EXCERPTA MEDICA

機械化プログラム

松村多美子*

1. ま え が き

医学知識の記録は、紀元前3世紀に、スメリヤ人やエジプト人が、クレイタブレットやパピルスに種々の薬の処方や病気の扱い方などを書き記した事に始まる。これらの記録が理論的に分類されていたという事実の最も古い裏づけは、Edfuにある古代エジプトの図書館の壁に彫られた本の目録であると言われている¹⁾。

時代が進んで、図書の個人的蒐集から、更に公共図書館の設立に至る迄には、長い歴史の流れを経るわけであるが、殊に科学の分野に於いては、近代に入って文献の量は増加の一途をたどり、遂に科学、医学等の専門図書館の誕生をみるに至った。priceはその著書 Science since Babylon の中で、約10の科学雑誌しか存在しなかった1750年以来、世界の自然科学関係の雑誌及び論文の数は、半世紀毎に10倍の増加を示したと述べている²⁾。医学生物学の分野もこの現象の例外ではない。又、論文の増加は、とりもおおざり研究者の世界的増加を意味するものに他ならない。しかしながら、この文献数及び研究者、利用者への驚くべき増加にもかかわらず、それに対応すべき医学・生物学関係図書館の発展及び情報管理制度の開発が平行して伴わなかったことも事実である。

コンピューターがこの問題にとり組むため

の新しい有力な武器としてとり上げられてから数年になるが、その図書館に於ける実際の利用の先がけとなったものの一つが、アメリカ国立医学図書館で行なわれている MEDLARS (Medical Literature Analysis and Retrieval System) である。同図書館が、1879年以来発行している Index Medicus の質的改善を直接の主要目的として1964年1月に始められた MEDLARS は、人間の知性とコンピューターの機械力とを組み合わせ、絶え間なく増加、複雑化する医学生物学関係の情報を処理しようと試みたものである。現在迄に、約750,000の文献が蓄積され、1969年迄には1,000,000以上になると予定されている。この蓄積された情報から、MEDLARSによって、月刊及び年刊 Index Medicus を始め10の recurring bibliographies (逐次書誌)、NLM Current Catalog 目録カード、そして数え切れぬ demand bibliographies (要求書誌) が作られている。

一方、Excerpta Medica は、1946年に設立されて以来、抄録、各種国際会議シリーズの出版等によって、医学及びその関係分野に於ける情報活動の国際的主動力の一つとなっている。主要出版物である抄録は、基礎、臨床及び特殊の34部門から成り、約70名に上る専門家のスタッフによって編集される。又、年間20,000冊に上る受入れ雑誌は、すべて全巻マイクロフィルムにとられ、このようにして1960年以降蓄積された論文の数は16,000,000頁に上り、マイクロフィッシュにして

* Tamiko MATSUMURA,

慶応義塾大学医学部 北里記念医学図書館

50万枚をこえるものとたっている。これらの情報を、必要な時に何時でも即時利用することができるようにするためと、又、抄録始め各種出版物の作製を迅速かつ能率的に行なうためとを目的として、Excerpta Medica 機械化のプログラムが進められて来たが、今日そのプランがほぼ完了し、1968年秋には実施の運びとなった。プログラムの概要は、A Brief Outline of the Excerpta Medica Foundation Program for the Automated Storage and Retrieval of Medical Information⁸⁾として発行されたので、そのあらましを紹介すると共に、既存の MEDLARS と比較してみよう。

2. Excerpta Medica 機械化プラン

Excerpta Medica の機械化プランでは、雑誌が受入れられてから最終的産物である抄録ができ上がるまで6つのステップから成り立っている。雑誌が受入れられて、マイクロフィルムに取られたあと、割当て編集者の手を経てそれぞれの論文は、その主題の属する部門別編集者の許に廻される。文献の探索の迅速性を最大の目的とするために、各種の書誌的データは順次コンピューターに入れられそれぞれのレベルで検索が可能のように壮組まれている。

1) 第1段階：参照

ここでは、論文に関する凡ゆる書誌的データ、即ち、雑誌名、著者名、その所属機関名等が与えられる。同時に、Excerpta Medica 34部門の中から、その論文の属する部門のコード番号が与えられる。これらデータは、flexowriter によってコンピューターに打ち込まれる。(附図1参照)

2) 第2段階：分類

前段階で、各部門別にわけられた論文は、ここで更に細かいカテゴリー番号を与えられる。これは、Excerpta Medica Classification System (EMCLASS) と呼ばれ、34部門それぞれについてカテゴリーに細分し、3次元の構造にくみ立てたものである。(附図1参

照) この分類表から、一論文について最高5つのカテゴリー番号が与えられるが、これら番号はコンピューターに入れられ、参照ファイルと一緒にされる。このようにして、雑誌受入れから1カ月以内に、完全な書誌的データを具え医学の専門分野別に分類された文献ができ上がるわけである。この段階に於て、コンピューターはこれら総てのインフォメーションを具えた抄録用紙と索引用紙を打出す。前者は、第5段階である抄録作業のために使われ後者は次の過程である主題索引及び項目索引の係りへ廻される。

3) 第3段階：主題索引

過去数年間に亘って、Excerpta Medica では、独自の索引技術を開発してきた。その結果、現在では、抄録の殆んど大部分について EMCO Index (Excerpta Medica Computer Index) が月刊及び年刊として出版されている。この索引に使われているシソーラスは、Master List of Medical Indexing Terms (MALIMET) と呼ばれるもので、これを使って、論文の主題にマッチする keywords を与えるのが、この段階の作業であり、オリジナルの論文からなされる。(附図3参照)

4) 第4段階：項目索引

前段階の主題索引に比べて、これは限られた数の一般的な性格の概念を表わすためのものであり、従ってそれほど適確ではあり得ない。しかし、性別、年齢、臨床か動物実験かの区別等文献に一応の制約を与えることになる。(附図4参照) このようにして、表題及び項目両索引のデータがコンピューターに入るのは、原文献が受入れられてから2乃至3カ月後ということになり、ここに於いて、高度に適確な主題についての文献検索が可能となる。

5) 第5段階：抄録

索引件名がコンピューターに入った所で、抄録用紙と原論文は抄録作業のためにその係りの許へ廻される。係りはここで、論文について、抄録する価値があるか否かを決定する

が、大体医学文献の 60% が抄録出版されることになる。又同時に、世界各国に散らばる抄録誌の中から論文の主題及び言語に関して最も適当なものを決定する。又、外国語の文献はすべて英語に翻訳される。このようにしてでき上った抄録のテキストは、コンピューターに入れられ、その時までに既に蓄積されている論文の他のデータと共にデータ・バンクに蓄積されることになる。(附図 5 参照) この段階に至る迄には、論文受入れから 4 乃至 5 カ月の時間が経過している。

6) 第 6 段階：主題索引-二次出版物

以上のようにして準備された抄録から、写真コピーが作られ、論文の二次の主題の部門に廻される。ここでは抄録をもとにして、他の分野から見た主題及び項目索引がなされる。(附図 6 参照)

以上見てきたように、コンピューターの利用によって原論文が受入れられてから、4～5 カ月後には最終産物である抄録が完成し、その過程に於て、各レベルで、それぞれ異なった適確性をもった文献検索が可能であるという事が、この Excerpta Medica 機械化プログラムの最大の特徴としている点である。

3. MEDLARS

一方、アメリカ国立医学図書館に於いて行なわれている MEDLARS は、1964 年 1 月、次のような目的をもって始められた。即ち

- 1) 主要出版物である Index Medicus の質を向上し、同時に収録誌の数をふやす。
- 2) 月刊 Index Medicus 発行に要する時間を 22 日から平均 5 日に迄引下げる。
- 3) 論文索引の深さを増す
- 4) Index Medicus と同様な形式・内容を持ち、なおかつ特殊主題分野に限られた書誌類を供給する。
- 5) 将来、単行本や雑誌論文以外の情報導入の可能性を持たせる。
- 6) コンピューターに蓄積された膨大な量の情報の迅速かつ効果的な検索を可能にする。

7) 米国内の他の団体機関によって行なわれている文献スクリーニング活動の重複をさける⁴⁾。

システムのデザインは 1962 年に完了し、2 年間のテスト期間を経た後実施されたが、現在 MEDLARS のために年間約 2,300 の世界の代表的医学生物学の雑誌から約 175,000 の論文記事が索引され、1969 年迄には、1,000,000 件の文献が蓄積される見込みになっている。

MEDLARS は、3 三つの主要な過程から構成されている。即ち、(1)索引及びコンピューター入力。(2)蓄積された情報の検索及び(3)蓄積された情報の出版である。(附図 7 参照) このそれぞれについて少し細かく見ると、

1) 索引及び入力—ここでは専門的索引者の知性と大規模なコンピューターの機械力が巧みに組み合わせられて文献の蓄積を行なう。

即ち、情報蓄積の第一段階に於いて、論文は訓練を受けた索引者によって分析され、主題に従って Medical Subject Headings (MeSH) から最も適切な件名が与えられる。索引作業には、NLM で開発考案された索引データシートが使用され、索引者は件名をそのシートの所定の位置にタイプするが、同時にそのシートには性別、in vitro, in vivo の区別、年齢、地名、人種等が予め印刷されていて、索引者は各論文ごとにこれらの項目を一つ一つチェックし、該当するものには印をつけることになっている。更に、この段階に於いて、論文の著者名、頁づけ等の書誌的データが与えられ、英語以外の論文のタイトルの英訳又はほん字もなされる。

その後、データシートは索引者の手をはなれて入力タイピストの許へ送られ、ここで flexowriter によってペーパーテープの上にコード化され、最も、基本的なユニット・レコードが作られるわけである。一方、flexowriter からは同時に hard paper のコピーが打ち出され、校正係りによって数回に亙る校

正, 訂正がなされた後始めてテープは継ぎ合わせられ, コンピューターに入れられる。このようにして文献の入った磁気テープから, 2つの主要なデータファイルが作られる。一つは Compressed Citation File で, これは後述する demand bibliographies (要求書誌) の検索のために使われる。他の一つは, Processed Citation File と呼ばれ, Index Medicus その他の recurring bibliographies (逐次書誌類) 製作のためのデータファイルである。

2) 検索—この段階で, demand bibliographies (要求書誌) の検索がなされる。医者, 研究者, 図書館員等からよせられた文献検索依頼は, 索引及び機械検索の論理についての十分なトレーニングを受けた検索者によって分析され, 適切な検索のためのデータエレメントが列挙され, そしてロジカルな検索式が作り上げられる。これがパンチカードによってコンピューターに入れられ, カードなり平板紙なり依頼者の希望のフォームで, 又, 著者のアルファベット順なり論文の年代順なりこれも又希望の排列方法で文献が検索される。これは従来の印刷された Index Medicus や種々の書誌索引類によるマニュアルな検索では効果の上らないような種類のトピックに対して行なうものである。本来出版を目的とした検索ではなく, 依頼者の個人的使用のために主として高度に適確な参照を提供するものであるが, 或る種のものは NLM Literature Search として印刷, はん布されている。

3) 出版—この段階で, Processed Citation File から Index Medicus 始め各種の recurring bibliographies (逐次書誌) が編纂され, GRACE (Graphic Arts Composing Equipment) によって印刷される。即ち, コンピューターに蓄積された文献は, MEDLARS によって種々の特殊分野別に編纂され, アメリカ政府の研究機関や科学学術団体によって研究者間に広く配布されている。例えば, Index of Rheumatology は American Rheuma-

tism Association によって, 又, Cerebrovascular Bibliography は National Heart Institute の National Institute of Neurological Diseases and Blindness によって販布されている。この他, 現在合わせて10タイトルの逐次書誌が MEDLARS によって発行されている。

4. SDI の立場から

さてここでこの2つのシステムを比較してみると, 両者ともそれぞれの主要出版物である抄録と索引の質的量的向上と作製のスピードアップを目的とした機械化という点で共通している。しかし乍ら, SDI (selective dissemination of information) の立場から見た場合には, Excerpta Medica は MEDLARS にはないすぐれた点を持っている。即ち, 医学情報の洪水の中にあつて, 資料プロフィールの作製を分類, 件名による主題索引, 及び項目索引の3段階にわけ, それぞれのレベルでの SDI を可能にし, 同時に, 収集・蓄積された資料を後の IR のための蓄積データとして使用する Excerpta Medica の機械化プランは, 情報サービスの2つの面, SDI と IR の両者をかね具えているという点で非常に強力なものと言える。Current awareness として従来あげられているものに, タイトル索引やコンテンツ・シートによるサービスがあるが, いずれも文献の内容把握の点からみて満足すべきものとはいえない。又図書館員が, マニュアル・サーチによって文献を選択的に探索し, 利用者に提供することも行なわれて来ている。しかし, 医学・生物学のように, 資料が日々増加し複雑化する分野に於いては, 利用者プロフィールに応じて, 最も適した情報を少しでも早く提供することが望ましいことは明白である。Excerpta Medica 機械化プランでは, 資料の受入れから約1カ月以内で情報の EMCLASS による分類がなされ, 書誌的データと共に探索可能になる。次いで, シソーラス MALIMET に

よる主題索引がなされ、更に第3段階である項目索引が完了するのは資料の受入れから2～3カ月以内といわれている。即ち、資料プロフィール作製を3段階に分けることによって、情報サービスの最新性と同時に選択的最適性をも確保しようと試みたものである。情報の遡及的検索 (retrospective search) の点からみれば、MEDLARS は決して劣ってはいない。しかし、医学・生物学関係の研究者達は過去の研究の成果をまとめて知りたいたいと思うと同様に、或いはそれ以上に、現在自分が研究している分野又は主題についての最も新しいそして適格な情報を入手したいと望むであろう。このような利用者側の要求に応ずるためには、IR と平行して SDI が効果的に進められて始めて適切な情報サービスと言えるのではないであろうか。そして、この点に於いて、Excerpta Medica の機械化プ

ランが、従来のシステムに見られなかった特色を持ったものとして一層注目されるわけである。

5. おわりに

以上、このたび公表された Excerpta Medica 機械化プランのあらましを紹介し、MEDLARS と比較し、殊に SDI の立場からその特色についてふれた。確かに、既存の各種システムの長所短所を研究して開発されたこのプランは、理論的に非常によく組織されていることは事実である。しかし、いかなるシステムも実際に動かした場合に、種々の予期せぬ結果が生じるものである。1968 年秋に実施を予定されているこの機械化プランが実行に移された後の成果は、医学分野の情報活動に従事するものにとって、大きな期待と注目のまとである。

附図 1 参 照

EFFECTS OF HYPERBARIC OXYGENATION
ON METABOLISM. IV. TIME SEQUENCE OF
BIOCHEMICAL CHANGES AT 5 ATMOSPHERES
100% OXYGEN—Sanders A. P. and Hall I.
H. Div. of Radiobiol., Dept. of Radiol., Duke
Univ. Med. Cent., Durham, N. Car.—PROC. SOC.
EXP. BIOL. MED. (N. Y.) 1967, 125/3 (716-721)

附図 2 分 類 (EMCLASS)

6. RESPIRATION	13. NEUROPHYSIOLOGY
6. 1. Function tests	13. 1. Techniques of investigation
6. 2. Mechanics of breathing	13. 2. Neuron
6. 2. 1. Breathing cycle	13. 2. 1. Conduction
6. 2. 2. Lung volume	13. 2. 2. Synapse.
6. 2. 3. Lung ventilation	13. 2. 3. Electric phenomenons
6. 3. Gas exchange and transportation	13. 3. Reflex systems
6. 3. 1. Lung gas diffusion	
6. 3. 2. Gas trasport	

附図 3 主題索引 (MALIMET)

ATHLETE, AEROBIC WORK, EXERCISE, OARSMAN, RUNNER, SPRINTER, THRO' ATMOSPHERIC PRESSURE, BEHAVIOR, MOTOR ACTIVITY, MOUSE, 735

ATP, HYPERBARIC 02, METABOLISM, 5 ATMOSPHERES, 536

LEUKOCYTE METABOLISM, LEUKOCYTE MITOCHONDRIA, MAMMAL, 619

ATPASE, ACID PHOSPHATASE, AORTA ENZYME, ATHEROGENIC DIET, MAGNES

附図 4 項 目 索 引

001 ADU	234 DRU	765 NEW
002 AGE	235 EPI	766 NOM
003 ANO	236 ETI	767 PRE
004 AUT	237 FOL	768 REV
005 CAS	238 EXP	769 MAL
006 COM		770 FEM

附図 5 抄 録

Fasted male Sprague-Dawley rats (150~200 g) were exposed to 100% oxygen at 5 ATA pressure for 30, 60, and 90 min, and ATP concentration, succinic dehydrogenase activity, % free cathepsin activity, % free ribonuclease activity, acid-soluble and total nitrogen, and tissue buffer capacity were determined on brain, liver and kidney, ATP Concentrations decreased at 30 min, returned toward normal

附図 6 主題索引 (二次出版物)

MESENTERIC ARTERY PERFUSION, ANGIOTENSIN, CALCIUM, SEROTONIN, 568

METABOLIC ACIDOSIS, ACID BASE BALANCE, EXERCISE, RUNNING, 836

ADH, ADRENALIN, BETA ADRENERGIC RECEPTOR BLOCKI METABOLISM, ATP, HYPERBARIC 02, 5 ATMOSPHERES, 536

METEOROLOGY, AEROSOL, ALTITUDE, DUST, RESPIRATORY TRACT, 825

METRAZOL, BRAIN METABOLISM, S 35, SEIZURE, CAT, 634

附図 1—6 は A Brief Outline of the Excerpta Medica Foundation Program for the Automated Storage and Retrieval of Medical Information. p. より

附図 7 MEDLARS フローチャート

CCF : Compressed Citation File

PCF : Processed Citation File

<参考文献>

- 1) Thompson, J. W., Ancient Libraries., Berkeley, Univ. of California Press, 1940. pp. 4~8.
- 2) Price, D. J. deSolla., Science since Babylon., New Haven, Conn., Yale Univ. Press, 1961, p. 97.
- 3) A Brief Outline of the Excerpta Medica Foundation Program for the Automated Storage and Retrieval of Medical Information. Amsterdam, Excerpta Medica Foundation, n. d. [28 p.]
- 4) The MEDLARS Story, U. S. Dept. of Health, Education and Welfare, Public Health Service, 1965, p. 3.