

# M E D I N F O '74 にみる

## 医用データベースの動向

上野晴樹(青学大・理工)

### 1 はじめに

医療における情報処理は、取り扱うデータの種類の利用目的は多様であるが、ファイル処理に集約されると考えられる。したがって、ファイルの統一的管理を可能ならしめるデータ・ベース・マネージメント・システム(以下DBMSあるいは単にDBと略す)の導入の必要性が高いと思われる。しかし一方我が国では、医療情報をもつデータの性質(属性, 質, 量, 構造など)や、DBの利用形態(目的, 処理技法, 応答形態, 利用範囲, 有効性, 障害など)に関する調査・研究が十分でなく、よってDBMSの稼働の実績は、未だないといつてもよいであろう。

MEDINFO '74では、Session 2.1 (Methodology of Health Data Base Development)と、Session 2.2 (Experience with Computer-based Medical Information Systems)で、DB関係の論文が多数取りあげられている。一般に、論文の内容と実体とは多少ずれている場合が多いが、これを差し引いても、医用DBに関する日本と欧米諸国との格差は非常に大ききことを認めざるを得ない。

また、DBMSの頭に医用という形容詞が付いたとき、一般くといつても色々あるが)のDBMSとどこが違うのかについて、色々議論されているが、これらの論文が一つの解答を示していると考えられる。以下に、上記の中からいくつかのシステムを紹介し、多少の考察を加えてみよう。なお、DBの定義には多少ありまいる点があるので、data bank, information system, information registerなど、内容がDBと考えられるものは同義とみられている。

### 2 システム例

#### 2.1 AVAS - 汎用情報評価システム

西ドイツのInstitut für Dokumentation, Information und Statistikで開発され、研究所の研究者およびHeidelberg大学病院の医師によつてルーチン的に利用されているシステムである。

AVAS (Allgemeines Variables Auswertungs-System)は、プログラムに未経験な者が、短時間の簡単な教育で、彼自身のデータも処理し評価できることを目的とした、パラメータ制御型の汎用データ検索/評価システムである。IBM-360/67の下で、TSSで働き、利用者はタイプライタあるいはディスプレイを通して、インタラクティブにシステムを利用できる。

AVASは、4つの主要モジュールCHECK, VARCOR, TABLEL, およびSEARCHからなり、各々次のような目的と使い方をとする。

- CHECK データベースへの入力データをチェックするモジュールで、不正な文字、データのレンジ、不正なコードなどをチェックする。
- VARCOR データベース中の記録をバッチモードで訂正するためのモジュールである。レコードのdeletion, addition, replacementおよびcorrectionが、数値のパラメータを指定するだけで行える。

T A B E L L 汎用の作表ルーチンである。利用者は下に示したような論理式  
で、作表すべきデータの条件を与え、更に作表のパラメータを  
与えることにより、各種の表を得ることが出来る。

ex) (A &  $\neg$  B) / (A & C &  $\neg$  B)

ここで、

A = *abratia mammae*  
B = *post-surgery radiation*  
C = *hormonal treatment*  
D = *cytostatic treatment*

入力は、タイプライタ、ディスプレイあるいはカードで与え、

出力はディスプレイあるいはラインプリンタで行われる。

このモジュールが最も多く利用されている。

以上の3つのモジュールは、バッチモードで処理される。

S E A R C H 利用者向きの簡易検索モジュールである。

このモジュールが対象とするファイルは、シーケンシャル、イン  
デックス、シーケンシャル、ランダムおよび階層構造などさ  
まざまな種類のファイルであり、レコードにも固定/可変長お  
よび固定/可変フォーマットが許されているので、利用者は、  
*data describing module* を使用して、ファイルとレコードの  
記述をしなければならないが、システムの方から対話的に問い  
まくるので容易である。

ex) 検索指定の例

D I A G N O S I S = 170 . A N D A G E < 16

出力は、条件に合ったレコードの数と、そのレコードのリスト  
であるが、T A B E L L にも連結されているので、作表できる。

このシステムは、*Integrated Scientific Support System* の第1ステップと  
して、1973年に稼働を開始した。第2ステップは、統計ルーチン S T A P R  
O (対話型) との連結を行う予定である。西ドイツがん研究所で使用される予定  
である。

## 2.2 . W A M I S

W A M I S (The Vienna General Medical Information System) は、オ  
ーストリアの Vienna 大学医学部で使用されており、開発に10年を要した。現在の  
のコンピュータ・システム構成は、IBM-370/145 (256KB), D  
I S K (6), M T (3), C R, C P, L P, O M R, および端末としてライ  
トペン付きディスプレイが14台と、タイプライタが36台である。また中央検  
査室には S Y S T E M / 7 (32KB) があり、中央のCPUと連結されている。

W A M I S は、患者に対する臨床活動の促進、合理化と、医学研究のサポート  
を主な目的としている。

W A M I S は、次の6つのモジュールで構成されており、患者データ・バンク  
(*patient orientated data bank*) に接続されている。

(1) 日常臨床活動のサポート・システム……省力化がねらい、各種レポート

(2) ドキュメンテーション・システム……各種データ入力のサポート

- (3) クリニカル・ラボ・システム (WIELAB) ----- SYSTEM/7 を使用
- (4) 医学診断システム ----- *diagnostic proposals* を得る, グループ代数の応用
- (5) 医用検索システム ----- 病歴情報を得る. バッチ/対話モードが可能.
- (6) 医用統計システム ----- (5) と連結している. バッチ/対話モードが可能.

### WAMIS の DB

CODASYL/DBTG の提案に従って、4つのレベルの階層構造ファイルで構成されている (図1)。ただし、これは参照用 (*intrinsic DB*) であり、この他に日常の新ししデータの記録ができるもの (*daily DB*) がある。前者は時々後者でアップ・デートされる。

データのリンクージュは患者 ID 番号でなされる。患者の生涯を通じて不変の、14桁の英数字で構成されている (PIF)。しかし、この ID 番号は桁数が多く不便であるので、日常処理には、1年以内有効な6桁の一連番号である "working no." を使用している (WNF)。

これらの階層構成には、検索の効率×重複をさけることが考慮されている。たとえば、WNF から3レベルで PDS (診療日順にならんだ患者データ・ファイル) に達するが、HDF (診断番号、治療番号等の創用番号ファイル) でセレクトされるので、PDS の一部を見るだけでよく、処理時間が大いに短縮される。なお、障害時のリカバリーは、当日の朝にさかのぼって、*log-files* によりバッチモードで行われる。

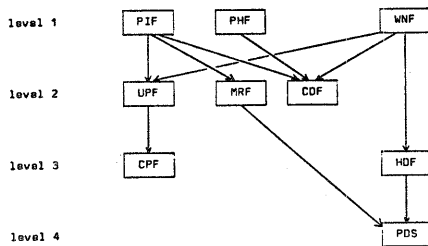


Fig. 1. Structure of the databank. PIF, Primary identification number file; PHF, phonetic file; WNF, working number file; UPF, surname and unchangeable personal data file; MRF, medical reference file; CDF, critical data file; CPF, changeable personal data file; HDF, highly-condensed data information file; PDS, primary data set.

### 2.3 4病院共同利用DB

このシステムは、医療の質の向上、医療事務と診療の効率化、研究の便宜を目的として、英口の United Cambridge Hospital, University College Hospital, Charing Cross Hospital および St Thomas' Hospital が協力して作成したものである。1967年に研究が開始されたが、目的達成のために次の3段階がとられた。才1は、各病院が同一会社のハードウェア (Rank Xerox の Sigma 6) を導入し、才2は、共同のソフトウェア班が共通のファイル構造をもったDB設計を行い、才3に、2つ以上の病院が協力してアプリケーションの開発に当たったことである。

DBはCODASYL/DBTGに従っているEDMS (Extended Data Management System) を使っている。共同利用システムであるので、通信機能に対して特に考慮されており、十分な2次元メモリとともに、障害からの自動回復機能、トランザクションのQueueの処理×型の識別×データベースバッティング機能をもたせてある。また、DBへはMulti-dimensional accessが可能であるとともに、全てのトランザクションは高水準言語で取り扱うことができる。

DBの設計に際して考慮されたのけ次の3点である。

#### (1) 項目 (items) 間の関係

これには、1:1 (患者とID), 1:M (患者と検査), M:M (患者と医師) および無関係 (患者間) の関係がある。

#### (2) 項目の記憶媒体への写像

項目はDB上に物理レコードとして記録され、一方利用の場合にはプログラムで論理レコードとして扱うが、両者は1:1に対応している。従って、応用の場合には、いくつかのDBレコードに關係する。

- (3) 項目とトランザクションとの關係  
この向にはm:mの關係がある。

また、処理も高速化するために、(1)ランダム編成、(2)慎重に配慮されたトリイ構造、(3)適度なリダンダンシー、(4)バッチ処理に対する前SORTの配慮などが考慮され、その結果、たとえば患者の来院時の両合せの80%は2回以内のディスク・アクセスで済んでいる。図2にDBのファイル構造を示す。

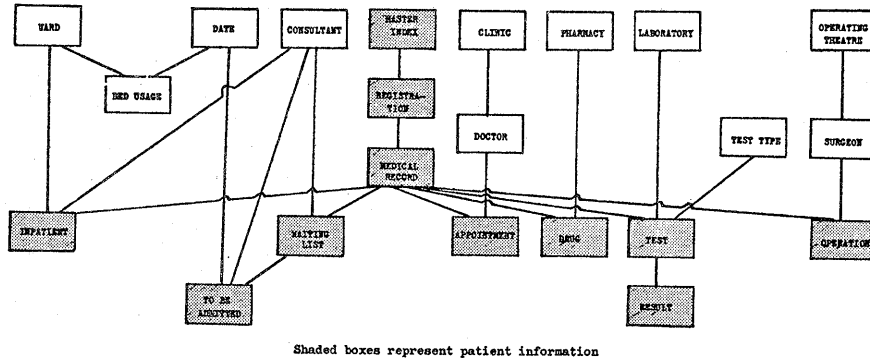


Fig. 2. **Global Database**. This highly simplified figure represents the actual database of 80 records and sheets.

### DBのドキュメント管理

共同利用型DBにおいては、ドキュメントの標準化が不可欠である。本プロジェクトでは、DB設計の初期の段階に、トランザクションの定義、論理レコードの定義および物理レコードの定義に關する、ドキュメント管理用のDBがEDMSを用いて作られ、Transaction Analysis Report, Data Dictionary Report, Physical Analysis Report および Data Base Reference Manual などのレポートを出力している。

このシステムは1974年にフル稼働を開始したが、cost/benefitも満すには数年を要する予想である。

### 2.4 その他のシステム例

南ア連邦のRed Cross War Memorial Hospitalでは、ミニコンをフロント・エンド・プロセッサにして、Cape Town大学の大型コンピュータ(Uminac 1106)にon-lineで連結し、経済的にDBMSを構成している。DBはDMIS 1100を使用し、中央のコンピュータの使用時間の制限により、現在はリモート・バッチ的に使用しているが、TSSによりリアルタイム処理も可能である。データは患者を中心に関層構造をもったダイナミックなトリーで表現され、構造の拡張は容易である。現在はこのシステムで、各種のレポートを作成している。

米国ではNIMHの補助で、6州にあたる100以上の施設で共同利用するMIS(Multi-State Information System)を開発した。中央のコンピュータは、Rockland State Hospitalの研究所に設置されているIBM-360/50とバッチ・ソフトウェアおよびタスクシェアリング用の360/44である。このシステムは、精神病患者の臨床と研究用に利用されている。比較的大きな施設にはカードリーダと

ラインプリンタのついた端末IBM2780が設置されているが、小さい施設は手紙を使用しなければならぬ。DBMSはDLE1が使用されており、データの多様性、更新の多さ、システムの拡張性および利用形態にバッチからリアルタイム処理が要求されることから、ファイル構造としては *indexed sequential hierarchical file structure* が用いられている。DBのデータの利用には、修得が極めて容易な利用者向き言語GALS (General Alphabetic Listor) およびSTAR GEN (Statistical Report Generator) が準備されている。

米国のNebraska大学のメディカル・センターでは、医療記録自動化に汎用DBを利用する場合の可能性を研究した。これは、英国で開発された"E-books"が手作業を前提としており不便であるので、コンピュータ化するためである。汎用DBMSとしてMRI Systems社のSYSTEM 2000が選ばれた。このシステムは医療では未だ実績がないが、米国において4巻目の利用者数をもっている。DBへのリトリフやアップデートは、バッチ/対話モードとも可能であり、各データ要素に可変長を許す *hierarchical tree* のデータ構造をとっている。また高速検索のための *inverted file* も可能である。このシステムは、英語表現に近い利用者向き簡易言語の(1) Immediate Access Feature と、FORTRAN/COBOLを親言語とする(2) Procedural Language Feature, 集計機能つきのリポート生成用(3) Report Writer Feature および磁気テープを対象とする(4) Sequential File Feature をもっている。使用予定のDBを設計し、FORTRANを用いてデータを生成してシミュレーションを行なったところ、応答速度と運用コストに満足できる結果を得た。Nebraska RMPの補助で設置中である。

その他にもミニコンPDP-10を用いたオンラインDB(仏)、医用DBのための簡易検索言語と高速化のための文法(独、仏)、DB定義の方法(米)、*free text* の *encoding / decoding* を行えるDB(加)などの報告があるが、参考文献を参照されたい。

### 3 おわりに

医用DBには各種の形態が考えられるが、臨床業務の効率化と医学的研究の両方を目的とする場合には、CODASYL/DBTG型の汎用DBMSを中心に、医療利用者向きの簡易言語および端末をまわりに付加するやり方が平均的と思われよう。

Cost/benefit を満たす手段としては、ミニコンをフロント・エンドに用いる方法や、データ構造/ファイル構造を単純化して、中型あるいはミニコンにインポートする方法が考えられる。また、特定の目的に対するDBの場合には、米国の共同利用システムに見られるような方法が効果的であろう。

Free text の処理機能を持たせることは、カルテのコンピュータ化が一つの理想であるように、医療では極めて重要であるが、用語やコードの標準化が極めて困難な現状にあるので、かなり先の問題として残さよう。

いすいにしても、医療側のニーズを調査、分析するとともに、技術的に振り下げた研究とともに、実際に試行実験することが不可欠である。

(文献) Proceedings of MEDINFO '74