・生物医学・薬学・動植物学・微生物学・農学・環境生物学・生物工学などライフサイエンス全般を対象としている。雑誌論文以外に学会抄録・会議録などからも収録され全体のレコードの約3分の1を占めているので、新しい知見などについては速報性も期待できる。その索引については、全部で約600の概念を示すコンセプトコードとウイルス・細菌・動植物を包含する生物分類コードがある。前者は一論文に平均約10個付与され、後者は生体を門・綱・目・科・(属)から系統的に検索できるように階層

的になっている。このほかにも農学を中心とした CAB Abstracts や AGRICOLA や遺伝子工学の DERWENT Biotechnology Abstracts があり、動植物・微生物・バイオテクノロジー関連の情報が得られる。

また広く化学・科学全般を収録するデータベースとしておもに以下のものがあげられる。まず国内の日本科学技術情報センターの作成する JICST 科学技術文献ファイルは世界50数か国から収集した雑誌、レポート、会議資料、公共資料から収録されその約半分が日本発行のものである。また先にも述べた JICST 科学技術用語シソーラスに基づいて階層関係を考慮した索引がなされ、世界の論文を日本語で検索し抄録が参照できるというのも大きな利点である。またこのうち1985年以降の国内発行雑誌分が英訳され JICST-E ファイルとして STN から提供されているので、世界各国からも日本国内の論文へ英語でアクセスでき、索引も JICST 科学技術用語シソーラスの英語版に基づいて統制されたものとなっている。ただ欲をいえばシソーラスの改訂が約5年ごとということで新しい概念がシソーラスに反映しにくい面があるので、そのあたりが改善されればより使いやすくなると思われる。Current Contents と SCISEARCH はともに ISI 社の作成する科学全般を対象とした速報性のあるデータベースである。Current Contents は雑誌の目次が参照できる目次レコードと個々の論文のレコードの二つのタイプのデータを持ち、後者の1991年以降のレコードには抄録がついているものもある。SCISEARCH は書誌事項・抄録の他に引用文献をも収録し、しかもその検索が可能なところに大きな特徴がある。このデータベースでは、ある論文が発表された以後にそれを引用して書かれた論文の検索、つまり逆孫引きができる。Chemical Abstracts(CA)についてはここでは詳述しないが、化学文献ファイルとしては最大のものであり、文献だけでなく27か国および2国際機関（ヨーロッパ特許庁、世界知的所有権機関）の特許も収録しているので、その化合物辞書ファイルとともに

に、化合物の新規性や合成法、植物の成分などの検索には欠かせないデータベースである。PASCAL はフランス国立科学研究センター(CNRS)が作成するもので英語のほかに、フランス語、ときにはスペイン語のディスクリプタが付与されており、抄録もフランス語で書かれているものもある。

2.2 化合物ディレクトリ型データベース

前項の文献データベースに対して、ここでは1物質を1データ単位とした化学物質・薬物のデータベースを化合物ディレクトリ型データベースと呼ぶこととして以下にとりあげる。このタイプのデータベースの利用法には大きく二通りあり、一つは特定の物質に関するデータの参照であり、冊子の代用的な利用法である。

もう一つの化合物ディレクトリ型データベースの有効な利用法は、ある条件にあてはまる化合物・薬物をリストアップすることにある。たとえば「アンジオテンシン拮抗薬」という検索を文献データベースで行うと得られる集合は論文単位であり1薬剤に対して複数の論文が対応することが充分考えられるので、得られるレコード数は決して該当薬剤の数ではない。これに対し作用機序の項目をもつ薬物ディレクトリ型データベースで同じ語を入力すると、得られるレコードの数は該当する薬物の数となり得る。化合物ディレクトリ型データベースは複数の情報源からその物質に関するデータをまとめてあるものなので、その詳細についてはさておき該当する物質の一覧を知りたいときに有用であるだけでなく、後でも述べるがその物質群に対する何らかの情報を得たい場合には、次の検索を行うための手段としても利用できるものである。表2に医学・薬学・生物学に関連の深いものをまとめた。

現在開発途中にある治験薬を収録対象としたものに、Pharmaprojects, IMSWorld R&D Focus, UNLISTED DRUGS, PHARMCAST-International, 明日の新薬がある。データ量やデータ項目については各データベースで異なりよく比較されるところであるが⁵⁾、おもに INN,

表2 薬物・化合物ディレクトリ型オンラインデータベース

データベース名称	収録対象	収録件数	更新	オンラインホスト	他の媒体	作成機関	主なデータ項目
Pharmaprojects	世界の開発薬	14,500	月1回 再装填	DIALOG STN DATA*STAR	冊子	PJB Publications	名称 ^{1,2} , 分子式, 薬効分類, 作用機序 ^{1,2} , 会社および特許関係, 国別開発段階, 開発概略, (構造図), CAS登録番号
IMSWorld R&D Focus	"	4,600	毎月	DIALOG DATA*STAR	冊子 CD-ROM	IMSWorld Publications LTD.	名称 ^{1,2} , 薬効分類, 会社および特許関係, 開発段階, 特許情報, 概要, 開発概略, CAS登録番号
UNLISTED DRUGS	"	14,000	毎月	DIALOG	冊子	Pharmaco Medical Documentation Inc	名称 ^{1,2} , 薬効分類, 会社名, 開発国, 分子式, CAS登録番号
PHARMCAST -International	"	4,000	毎月	DRUGBASE	冊子	デラミッド・フアム・イン・サ・ピ・ス	名称 ^{1,2} , 薬効分類, 会社名, 国別開発段階および会社名, 概要, 剤型, 効能効果(国別), 申請日・承認日・薬価収載日
明日の新薬	国内の開発薬	2,500	毎月	DRUGBASE G-SEARCH	冊子・ 磁気テープ	"	名称 ^{1,2} , 薬効分類, 会社名, 開発段階, 開発概略, 基礎・臨床データ, 剤型, 効能効果, 申請日・承認日・薬価収載日
Derwent Drug Registry (Standard Drug File)	DDF ¹ 2回以上登場の 化合物, その他	23,800	不定期 (年3回)	DIALOG	磁気テープ	DERWENT Publications LTD.	DDF登録名, 薬物名, 作用機序 ^{1,2} , 構造コード, CAS登録番号
医薬品情報 (添付文書)	国内販売医薬品の 添付文書	16,500	毎月	G-SEARCH	CD-ROM	MEDIS-DC 日本医薬情報センター	商品名, 一般名, 会社名, 薬効分類, 効能効果, 用法用量, 薬効薬理, 体内動態など図表・引用以外の添付文書項目
Drug Information Fulltext	米国内販医薬品の 評価的モナグラフ	1,300	年4回 ^{1,5}	DIALOG BRS DATA*STAR		American society of Hospital Pharmacists	商品名, 一般名, 会社名, 薬効分類, 効能効果, 用法用量, 薬効薬理, 薬物動態, 毒性, 注意事項, 引用文献書誌事項
IMSWorld New Product Launches	世界の新薬情報	47,000	毎月	DIALOG DATA*STAR	冊子 CD-ROM	IMSWorld Publications LTD.	製品名, 成分名, 国名, 会社名 薬効分類, 適応症, 発売年月, 剤型および薬価
IMSWorld Product Monographs	世界の上市薬物リスト	345,000	毎月	DATA*STAR	冊子	"	製品名, 成分名, 国名, 会社名 剤型毎: 薬効分類, 発売年月, 医療・一般用別, 薬価
Martindale ONLINE	主要国の医薬品	5,200	年1回	DATA*STAR	冊子 CD-ROM	The Pharmaceutical Society of G. B. ^{1,3}	名称 ^{1,2} , 薬効, 物性, 用法用量, 注意, 副作用, 薬物動態, 各国医薬品集への収載, CAS登録番号
CHEMSEARCH, CHENNAME CHEMDEX	Chemical Abstracts 収載化合物('57~)	12,1百万	毎月	DIALOG DATA *STAR ORBIT		Chemical Abstracts Service (CAS)	CAS登録番号 ^{1,4} , 分子式, CA索引名(基本母核+置換基), 名称 ^{1,2} , 環系データ(数・大きさ・構成原子など), 成分数
REGISTRY	"	12,1百万	毎週	STN		"	構造図, CAS登録番号 ^{1,4} , 分子式, CA索引名(名称モリット) 名称 ^{1,2} , 環系データ(数・大きさ・構成原子など), 成分数
Merck Index Online	重要な化学物質 生物学的物質	10,000	年2回	DIALOG BRS	冊子	Merck & Co., Inc	名称 ^{1,2} , 物性, 用途・毒性・化学・生物医学・特許文献への 参照事項, 薬効分類, 分子式, CAS登録番号
Registry of Toxic Effects of Chemical Substances (RTECS)	化合物の毒性データ	117,000	年4回	DIALOG STN DATA*STAR CIS	冊子	National Institute of Occupational Safety and Health (NIOSH)	名称 ^{1,2} , 毒性データ, 変異原性データ, 生殖毒性データ, 発癌性 データ, 刺激性データ(各々動物種, 投与経路, 投与量, 出典 情報の表)

収録件数は'93年1月を基本とした

¹ DDF: Derwent Drug File (JERINGDOC), Derwent社作成の薬学文献データベース² 各一般名, 商品名, 治験番号, 慣用名など³ G. B.: Great Britain⁴ 成分登録番号, 廃止された登録番号, 置換された登録番号, 優先登録番号, 非優先登録番号などを含む⁵ 年4回の更新の他, 毎年再装填される

USAN などの一般名・治験番号・商品名・化学名・分子式・CAS 登録番号などの化合物同定情報、開発会社およびライセンス関係、開発段階（国別）、薬効分類、作用メカニズム、開発概略などを中心としたその治験薬のプロフィールとなっている。治験薬のデータベースはその開発進行状況をいかにタイムリーに提供するかといったところも大きな評価点となるが、Pharmaprojects はデータ更新頻度がまもなく月 1 回から週 1 回になるとの情報もある。

一方上市済みの薬物については、国内では医療用医薬品の添付文書のデータベースがあるが、これはデータ単位が添付文書の単位となっており、検索できる項目も限られている。また世界の医薬品について、Drug Information Fulltext は米国市販医薬品について、実際の臨床経験をも含めた評価的モノグラフで、AHFS(American Hospital Formulary Service) Drug Information と Handbook on Injectable Drugs の両冊子に対応するものである。Martindale ONLINE は英国薬学会の出版する Martindale, The Extra Pharmacopeia を基本とした世界の医薬品の特性を提供するものである。最近登場したものには、医薬品の市場・処方統計を提供する IMS 社が作成するデータベース IMSWorld New Product Launches と IMSWorld Product Monographs がある。前者は発売時のデータを収録しており 1 つのレコードに関するデータ内容の更新は行われず、品目の追加のみである。これに対して後者では薬価などの改定や剤型の追加とともにデータが更新される。このデータベースでは医薬品の各国での発売状況や商品名・薬価・発売時の適応症などが容易に調べられる便利なものである。Merck Index と Registry of Toxic Effects of Chemical Substances (RTECS) はそれぞれ同名の冊子に相当するものである。RTECS は化学物質の毒性関連データの数値をその出典とともに提供し、ホストシステムによっては数値検索、つまりある毒性値範囲に当てはまる化合物をリストするといった使い方もできる。

また化合物ディレクトリ型データベースの中に

は、文献データベースを収録源としてそれとリンクした検索を目的とするものもある。その一つに薬学論文データベース Derwent Drug File(DDF) から薬物をピックアップした Derwent Drug Registry (DDR) がある。これは構造キーワード、構造コードを使ってある程度の構造検索が可能であり、場合によっては薬効と組み合わせて検索した薬物の集合を DDF や他のデータベースへクロスファイルした後、他の要素と組み合わせて検索を行うことができる。DDR を基本として各国医薬品集なども収録対象にした Standard Drug File with Graphics(SDF-G)はオンラインで提供されているものではないが In-House でのグラフィック構造検索が可能である。Chemical Abstracts (CA)の化合物辞書としては STN での REGISTRY ファイル、DIALOG での CHEMSEARCH などがありそれぞれの特徴をもっている⁶⁾が、おもに化合物同定情報を収録している。STN の REGISTRY ファイルではオンライン上でのグラフィック構造検索が可能で、その機能は年々レベルアップしており、かなり綿密な構造指定での部分構造検索や、環数と構成元素のみを指定したようなあいまいな質問式にも対応できる。また REGISTRY ファイルには核酸とタンパク質の配列データも含まれておりこれらの配列検索も可能である。ほかにも配列検索で有名なものとして米国 IntelliGenetics 社の BIONET タイムシェアリングサービスにおける GenBank, EMBL 等のデータバンクのほか、特許を収録源とした Derwent 社の Geneseq がある。

2.3 ニュース記事を収録するデータベース

表 3 には医学・薬学・生物学関連の最近の動向や経済的側面、規制に関する最新情報をニュースのような形で提供するおもなデータベースを一覧にした。これらには印刷体のニュースレターに対応するものや、文献データベースのように様々な一次情報を収録源としてその所在を示しているものもある。また前者でも、以前は見出し・要約しか提供されていなかったものも全文で提供される

表3 医学・薬学・生物学関連ニュース オンラインデータベース

データベース名称	収録内容	収録件数	更新	オンライン	収録	作成機関	対応出版物/収録源
Pharmaceutical News Index (PNI)	医薬品, 医療機器など衛生産業に関する最新情報	378,000	毎週	DIALOG STN BRS	1975- ^{*2}	UMI/Data Courier	SCRIP World Pharmaceutical News, FDC Reports, Pharma - Japan, Clinica World Medical Device News など20誌
Pharmaceutical and Healthcare Industry News Database (PHIND)	製薬/医療機器/動物の健康/農作物・農芸化学産業の製品・研究- ^{*1} , 会社情報など	201,000	毎日	DIALOG BRS DATA*STAR	1980- ^{*2}	PJB Publications	SCRIP World Pharmaceutical News, CLINICA World Medical Device and Diagnostic News, ANIMAL PHARM World Animal Health News, AGROW World Crop Protection News の4誌
INSWorld R&D Focus	世界の医薬品の研究開発についての最新情報	2,500 40/週	毎週	DIALOG DATA*STAR	1991-	INSWorld Publications Ltd.	INSWorld R&D Focus Drug News (週刊)
ADIS Drug News	新薬開発, 薬物療法, 市場, 副作用などの傾向, 最新情報	55,000	毎日	DATA*STAR BRS	1983-	ADIS International, Ltd.	INPHARMA (週刊), REACTIONS (週刊) の2誌
日刊薬業新聞	国内を中心とした医薬品・薬業界に関する最新情報	22,000	毎日	G-SEARCH	1990-	薬業時報社	日刊薬業新聞
FEDERAL REGISTER	米国政府 (FDA, EPA など) の官報の全文。規制情報	156,000 135/日	毎日	DIALOG BRS ^{*1}	1988-	Dialog Information Services, Inc.	Federal Register (日刊)
FDC-REPORTS	世界の保健衛生産業の会社/製品/市場/規制など最新情報	45,000 180/週	毎週	DIALOG BRS DATA*STAR	1987-	FDC Reports, Inc.	The Pink Sheet (製薬), The Gray Sheet (医療機器) The Rose Sheet (化粧品) の3誌 全て週刊
DIAGENES	医薬品・医療器具の米国政府規制関連のニュース, 未出版文献	182万 250/週	毎週	DIALOG BRS DATA*STAR	1981- ^{*3}	FOI Services, Inc.	Washington Drug Letter, The GMP Letter などのニュース, 情報公開法下で入手されたFDA未出版文献, FDA承認品リスト等
BIOBUSINESS	生物学を中心とした生命科学分野の経済的側面	491,000 3600/月	毎月	DIALOG STN BRS DATA*STAR	1985-	Biological Abstracts社 (BIOSIS)	世界中の技術雑誌とビジネス誌, 一般誌, ニュース, 会議録, 単行本, 米国特許など
Chemical Industry Note (CIN)	世界の化学工業に関する製品, 市場などの最新情報	953,000 2000/2週	隔週	DIALOG ORBIT DATA*STAR STN	1974-	Chemical Abstracts Service (CAS)	Chemical Industry Notes (週刊)
Chemical Business Newbase	化学関連産業に関するビジネスニュース (3-ロツパ 中心)	262,000 700/週	毎週	DIALOG STN DATA*STAR	1984-	Royal Society of Chemistry	雑誌, 新聞, 企業の年次報告, 報道機関向け発表, 市場調査報告書, 台帳など

収録件数は'93年1月を基本とした

^{*1} DIALOGでは全文が提供されているがBRSでは抄録のみ^{*2} 収録誌により開始年月が異なる^{*3} 未出版文献は1976年以降

ことが多くなってきた。例えば製薬業界に関連する最近のトピックを週に2回印刷物として発行しているSCRIPという情報誌がある。これについてはPharmaceutical News Index(PNI)ではそのタイトルとキーワードが収録されているが、Pharmaceutical and Healthcare Industry News Database(PHIND)ではその全文を参照できる。さらにPHINDでは出版前の原稿段階のデータが載っており、印刷物よりも速くその情報を入手することも可能である。

3. 文献データベースの比較

同一論文の収録内容について、Derwent Drug File(以下DDF)、EMBASE、MEDLINE、BIOSIS、Current Contents(以下CC)の5つのデータベースを比較する。CC以外の4つのデータベースについてはその内容・検索法や提供システムによる比較について紹介された論文があるので参照されたい³⁾⁷⁾⁸⁾⁹⁾。図1・図2・図3はジクロフェナク徐放剤が原因となった副作用の回盲腸潰瘍について述べられた同一論文をそれぞれ上記5つのデータベースから出力したものである。この論文の収録内容について、抄録・論文中の薬物の索引・索引語の階層関係・索引語に対する役割付けとそれらの関連性という4つの観点から述べることにする。

まず抄録について、DDFでは原著論文の内容をできるだけ反映するためデータベース作成者により書き直された第三者抄録の方式がとられ、実験方法や数値などを盛り込んだ詳しい内容となっている。これに対し、他の4つのデータベースでは著者抄録が採用されており、内容の面では劣るが速報性はDDFよりも期待できる。

次に適合性、網羅性の高い検索を行うために重要な索引部分を比較する。論文中に記載のある薬剤をどれだけ索引語として採用しているかについて、DDFでは原因薬剤のジクロフェナク、併用薬のアザチオプリンとプレドニゾロン、副作用に対する処置として用いられたシメチジンが索引され

ている。EMBASEではシメチジン以外の3薬剤、MEDLINEでは抄録に記載のあるジクロフェナクのみである。BIOSIS、CCでは共に具体的な薬剤名は索引語として特に付与されていない。この場合DDF、EMBASEではタイトル、抄録中に出現しなくても原論文中に記述のあるアザチオプリン、プレドニゾロンでもこの論文を検索することができる。

また索引語の階層関係について、具体的な薬物名ではなく「抗炎症剤」という総称からの検索を行う場合を想定する。DDFでは各薬剤に対し上位語となる薬効分類語「ANTIINFLAMMATORIES」が付与されているし、MEDLINEやEMBASEでは各索引語を階層的に整理したシソーラスを持っており、下位概念まで含めた検索が可能のように工夫されている。これに対しBIOSIS、CCでは「抗炎症剤」を表すキーワードはあるもののこれらに相当する具体的な薬剤をすべて検索できるようには設計されておらず、「抗炎症剤」という記載そのものがなければヒットしないし、特にCCでは統制語ではないため考えられる表現を様々入力する必要がある。

最後に索引語の役割とそれらの関連性について、ここでは副作用と原因薬剤の因果関係について見てみよう。この論文は「ジクロフェナクによる潰瘍」を述べた論文であるので、「プレドニゾロンによる潰瘍」という検索を行ったときこの論文は回答集合に含まれないほうが適合性の高い検索のできる索引体系をもったデータベースといえる。

図1に示すとおりDDFでは索引された4つの薬剤が*01*～*04*の別々のセンテンスに分割されており、また各薬剤に関する事項はその論文中的役割を表すクオリファイアとともに関連ある薬剤のセンテンスのみに入っている。検索時には「同一センテンス中に存在する」という検索が可能のように設計されているので、原因薬剤のジクロフェナクと副作用として発生した潰瘍の因果関係を規定して検索ができる。またプレドニゾロンによる副作用「ST=PREDNISOLONE」と副作用として発生した潰瘍「ST=ULCER」は別々

< Derwent Drug File サンプルレコード >

0535741

ACCESSION NUMBER: 93-24016

THEMATIC GROUPS: T S E

LANGUAGE: E

SDI PROFILES: 16 24 35 43 46

AUTHORS: Hudson N; Wilkinson M J; Swannell A J; Steele R J; Hawkey C

LOCATION: Nottingham, U.K.

CODEN: APTHEN

JOURNAL REFERENCE: Aliment. Pharmacol. Ther. (7, No. 2, 197-200, 1993)

TITLE: Ileo-Caecal Ulceration Associated with the use of Diclofenac Slow Release.

ABSTRACT: High dose slow release diclofenac (Voltarol Retard) induced ileo-cecal ulceration in a case of seropositive arthropathy. Other medication included prednisolone (CIBA-Geigy) and azathioprine. Dyspeptic symptoms responded to cimetidine.

EXTENSION: A 38 yr-old man developed a severe destructive arthropathy associated with a positive rheumatoid factor, a high ESR and thrombocytosis at the age of 30 yr. He was stabilized on diclofenac SR (Voltarol Retard, 100 mg, b.i.d.), prednisolone (CIBA-Geigy, 7.5 mg) and azathioprine (50 mg) every morning. 5 Yr later he developed an iron deficiency anemia with epigastric pain and diarrhea. He was transfused and his dyspepsia improved on cimetidine. 5 Mth later he developed right-sided lower abdominal pain. Although his dyspeptic symptoms resolved, episodic diarrhea continued and hemoglobin levels fell. A leukocyte scan revealed an intense ring of activity in the right iliac fossa. Biopsies revealed granulation tissue and ulcer slough adjacent to normal colonic mucosa which contained only a few inflammatory cells. The ulcers in the ileum, cecum and ascending colon all showed similar features differing only in depth of penetration. The ulcers in the colon and ileum penetrated through the muscularis propria, the base of these ulcers consisting of fibrous scar tissue. In the cecum the ulcer had perforated, the resulting tract being lined by fibrous scar tissue, granulation tissue and slough. Overall, the ulcers had a striking similarity to peptic ulcers and NSAID-induced ulcers seen in the proximal gastrointestinal tract. There was no evidence of vasculitis, granulomata or other features suggesting a specific aetiology within the ulcers or surrounding mucosa. No organisms were seen on acid fast or fungal stains in these areas and bacterial culture of the ileum and cecum were negative.

4 Fig. 28 Ref. (JP)

Department of Therapeutics, University Hospital, Nottingham NG7 2UH, England. (C. J. H.).

COMMON TERMS: FT=IN-VIVO FT=CASE-HISTORY FT=CASES

LINK TERMS:

01 TR=DICLOFENAC ST=DICLOFENAC TR=VOLTAROL-RETARD ST=VOLTAROL-RETARD
TR=RHEUMATOID TR=ARTHRITIS TR=JOINT-DISEASE ST=ULCER
ST=GASTROENTEROPATHY ST=CECUM ST=ILEUM ST=DYSPEPSIA FT=ANTIINFLAMMATORY
FT=INTESTINE FT=ANTIINFLAMMATORIES FT=ANALGESICS
02 TR=PROSTAGLANDIN-ANTAGONISTS FT=15307-86-5 RN=DICLOFENAC FT=TR FT=ST
TR=AZATHIOPRINE TR=RHEUMATOID TR=ARTHRITIS TR=JOINT-DISEASE
FT=CYTOSTATICS FT=IMMUNOSUPPRESSIVES FT=446-86-6 RN=AZATHIOPR FT=TR
03 TR=PREDNISOLONE ST=PREDNISOLONE FT=CIBA-GEIGY TR=RHEUMATOID
TR=ARTHRITIS TR=JOINT-DISEASE FT=CORTICOSTEROID FT=CORTICOSTEROIDS
FT=50-24-8 RN=PDNISOLON FT=TR FT=ST
04 TR=CIMETIDINE TR=DYSPEPSIA TR=ULCER TR=GASTROENTEROPATHY TR=CECUM
TR=ILEUM FT=ANTIULCER FT=INTESTINE FT=ANTIHISTAMINES-H2 FT=ANTIULCERS
FT=GASTRIC-SECRETION-INHIBITORS FT=51481-61-9 RN=CIMETIDIN FT=TR

線部 ST=は「副作用」を示すキーワード。*01*ではDICLOFENACの副作用として
ULCERが生じたことを表し、その部位についてもST=で索引されている。
TR=は「治療」を示すキーワード。*01*ではRHEUMATOID ARTHRITIS の治療の
ためにDICLOFENACが投与され、*04*ではULCERおよびその部位は治療の
対象であり、治療薬としてCIMETIDINEが投与されたことを表している。

線部 各索引語に対しその上位語として付与されたもの。*01*では薬剤の
DICLOFENACが属する薬効分類のANTIINFLAMMATORIES, ANALGESICS,
PROSTAGLANDIN-ANTAGONISTSの3つが、*01*、*04*では疾病のULCER、
CECUM、ILEUM、DYSPEPSIAにはGASTROENTEROPATHYがそれぞれ対応す
るものと同じキーワードで付与されている。

図1 Derwent Drug File サンプルレコード

< EMBASEデータベース >

8805735 EMBASE No: 93109467

Ileo-caecal ulceration associated with the use of diclofenac slow release
Hudson N.; Wilkinson M.J.; Swannell A.J.; Steele R.J.; Hawkey C.J.

Department of Therapeutics, University Hospital, Nottingham NG7 2UH
United Kingdom

ALIMENT. PHARMACOL. THER. (United Kingdom), 1993, 7/2 (197-200) CODEN:

APTHE ISSN: 0269-2813 ADONIS ORDER NUMBER: 026928139300026G

LANGUAGES: English SUMMARY LANGUAGES: English

SUBFILES: 030; 031; 048

A 38-year-old man with seropositive arthropathy on long-standing treatment with diclofenac slow release presented with abdominal pain, diarrhoea and anaemia. Upper gastrointestinal endoscopy and barium series were unremarkable. Five months later he presented with a right iliac fossa mass. Investigations revealed discrete ulceration in the terminal ileum and caecum confirmed at laparotomy. Other possible causes of intestinal ulceration were excluded suggesting a diagnosis of non-steroidal anti-inflammatory drug associated intestinal ulceration.

TRADE NAME/MANUFACTURER NAME: voltarol; GBR ciba geigy

ENTACS:

Therapy 0160; Digestive system 0935; Small intestine 0941; Large intestine

0940; Diagnosis 0140; Mammal 0738; Human 0888; Male 0041; Case report 0151;

Adult 0018; Oral and intragastric drug administration 0181; Article 0060;

Adverse drug reaction 0198; Iatrogenic disease 0300

DRUG DESCRIPTORS:

*diclofenac--adverse drug reaction--ae; *diclofenac--drug therapy--dt; *

nonsteroid antiinflammatory agent

sustained release preparation; prednisolone--drug therapy--dt; azathioprine

--drug therapy--dt; barium

MEDICAL DESCRIPTORS:

*intestine ulcer--surgery--su; *intestine ulcer--side effect--si; *

arthropathy--drug therapy--dt

ileum; cecum; abdominal pain--side effect--si; diarrhea--side effect--si;

anemia--side effect--si; gastrointestinal endoscopy; human; male; case

report; adult; oral drug administration; article

ENCLAS DRUG CODES:

03700000000; 038000000000

CAS REGISTRY NO.: 15307-79-6; 15307-86-5; 50-24-8; 446-86-6; 7440-39-3

< MEDLINEデータベース >

08540228 93250228

Ileo-caecal ulceration associated with the use of diclofenac slow release.

Hudson N; Wilkinson MJ; Swannell AJ; Steele RJ; Hawkey CJ

Department of Therapeutics, University Hospital, Queen's Medical Centre, Nottingham, UK.

Aliment Pharmacol Ther (ENGLAND) Apr 1993, 7 (2) p197-200, ISSN

0269-2813 Journal Code: ASD

Languages: ENGLISH

Document type: JOURNAL ARTICLE

JOURNAL ANNOUNCEMENT: 9308

Subfile: INDEX MEDICUS

A 38-year-old man with seropositive arthropathy on long-standing treatment with diclofenac slow release presented with abdominal pain, diarrhoea and anaemia. Upper gastrointestinal endoscopy and barium series were unremarkable. Five months later he presented with a right iliac fossa mass. Investigations revealed discrete ulceration in the terminal ileum and caecum confirmed at laparotomy. Other possible causes of intestinal ulceration were excluded suggesting a diagnosis of non-steroidal anti-inflammatory drug associated intestinal ulceration.

Tags: Case Report; Human; Male

Descriptors: *Diclofenac--Adverse Effects--AE; *Intestinal Diseases

--Chemically Induced--CI; *Ulcer--Chemically Induced--CI; Adult;

Delayed-Action Preparations; Diclofenac--Administration and Dosage--AD;

Diclofenac--Therapeutic Use--TU; Intestinal Diseases--Diagnosis--DI; Joint

Diseases--Drug Therapy--DT; Ulcer--Diagnosis--DI

CAS Registry No.: 0 (Delayed-Action Preparations): 15307-86-5

(Diclofenac)

← この例ではDRUG DESCRIPTORのdiclofenac,

prednisolone, azathioprineに薬物療法を

示す「DT」、diclofenacには副作用を示す

「AE」のツグ用語が組み合わされて索引され

ている。MEDICAL DESCRIPTORでは治療「DT」

の対象となったのはarthropathyで、

intestine ulcerは副作用「Si」として発現

したものであり、これに対して手術「SU」が

行われたことが索引上で判断できる。

↑

索引の対象となったdiclofenacに

は治療的利用「TU」と副作用「AE」の

ツグ用語が組み合わされ、Ulcer

, Intestinal Diseasesには化学的

に誘発されたことを表す「CI」、

Joint Diseasesには薬物療法の対

象であることを表す「DT」が付与さ

れている。

図2 EMBASE・MEDLINE サンプルレコード

- < BIOSISサーチワード >
 10411825 BIOSIS Number: 96011825
 ILEO-CAECAL ULCERATION ASSOCIATED WITH THE USE OF DICLOFENAC SLOW RELEASE
 HUDSON N; WILKINSON M J; SWANNELL A J; STEELE R J; HAWKEY C J
 DEP. THERAPEUTICS, UNIV. HOSPITAL, NOTTINGHAM NG7 2UH, UK.
 ALIMENT PHARMACOL THER 7 (2). 1993. 197-200. CODEN: APTHE
 Full Journal Title: Alimentary Pharmacology & Therapeutics
 Language: ENGLISH
 Subfile: BA (Biological Abstracts)
 A 38-year-old man with seropositive arthropathy on long-standing treatment with diclofenac slow release presented with abdominal pain, diarrhoea and anaemia. Upper gastrointestinal endoscopy and barium series were unremarkable. Five months later he presented with a right iliac fossa mass. Investigations revealed discrete ulceration in the terminal ileum and caecum confirmed at laparotomy. Other possible causes of intestinal ulceration were excluded suggesting a diagnosis of non-steroidal anti-inflammatory drug associated intestinal ulceration.
 Descriptors/Keywords: HUMAN NON-STEROIDAL ANTIINFLAMMATORY DRUG TOXICITY
 SEROPOSITIVE ARTHROPATHY CECAL
 Concept Codes:
 *12508 Pathology, General and Miscellaneous-Inflammation and Inflammatory Disease
 *14006 Digestive System-Pathology
 *18006 Bones, Joints, Fasciae, Connective and Adipose Tissue-Pathology
 *22005 Pharmacology-Clinical Pharmacology (1972-)
 *22012 Pharmacology-Connective Tissue, Bone and Collagen-Acting Drugs
 *22504 Toxicology-Pharmacological Toxicology (1972-)
 10060 Biochemical Studies-General
 12512 Pathology, General and Miscellaneous-Therapy (1971-)
 Biosystematic Codes:
 86215 Hominidae
 Super Taxa:
 Animals; Chordates; Vertebrates; Mammals; Primates; Humans
- < Current Contentsサーチワード >
 04494683 Genuine Article#: KJ964 Number of References: 28
 Title: ILEO-CAECAL ULCERATION ASSOCIATED WITH THE USE OF DICLOFENAC SLOW RELEASE
 Author(s): HUDSON N; WILKINSON M J; SWANNELL A J; STEELE R J; HAWKEY C J (Reprint)
 Corporate Source: UNIV NOTTINGHAM HOSP, QUEENS MED CTR, DEPT THERAPEUT/NOTTINGHAM NG7 2UH//ENGLAND/ (Reprint); UNIV NOTTINGHAM HOSP, QUEENS MED CTR, DEPT THERAPEUT/NOTTINGHAM NG7 2UH//ENGLAND/
 Journal: ALIMENTARY PHARMACOLOGY & THERAPEUTICS, 1993, V7, N2 (APR), P 197-200
 ISSN: 0269-2813
 Current Contents Journal Announcement: CC CLIN, V21, N20; CC LIFE, V36, N20
 Language: ENGLISH Document Type: ARTICLE
 Geographic Location: ENGLAND
 Subfile: SciSearch; CC LIFE—Current Contents, Life Sciences; CC CLIN—Current Contents, Clinical Medicine
 Journal Subject Category: PHARMACOLOGY/TOXICOLOGY
 Abstract: A 38-year-old man with seropositive arthropathy on long-standing treatment with diclofenac slow release presented with abdominal pain, diarrhoea and anaemia. Upper gastrointestinal endoscopy and barium series were unremarkable. Five months later he presented with a right iliac fossa mass. Investigations revealed discrete ulceration in the terminal ileum and caecum confirmed at laparotomy. Other possible causes of intestinal ulceration were excluded suggesting a diagnosis of non-steroidal anti-inflammatory drug associated intestinal ulceration.
 Identifiers—Keywords Plus: ANTI-INFLAMMATORY DRUGS; TERM INDOMETHACIN; INTESTINAL PERMEABILITY; BOWEL PERFORATIONS; HEMORRHAGE; THERAPY; INFANTS; ULCERS
- ← BIOSISの索引は個々の薬物、疾患名よりも主題の概念を表すConcept Codesと、生物系統を表すBiosystematic Codesに重点が置かれている。この例では「消化器系の病理学・14006」「結合組織・骨・コラーゲン作用性薬物の薬理・22012」「医薬毒物学・22504」などのかなり広い概念で索引が付与されている。
- ↑ Current Contentsではこの例には存在しないが著者のキーワードと、索引者によるKeywords Plusがあるが、語の統制や階層関係などは充分ではない。

図3 BIOSIS・Current Contents サンプルレコード

のセンテンスにあるので「同一センテンス中」の検索を行えば、この論文は回答集合に含まれず事実関係とも一致する。

これに対して EMBASE では図 2 のサンプルレコードのとおり、副作用を示すリンク用語「adverse drug reaction(ae)」は索引された薬剤のうちジクロフェナクのみを組み合わされているので「プレドニゾロンの副作用」という段階でこの論文は検索されてこない。ところがもし仮にジクロフェナクによる潰瘍の他にプレドニゾロンによる副作用としての浮腫が報告されていれば、DDF・EMBASE の索引はそれぞれ図 4 のようになり、EMBASE では「adverse drug reaction(ae)」のリンク用語が組み合わされた薬剤はジクロフェナクとプレドニゾロンの 2 つになり、症状の方では「side effect(si)」が組み合わされたものは「intestine ulcer」と「edema」の二つになる。これらの原因薬剤と症状の各二つずつを関連付けることはでき

ないので、プレドニゾロンによる副作用「prednisolone—adverse drug reaction—ae」と副作用として発生した潰瘍「intestine ulcer—side effect—si」は同一論文中に存在するため、実際の因果関係とは異なるにもかかわらず検索ノイズとして回答集合に含まれてしまう。これは MEDLINE でも同様で、サブヘディングによってその索引語の役割までは規定できてもそれらの関連付けまで行うことはできない。

BIOSIS, CC では薬物名の索引もなく、サブヘディングのような役割を表すものも設定されていないのでさらに単純な語の組み合わせによる検索しか期待できない。このような複雑な検索意図に対応できる DDF はむしろ特殊なデータベースであり、通常のデータベースでは EMBASE のように索引語どうしの関係までは規定できず、このようなノイズは機械検索の限界として許容し、最終的には人間の目での判断が必要になってくる。こ

こではほんの 1 例についてその索引を比較したにすぎないが、各データベースの索引方針や特徴を考慮において検索式をたてる必要があることは言うまでもない。

4. 利用例

4. 1 複数データベースの検索

各データベースでそれぞれ対象となる主題領域が異なるので、検索式を組み立てる前にまずデータベースの選択が必要である。最近では各オンラインシステムにおいてその中の各データベースでの該当件数を参照できるインデックスファイル (例: DIALOG の DIALINDEX, BRS の CROS) や複数ファイルを同時に検索できる機能 (例: DIALOG の ONESEARCH, STN のマルチファイル検索) およびその結果に対し重複レコード、つまり各データベースへの収録対象となった

< Derwent Drug File 仮定索引 >

COMMON TERMS: FT=IN-VIVO FT=CASE-HISTORY FT=CASES
LINK TERMS:
01 TR=DICLOFENAC ST=DICLOFENAC TR=VOLTAROL-RETARD ST=VOLTAROL-RETARD
TR=RHEUMATOID TR=ARTHRITIS TR=JOINT-DISEASE ST=ULCER
ST=GASTROENTEROPATHY ST=CECUM ST=ILEUM ST=DYSPEPSIA FT=ANTIINFLAMMATORY
FT=INTESTINE FT=ANTIINFLAMMATORIES FT=ANALGESICS
FT=PROSTAGLANDIN-ANTAGONISTS FT=15307-86-5 RN=DICLOFENC FT=TR FT=ST
02 TR=AZATHIOPRINE TR=RHEUMATOID TR=ARTHRITIS TR=JOINT-DISEASE
FT=CYTOSTATICS FT=IMMUNOSUPPRESSIVES FT=446-86-6 RN=AZATHIOPR FT=TR
03 TR=PREDNISOLONE ST=PREDNISOLONE FT=CIBA-GEIGY TR=RHEUMATOID
TR=ARTHRITIS TR=JOINT-DISEASE FT=CORTICOSTEROID FT=CORTICOSTEROIDS
FT=50-24-8 RN=PDNISOLON FT=TR FT=ST ST=EDEMA
04 TR=CIMETIDINE TR=DYSPEPSIA TR=ULCER TR=GASTROENTEROPATHY TR=CECUM
TR=ILEUM FT=ANTIULCER FT=INTESTINE FT=ANTIHISTAMINES-H2 FT=ANTIULCERS
FT=GASTRIC-SECRETION-INHIBITORS FT=51481-61-9 RN=CIMETIDIN FT=TR

< EMBASE 仮定索引 >

DRUG DESCRIPTORS:
*diclofenac—adverse drug reaction—ae; *diclofenac—drug therapy—dt; *
nonsteroid antiinflammatory agent
sustained release preparation; prednisolone—drug therapy—dt; azathioprine
—drug therapy—dt; barium; prednisolone—adverse drug reaction—ae
MEDICAL DESCRIPTORS:
*intestine ulcer—surgery—su; *intestine ulcer—side effect—si; *
arthropathy—drug therapy—dt
ileum; cecum; abdominal pain—side effect—si; diarrhea—side effect—si;
anemia—side effect—si; gastrointestinal endoscopy; human; male; case
report; adult; oral drug administration; article; edema—side effect—si

図 4 Derwent Drug File と EMBASE の仮定索引

```

***** Welcome to STN International *****
=> FILE PHAR
FILE 'PHAR' ENTERED AT 09:14:28 ON 31 AUG 93
COPYRIGHT (C) 1993 PJB Publications Ltd. (PJB)

=> E DOPAMINE D1 /PHCD 5
E1 1 DOPA-D5-AG/PHCD
E2 1 DOPA-D5-AM/PHCD
E3 0 --> DOPAMINE D1/PHCD
E4 16 DOPAMINE D1 AGONIST/PHCD
E5 20 DOPAMINE D1 ANTAGONIST/PHCD

=> S E4
L1 16 "DOPAMINE D1 AGONIST"/PHCD

=> SELECT CHEM L1 1-16
E1 THROUGH E50 ASSIGNED

=> DIS SEL
E1 2 1,2-BENZENEDIOL, 4-(2-(BUTYLPROPYLAMINO)ETHYL)-, HYDRO
    CHLORIDE/BI
E2 2 57464-71-8/BI
E3 2 65340-90-1/BI
E4 2 71800-28-7/BI
E5 2 SK&F-38393/BI
E6 1 (-)-(3R*,4R*)-3-(2-CHLORO-3-HYDROXYPHENYL)-4-(3,4-DIHY
    DROXYPHENYL)PIROLIDINE HCL/BI
E7 1 (1R,3S)-1-AMINOMETHYL-5,6-DIHYDROXY-3-PHENYLISOCROMAN
    /BI

< 途中省略 >
E45 1 SK&F-38393-C/BI
E46 1 SK&F-77434/BI
E47 1 SK&F-82958/BI
E48 1 SK&F-87516/BI
E49 1 SK&F-R-87516/BI
E50 1 SULFAMIDE, N,N-DIETHYL-N'-(8,ALPHA.)-6-METHYLGOLIN-
    8-YL)-/BI

=> SET INTERPRET OFF
SET COMMAND COMPLETED
=> SET POSTINGS OFF
SET COMMAND COMPLETED

=> QUE E2-E5,E10-E13,E16,E22-E34,E36-E38,E40-E49
L2 QUERY CREATED

=> FILE HCA

```

***** * * * * *
 開発薬「ルトリヂン」へ入る
 「PHARMAPROJECTS」へ入る
 「Dopamine D17」エントリの作用
 を持つ開発薬を検索する
 16品目が回答として得られた
 16品目の化合物名称,CAS登録番号
 をどアップし,そのリストを表示する
 検索の途中経過表示を
 簡略化するコメント
 先にどアップした50語から
 次の検索語として使用する
 ため,治験番号とCAS登録番号を選択し質問式を作成

図5 クロスファイル検索・マルチファイル検索・重複除去の例(1)

THE CA FILE FOR 1967-PRESENT WITH HOUR-BASED PRICING.
 FILE COVERS 1967 - 21 Aug 93 (930821/ED) VOL 119 ISS 8.

=> S L2

L3 167 (57464-71-8/BI OR 65340-90-1/BI OR 71800-28-7/BI OR "SK&F-38393"/BI OR 100999-26-6/BI OR 105016-30-6/BI OR 112431-33-1/BI OR 112460-74-9/BI OR 20012-10-6/BI OR 62717-42-4/BI OR 62751-59-1/BI OR 64795-23-9/BI OR 74627-35-3/BI OR A-68930/BI OR A-77636/BI OR "ANTIPARKINSONIANS, NOVO"/BI OR CI ANERGOLINE/BI OR CQ-32-084/BI OR CY-208-243/BI OR "DOPAMINE AGONISTS, LILLY"/BI OR "DOPAMINE AGONISTS, PARKE-DAVIS"/BI OR E-4101/BI OR ETISULERGINE/BI OR FPL-63012AR/BI OR FP L-65447AA/BI OR "PBDA, CORNELL"/BI OR "PBDA, SKB"/BI OR "PROPYLBUTYLDOPAMINE, CORNELL"/BI OR "PROPYLBUTYLDOPAMINE, S KB"/BI OR "SK&F-38393-A"/BI OR "SK&F-38393-C"/BI OR "SK&F-77434"/BI OR "SK&F-82958"/BI OR "SK&F-87516"/BI OR "SK&F-R-87516"/BI)

=> S L3 AND (STRU?(W)ACTIV?)/BI

L4 8 L3 AND (STRU?(W)ACTIV?)/BI

=> FILE MEDLINE ENBASE BIOSIS

FILE 'MEDLINE' ENTERED AT 09:27:29 ON 31 AUG 93

COPYRIGHT (C) 1993 U.S. National Library of Medicine (NLM)

FILE 'ENBASE' ENTERED AT 09:27:29 ON 31 AUG 93

Copyright (C) 1993 Elsevier Science Publishers B. V. Amsterdam.
 All rights reserved. (ELSEVIER AMS)

FILE 'BIOSIS' ENTERED AT 09:27:29 ON 31 AUG 93

COPYRIGHT (C) 1993 BIOSIS(R)

=> S L2

FILE 'MEDLINE'

L5

924 (57464-71-8/BI OR 65340-90-1/BI OR 71800-28-7/BI OR "SK&F-38393"/BI OR 100999-26-6/BI OR 105016-30-6/BI OR 112431-33-1/BI OR 112460-74-9/BI OR 20012-10-6/BI OR 62717-42-4/BI OR 62751-59-1/BI OR 64795-23-9/BI OR 74627-35-3/BI OR A-68930/BI OR A-77636/BI OR "ANTIPARKINSONIANS, NOVO"/BI OR CI ANERGOLINE/BI OR CQ-32-084/BI OR CY-208-243/BI OR "DOPAMINE AGONISTS, LILLY"/BI OR "DOPAMINE AGONISTS, PARKE-DAVIS"/BI OR E-4101/BI OR ETISULERGINE/BI OR FPL-63012AR/BI OR FP L-65447AA/BI OR "PBDA, CORNELL"/BI OR "PBDA, SKB"/BI OR "PROPYLBUTYLDOPAMINE, CORNELL"/BI OR "PROPYLBUTYLDOPAMINE, S KB"/BI OR "SK&F-38393-A"/BI OR "SK&F-38393-C"/BI OR "SK&F-77434"/BI OR "SK&F-82958"/BI OR "SK&F-87516"/BI OR "SK&F-R-87516"/BI)

HCA77体で先の検索式を
 実行し,「構造活性相関」
 の語とAND演算する

```

FILE 'EMBASE'
L6      197 (57464-71-8/BI OR 65340-90-1/BI OR 71800-28-7/BI OR "SK&F-
38393"/BI OR 100999-26-6/BI OR 105016-30-6/BI OR 112431-33
-1/BI OR 112460-74-9/BI OR 20012-10-6/BI OR 62717-42-4/BI
OR 62751-59-1/BI OR 64795-23-9/BI OR 74627-35-3/BI OR A-68
930/BI OR A-77636/BI OR "ANTIPARKINSONIANS, NOVO"/BI OR CI
ANERGOLINE/BI OR CQ-32-084/BI OR CY-208-243/BI OR "DOPAMIN
E AGONISTS, LILLY"/BI OR "DOPAMINE AGONISTS, PARKE-DAVIS"/
BI OR E-4101/BI OR ETISULERGINE/BI OR FPL-63012AR/BI OR FP
L-65447AA/BI OR "PBDA, CORNELL"/BI OR "PBDA, SKB"/BI OR "P
ROPYLBUTYLDOPAMINE, CORNELL"/BI OR "PROPYLBUTYLDOPAMINE, S
KB"/BI OR "SK&F-38393-A"/BI OR "SK&F-38393-C"/BI OR "SK&F-
77434"/BI OR "SK&F-82958"/BI OR "SK&F-87516"/BI OR "SK&F-R
-87516"/BI)

FILE 'BIOSIS'
L7      222 (57464-71-8/BI OR 65340-90-1/BI OR 71800-28-7/BI OR "SK&F-
38393"/BI OR 100999-26-6/BI OR 105016-30-6/BI OR 112431-33
-1/BI OR 112460-74-9/BI OR 20012-10-6/BI OR 62717-42-4/BI
OR 62751-59-1/BI OR 64795-23-9/BI OR 74627-35-3/BI OR A-68
930/BI OR A-77636/BI OR "ANTIPARKINSONIANS, NOVO"/BI OR CI
ANERGOLINE/BI OR CQ-32-084/BI OR CY-208-243/BI OR "DOPAMIN
E AGONISTS, LILLY"/BI OR "DOPAMINE AGONISTS, PARKE-DAVIS"/
BI OR E-4101/BI OR ETISULERGINE/BI OR FPL-63012AR/BI OR FP
L-65447AA/BI OR "PBDA, CORNELL"/BI OR "PBDA, SKB"/BI OR "P
ROPYLBUTYLDOPAMINE, CORNELL"/BI OR "PROPYLBUTYLDOPAMINE, S
KB"/BI OR "SK&F-38393-A"/BI OR "SK&F-38393-C"/BI OR "SK&F-
77434"/BI OR "SK&F-82958"/BI OR "SK&F-87516"/BI OR "SK&F-R
-87516"/BI)

TOTAL FOR ALL FILES
L8      1343 L2

=> S L8 AND STRU? (W) ACTIV?
FILE 'MEDLINE'
L9      21 L5 AND STRU? (W) ACTIV?
FILE 'EMBASE'
L10     14 L6 AND STRU? (W) ACTIV?
FILE 'BIOSIS'
L11     3 L7 AND STRU? (W) ACTIV?
TOTAL FOR ALL FILES
L12     38 L8 AND STRU? (W) ACTIV?

=> SET DUPORDER
ENTER FILE OR (DEFAULT): FILE
SET COMMAND COMPLETED

```

HCA および3つの
71800-28-7/BI, 計477件
の検索結果に対し
重複除去を行う

HCA から1件目のみ表示

重複除去
結果の表示

各71800-28-7/BIから1件目のみ表示

重複除去を行った検索と同様の検索
を、MEDLINE, EMBASE, BIOSIS71800の
71800-28-7/BI検索で実行する

重複除去を行った結果を
71800-28-7/BI順に並べるためのコマンド

図6 クロスファイル検索・マルチファイル検索・重複除去の例(2) (図5の続き)

一次情報が同じものを検出または除去する機能もある。これらを有効に活用すれば、各データベースの収録内容や重複の度合いについて詳細な比較をすることにそれほど重要な意味はなく、その場合に応じて該当件数を参照しながら必要なデータベースを選択すればよい。

また複数のデータベースの使い方としては、類似タイプのデータベースを同時に検索するだけでなく、一つのデータベースで得られた結果を次のステップの検索語に利用するという方法もある。STN の REGISTRY ファイルと CA ファイルのように最初から横断検索を意識して作られたデータベースもあるが、レコード中の指定フィールドのデータをピックアップし検索式を作成および編集する機能を持つホストシステムもある（例：DIALOG の MAP コマンド、STN の SELECT コマンド）。

図5・図6に STN での SELECT コマンドを利用した横断検索、マルチファイル検索、重複除去機能を使用した例を示す。これは「開発中の DOPAMINE D1 アゴニストの構造活性相関に関する論文」を検索したものである。検索対象ファイルは MEDLINE, EMBASE, BIOSIS, CA (ここでは検索語に対する課金がなく接続時間、出力データに依存する HCA ファイルを使用) であるが、「DOPAMINE D1 アゴニスト」という総称に該当するものを検索する場合、MEDLINE, EMBASE については索引語が階層的に構成され下位概念まで含めて検索できるようなシステムにはなっているので該当する索引語があれば検索は容易である。しかしどちらも "DOPAMINE AGONISTS" を示す概念はもっているがその下に "DOPAMINE D1 AGONIST" "DOPAMINE D2 AGONIST" などという細分類はまだ作られていないので単純には対応できない。また原論文に個々の薬物名のみで記載され詳細な作用メカニズムにまで言及されていないことも考えられる。そこで開発薬リストのデータベースである PHARMAPROJECTS を利用する。このデータベースは薬効分類のほかにその薬剤の作用メカニズムを

示す項目が最近追加され、先の文献データベースに比べてその分類が細かいのでこのデータベースで「DOPAMINE D1 アゴニスト」を検索し、その結果得られた薬物名を次の検索語として使用する。該当する16薬剤の集合に対し、SELECT コマンドを用いて50の名称およびCAS登録番号がリストアップされた。ここから化学名を除き次の検索に有効だと思われる治験番号とCAS登録番号を検索式としてシステムに記憶させた後、データベースを移ってその検索式を実行させれば個々の薬剤名をいちいち入力しなくてもよい。薬剤名でヒットする集合と構造活性相関を示す語のAND演算による検索の結果CAファイルから8件の論文が得られた。より網羅性をあげるため、さらにMEDLINE, EMBASE, BIOSISでも同様の検索を行うが、三つのファイルを同時に指定し検索を行うこともできる。それぞれのファイルで21件、14件、3件のヒットが得られたが、先のCAファイルの8件も含めて各ファイル間での重複が考えられるため、これらに対して重複除去処理を行う。その結果12件のレコードが重複とみなされて除去され、計34のレコードがユニークな回答として得られた。

4.2 検索結果の統計処理

検索によって論文のリストを出力するだけでなく、検索結果集合に対し、統計処理を施すことができるオンラインシステムもある。1993年初めにDIALOGに登場した統計処理コマンドRANKを用いての検索例を図7に示す。SCISEARCHファイルは引用文献を収録し、かつ検索も可能なデータベースである。例としてアドレナリン α 1レセプターのクローニングに関する論文76件に対し、引用文献の項目について統計処理を行い、その上位10位を表示させた。このランキングは先の76件の論文に多く引用された論文、つまりこの研究においてコアとなる論文と推察することができる。SCISEARCHファイルでは引用文献の著者と収載雑誌・巻号頁しか直接出力することはできないが、あらためてこれらの論文を書誌事項から検

File 434:SCISEARCH(R) 1974 - 9307W4
(c) 1993 ISI Inc.

S1 1012 (ALPHA1? OR ALPHA()1? ?)(1N)RECEPTOR?(1N)ADRENER?
S2 2387 (ALPHA1? OR ALPHA()1? ?)(2N)ADRENOCEPTOR?
S3 76 (S1 OR S2) AND CLONING?

SCISEARCHファイルで「α1レセプターのクローニング」に関する論文を検索する。
76論文が回答として得られた。

?RANK CR CONT

Started processing RANK
Completed Ranking 76 records
Press ENTER to view the TOP 50 terms
or enter a number N to view the top N terms
or >N to view terms with more than N terms
or enter ALL to view all terms

?10

Enter title for continuous output
or press ENTER for current title option
?(リターン)

DIALOG RANK Results

RANK: S3/1-76 Field: CR= File(s): 34
(Rank fields found in 74 records - 1750 unique terms)

RANK	No. Items	
No.	Ranked	Term

1	46	COTECCHIA S, 1988, V85, P7159, P NATL ACAD SCI
2	43	SCHWINN DA, 1990, V265, P8183, J BIOL CHEM
3	31	LOMASNEY JW, 1991, V266, P6365, J BIOL CHEM
4	27	MORROW AL, 1986, V29, P321, MOL PHARMACOL
5	24	MINNEMAN KP, 1988, V40, P87, PHARMACOL REV
6	19	HAN C, 1987, V32, P505, MOL PHARMACOL
7	19	HAN C, 1987, V329, P333, NATURE
8	18	KOBILKA BK, 1987, V238, P650, SCIENCE
9	18	MINNEMAN KP, 1988, V33, P509, MOL PHARMACOL
10	17	GROSS G, 1988, V151, P333, EUROPEAN J PHARM

引用文献をランキングし
上位10位を表示

?RANK PY

Started processing RANK
Completed Ranking 76 records
DIALOG RANK Results

RANK: S3/1-76 Field: PY= File(s): 34
(Rank fields found in 76 records -- 5 unique terms)

RANK No. Items

No.	Ranked	Term
1	39	1992
2	19	1991
3	14	1993
4	2	1988
5	2	1990

---end of results---

出版年でランキング

各出版年ごとの論文数が
表示される
(1992年発行は39件etc)

<上位10位の引用論文のタイトル (MEDLINEファイルより別に調査したもの)>

- 1 Molecular cloning and expression of the cDNA for the hamster alpha1-adrenergic receptor.
- 2 Molecular cloning and expression of the cDNA for a novel alpha1-adrenergic receptor subtype.
- 3 Molecular cloning and expression of the cDNA for the alpha 1A-adrenergic receptor. The gene for which is located on human chromosome 5.
- 4 Characterization of alpha 1-adrenergic receptor subtypes in rat brain: a reevaluation of [3H]WB4104 and [3H]prazosin binding.
- 5 Alpha 1-adrenergic receptor subtypes, inositol phosphates, and sources of cell Ca2+.
- 6 Heterogeneity of alpha 1-adrenergic receptors revealed by chlorethylclonidine.
- 7 Alpha 1-adrenoceptor subtypes linked to different mechanisms for increasing intracellular Ca2+ in smooth muscle.
- 8 Cloning, sequencing, and expression of the gene coding for the human platelet alpha 2-adrenergic receptor.
- 9 Comparison of alpha 1-adrenergic receptor subtypes distinguished by chlorethylclonidine and WB 4101.
- 10 5-Methyl-urapidil discriminates between subtypes of the alpha1-adrenoceptor.

← 76件中この論文を引用したものが46件
以下同様

図7 統計処理コマンド利用例

索し、そのタイトルを参照するとその結果は明らかである。またこの76件の論文をその出版年について統計処理すると、この研究がいつ頃を中心になされているかといったこともわかる。このほか、収載雑誌、著者などのランキングを行うと、注目すべき雑誌、研究の中心人物などの推測も可能である。

5. お わ り に

以上医学・薬学・生物学領域でのおもなデータベースの概観とその利用法について述べてきたが、その規模・内容・範囲は幅広いものであり、紙面の都合もあって代表的なものを紹介するにとどめた。

次に著者の立場で今後のこの分野における要望などについて私見を加えたい。まず第一として、この分野のデータベースはその索引構造等が複雑なものもあり、情報検索の専門家でないといえなせない場合もある。しかし、近年CD-ROMなどの急速な発展により、エンドユーザ検索も普及してきた。商用データベースにおいても簡単な検索だけでなく専門的な検索がエンドユーザにもでき

るようなメニュー方式や、各ホストシステム間のコマンドの違いを意識する事なく検索できるようなソフトウェアの充実も望まれる。

第二に現在提供されているデータベースの質の向上も期待したい。本稿で紹介したように、データベースによってその索引方針には特徴があり、それぞれのよさはあるが、広さ深さにかなりの差がある。しかし内容的な意味における索引についてはできるだけ詳細な情報がほしいものである。なぜなら著者抄録は検索を意識して書かれたものではなく、その中にはかならずしも必要とする情報が盛り込まれていない場合も多いので、それを補うような索引を付与してほしい。

最後に新しいデータベースに対する期待も少し述べておく。最近ディレクトリデータベースが充実してきている。たとえば、開発中のものを含む医薬品について、医薬品単位でその詳細なプロフィールがわかるような総合的なデータベースの実現を願う。また最新の情報を得るために学会のプログラムやその抄録に関するデータベースの充実を希望する。

こうしたものの実現によって生命科学分野のデータベースのますますの発展を期待したい。

参 考 文 献

- 1) 西川隆也・勝山麗, 医学・医薬品情報, 情報管理, 29(6)491-509(1986)
- 2) 平松賢二, EMBASE plus の評価, 1990 第27回情報科学技術研究集会発表論文集, 195-201(1991)
- 3) 笹川統, Ringdoc の新しいシステム Unified Data Base, 薬学図書館, 29(2/3) 99-117(1984)
- 4) 大澤礼子, 副作用情報のオンラインデータベース-SEDBASE, 薬学図書館, 33(2)125-133(1988)
- 5) 安藤聡子・浦川親紘・渡辺智佳子, 治験薬情報の比較-PHARMAPROJECTS と明日の新薬, 薬学図書館, 35(3)155-162(1990)
- 6) 松谷貴己, 複数システムのデータベース利用評価(2) Chemical Abstracts Service の提供する化学物質辞書ファイルの比較, 情報の科学と技術, 42(12)1173-1183(1992)
- 7) 大澤礼子, 複数システムのデータベース利用評価(6) EMBASE, 情報の科学と技術, 43(6)573-580(1993)
- 8) 原田智子, 複数システムのデータベース利用評価(5) MEDLINE, 情報の科学と技術, 43(4)381-395(1993)
- 9) 原修, 複数システムのデータベース利用評価(4) BIOSIS Previews, 情報の科学と技術, 43(2)165-173(1993)