

## 講演

## 医療データベース\*

古川 俊之\*\*

## 1. はじめに

保健医療は社会の高度化と健康の価値観の向上に伴って、先進諸国共通の国家的問題として取上げられ、具体的な対策の一つとして医療の再編成によるシステム化の実現が急がれている。

医療のシステム化が必要となった背景を分析すると、医療の需給両面に要因を見出すことができる。医療への需要は質的にも量的にも急速に増大しつつあるが、その主因は、住民の保健意識の変化に伴う潜在需要の顕在化と、人口構成の老化化であって、前者は病気の治療のためだけでなく、病気の予防、早期発見、健康の増進など包括的な医療を積極的に要請し、後者は成人病や慢性疾患の増加となって医療需要の増大を生じている。この他、公害病や災害病の増加も見落すことができない。

一方、医療の供給面に関しては、医学の進歩により医療技術はますます高度化、複雑化、巨大化し、診療内容が濃密化した反面、医療従事者の養成や医療施設の増設は需要の増加に大きく遅れていて、需給の差が開く一方になっている(図-1)。この結果、住民の要望に応じて良質の医療を円滑に提供することが容易ではなくなり、ある種の社会的不安をも生みつつある。こうした現状への対策の一つとして限られた医療資源を最も効果的に活用するためのシステム化が、国家的及び地域レベルで多くの角度から検討されるようになったが、これを技術的に見た場合、医療情報の管理と利用を効率化することが最も重要な課題とされ、その中核となる技術として、医療データ処理システムの構成、なにかんずく医療データベースの開発に期待が寄せられている。

以下、データベースの概念を簡単に述べ、次いで医療情報の特徴、医療用データベースの事例を紹介しながら、医療用データベースに関する問題点を指摘した

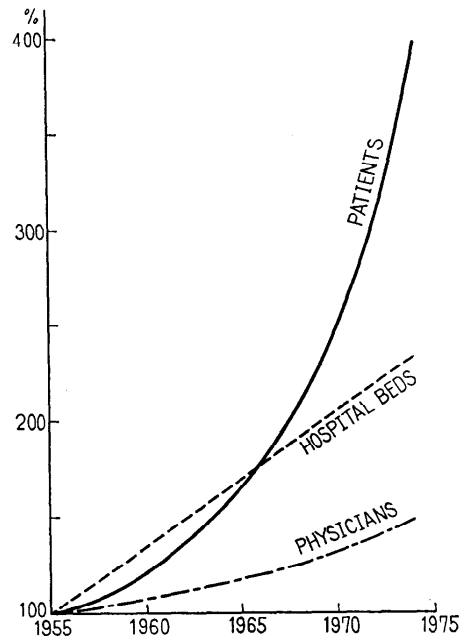


図-1 わが国における医療需給の年次推移

い。

## 2. データベースの概念

データベースは、単に、データを集めるだけではなく集まったデータを有機的に結合し、統合されたデータをさまざまな利用者が、それぞれの目的に応じて利用できるように、ソフトウェア及びハードウェアが工夫された「データベース操作システム」を備えていることが特徴である。従来のファイル管理方式であると、一つの利用目的あるいは1人の利用者一つのデータファイルが従属して保管され、これが他の利用目的や他の利用者に利用されることは技術的に困難であった。その結果、重複したデータファイルが保管されるという無駄が生じたり、ファイルの更新や検索のためのソフトウェアも個別的で共同利用することは容易ではなかった。

\* 情報処理学会創立 15 周年記念大会招待講演 (昭和 49 年 12 月 6 日)

\*\* 大阪大学医学部第一内科



創立 15 周年記念大会で講演中の筆者

これに対しデータベースでは、後述するようにデータの登録に際し、規準に従って「データ構造」が記述されて格納されるために、システム内におけるファイル間の編成が可能で、重複したデータの格納が避けられ、また、データベース利用のための特別な言語を用意することによってファイルの利用と管理を容易にしている(図-2)。この他、データベース方式は、従来のデータ管理方式に比し、データやプログラムの管理、

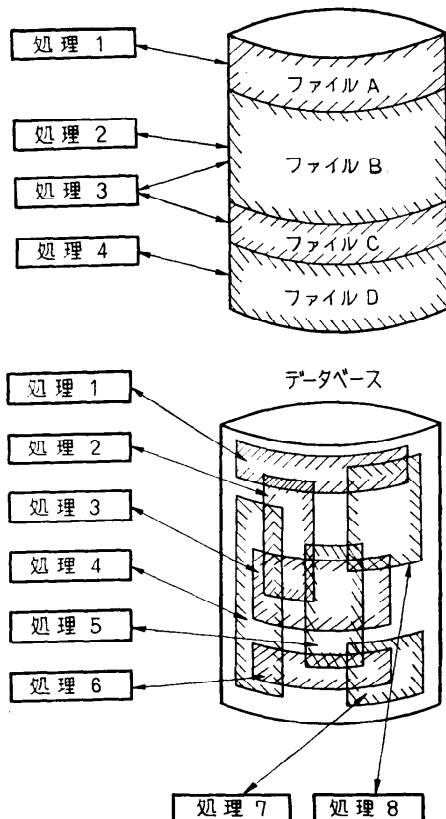


図-2 従来のファイル管理(上)とデータベースによる管理(下)

表-1 従来のファイル管理方式とデータベース方式との比較

	従来のファイル方式	データベース方式
プログラムとファイルの関係	従属	独立
プログラムとファイルの管理	困難	容易
データの重複	多い	少ない
業務の拡張	困難	容易

機密保持が容易であること、業務の変更や拡張に対し柔軟性があるといった優れた特徴を有している(表-1)。

データベースのこのような利点は、格納されるデータの「データ構造」が明確にされていることと、「データベース言語」と呼ばれる利用に便利な特別な言語が用意されることによって達成される。

2.1 データ構造の記述

データベースでは、データはその構造の記述とともに登録されている。データ使用の場合も、データ構造を記述することによってデータを引出すが、この利点は、データの利用目的に応じ、登録されているデータ構造を検索して必要なデータファイルを作成し得ることである。

一般に、ある目的に使われるデータは階層的に、「項目」、「集団」、「エントリ」、「ファイル」といったように分類することができるが、こうして分類されたデータ間の論理的な関係をデータ構造と呼ぶ。データベースでデータ構造を記述する場合は、データ構造をスキーマで表わしておき、これをデータ構造定義用の言語で記述する。MUMPSにおける患者ファイルのデータ構造をスキーマとして記述すると、これが階層構造をもつことが判る(図-3(次頁参照))。データ構造には、このような階層構造の他、単純な線構造や、より複雑な網目構造のものもある。

記述されたデータ構造は、システムで処理されると、システム内にデータの格納とともにデータの構造が定義される。

2.2 データベースのため言語方式

データベースでは、上記のデータ構造の記述や、ファイルの更新、問合せ及び報告書作成に便利のように専用の言語方式が備えられている。

この言語方式には大別して、親言語方式と独立言語方式の2通りがある。この両者の厳密な区分は難しいが、親言語方式は、COBOL, PL/I, アセンブラ言語などの手続き型言語の機能の上に、データベース利用に適した機能が付加されたシステムであり、独立言語方式は手続き型言語と関係なく、それ自体で独立して

閉じたシステムである。一般には、独立言語方式では、あらかじめ条件と動作との組合せを固定し、パラメータの変更のみで使用可能にしているのでプログラムとして書く量は少なく、プログラマでない人達にも容易に使用し得るようになっているが、その代償として応用の範囲は親言語方式より限定される。どのような言語方式が採用されるかは、そのデータベースの目的によって考慮される。

3. 医療情報と医療用データベース

以上述べたデータベースの有する利点は、医療情報のシステム化についても当てはまると思われるが、実際に医療データベースの設計が行なわれる際には、医療情報をよく分析し問題点を把握しておく必要がある。

3.1 医療情報の複雑性

医療情報が他の領域と異なることは、一口で言えば、非常に多種多様であることといえる。医療情報は、1) 患者、住民の健康状態に関して発生するもの、2) 住民の所属する社会や環境に関したのもの、3) 医療従事者に関して発生するもの、4) 医療施設や医療機器、物品に関したのもの、5) 医学文献や医学研究に関したのもの、など多くの種類があり、しかもこれらの情報が、1) 地域的に分散して発生することはもちろん、病院内においても部門別に分れ、2) それが時間的に不規則に発生すること、さらに、3) そのデータの型は、数値データ、定性的データ、波形データ、画像データ、記述的データなどが混然としたものであり、その様式も定まっていない、ことなどによって一層複雑なものになっている。

3.2 医療情報の多目的性

一方、これらの医療情報はさまざまな目的に、さまざまな人によって使用される。使用目的によって大別してみると、1) 患者、住民へのサービス、2) 医師への情報提供による診療の補助、3) 医療機関の管理のための情報提供、4) 医学研究、ということになるが、その各々はさらに細分可能である(表-2)。また、これらの医療情報の流れを、一般の診療所から発生したものに限って図示してみると、その複雑性がよく理解されよう(図-3)。

これらの医療情報の利用は、どれ一つを取り上げてみても重要なものばかりであり、医療水準の向上は、

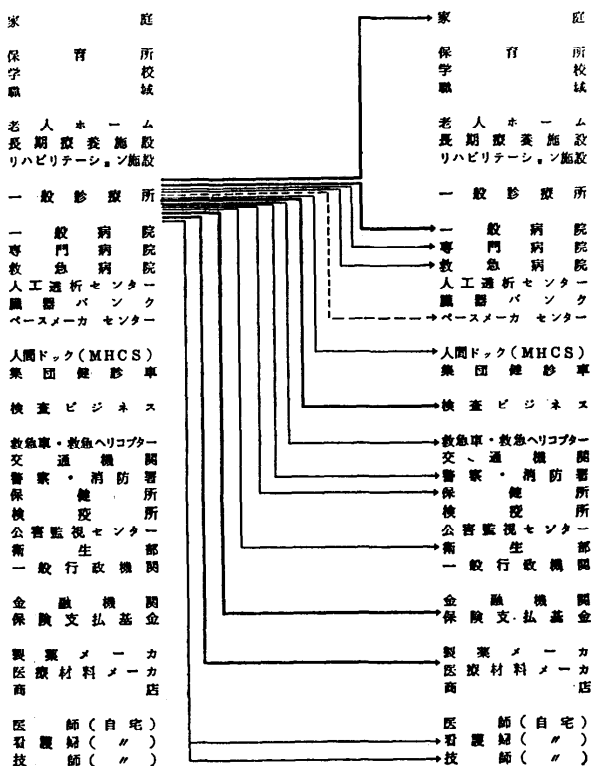


図-3 診療所から発生した情報の流れ

この一つ一つを効率化して行くことによるのみ達成されるといっても過言ではない。

3.3 医療情報管理の現状

現在まで医療情報はどのように管理され、利用され

表-2 医療情報の使用目的

I. 患者、住民へのサービス
病院の窓口業務
予約、診療スケジュール
病院内案内と待ち時間の減少
救急医療
環境汚染情報、災害情報
保健教育
II. 医師への情報提供と診療補助
病歴、検査情報の提供
CAD (Computer Aided Diagnosis)
毒物、薬剤に関する情報提供
血液及び移植用臓器に関する情報提供
III. 医療機関の管理に関するもの
職員の人事管理、給与計算
機器、物品の在庫管理
IV. 医学研究に関するもの
医学統計資料の作成、疫学
薬効の判定、副作用のチェック
医学文献
医学教育の補助

てきたかという、全般的に次のことが指摘されよう。すなわち、1) EDPS 化は、現代社会の他の領域に比し遅れている、2) データは個人の医師単位あるいは、せいぜい病院単位で保管され、他との情報の交換や統合はほとんどなされていない、3) 問題の発生毎に、retrospective または prospective に必要なだけのデータをかき集める努力をするなどである。この結果、たとえば、個人の病歴管理を例にあげて問題点を指摘すると、1) いわゆるカルテは最低5年間保存され、病院の中に山積されているが、検索の方法が備わっていないので必要なデータを収集するのに大変な時間を要し、ほとんどの病院ではこれを利用しようとするものがない、2) X線フィルムや心電図は、それぞれのデータ発生場所だけで保管されることが多いが、カルテ以上に物理的に大きな場所を占有し、その検索も容易ではない、3) 医療機関相互の病歴情報の交換は、医師間の通話と紹介状のみに頼っている、といった状態であり、このため医師が患者の一連の病歴を把握するのに長時日を要したり、医療機関の変更によって同じ検査を重複して行なうことは珍らしくなく、患者に余分な負担をかけるとともに診療機能自体の低下が懸念されることがある。また、地域的に広く多数の症例の病歴を集計し統計的資料を作成することも非常に困難であるから、よりよい診断基準の確立や病気の原因や予後に関する疫学的研究がはかどらなればかりか、薬剤の副作用や公害病の発見が遅れる原因の一つとなっている。

こうした現状を少しでも改善する手段として、病歴(カルテ)をマイクロフィルム化することや、特定患者の病歴のみをパンチカードやマークカードを利用して保管することが試みられてきた。しかし、マイクロフィルムにしてもカルテの山を縮小する物理的效果はあっても検索時間に大差はなく、特定疾患を対象としたEDPSも特定の目的に限局し、広範な利用が念頭に置かれたものではなかった。

#### 3.4 医療情報とプライバシー

このように医療情報の管理が、他の領域に比し立遅れているのは、前述した医療情報の複雑性や経費の問題も大きな理由となっているが、医療情報、特に個人の病気や健康の情報にプライバシーの問題が関係していることも隘路となっている。

診療は、患者と主治医との信頼関係の上に成立し、診療に関する責任は主治医に存在するから、患者からの情報収集とその管理は原則として主治医に委ねられ

ている。一方、病歴の集中管理とその効果的な利用のためには、データが標準化されていることが必要であり、そのために個々の医師に対しデータの収集や記述方法について統制を加えねばならない。こうした統制に対し医師の理解を十分に得る努力とともに、統制や集中管理によって医師の責任や機密の保持が損なわれることのないように十分に保障する必要がある。

### 4. 海外における医療データベースの事例

わが国においては医療データベースが実用化されている例はまだ見当たらないが、海外においては数年前より研究が始められ、実用化の段階に至っているものはいくつかある。この中から、代表的な例を参考までに紹介しておく。

#### 4.1 MUMPS

MUMPS は、Massachusetts General Hospital Utility Multi-Programming System の頭文字をとったもので、1966年に、アメリカのボストン市にあるマサチューセッツ総合病院のコンピュータサイエンス研究室とハーバード大学医学部において、病院情報システムの要求を充つようなプログラミングシステムの開発が始められたのがその起源である。その特徴の主なものを拾って要約すれば以下の通りである。

##### 4.1.1 モジュラーシステム

開発の方針は、病院内の情報を一括して処理する、いわゆるトータルシステム方式ではなく、いくつかに分割して積上げて行くモジュラーシステム方式になっている。モジュラー方式が採用されたのは、1) 初期出費が少ないこと、2) 比較的少数の専門技術者で足りること、3) システムの変更が比較的容易にできること、4) 移動が容易なこと、などの利点を有しているからである。

##### 4.1.2 データファイル

MUMPS のデータファイルは、図-3のような分枝構造のデータ構造で組織化されている。分枝構造のレベルの深さや1階層あたりのデータの数、データの長さには制限がない。

MUMPS は、診療記録の管理に目標の重点が置かれたデータベースであるが、従来の診療記録は、内容、様式とも不規則、冗漫、主観的、不完全であるので、このまま格納されたのでは利用の範囲が著しく限られ意義が薄れるため、病歴記述を Weed 博士によって提唱された「問題志向型の病歴記載法」(POMR)で行なうことに統一されている。POMR は、従来の病歴

記載法が無目的、羅列的であったのを改め、患者の有する問題に出発点を置いて、これに関連する情報を階層的に組立てて記述しようとする病歴記載方法で、データベースの活用を念頭に置いて考案されたものである。

4.1.3 言語方式

MUMPS に用いられる言語は、このシステムだけに適用する一種の独立言語方式であるが、会話形式によるプログラムが容易なようにインタープリタ言語になっている。このため、言語の習得が容易で誰でも簡単に使用でき、多くの端末から TSS によって同時使用することが可能である。また、記述的なデータに適するようにストリング処理機能に重点が置かれている。しかし、逆に、数値演算機能が非常に乏しく、医学研究的な統計処理には不適であること、演算処理時間が大きいこと、種々の医用機器との on-line 化が困難なこと、などが MUMPS の欠点となっている。

4.1.4 ハノーバー医療システム (M.S.H.) におけるデータベース

ハノーバー医療システムにおけるデータベースは、IBM により開発された親言語方式の DL-1 を使用していることが特徴である。

(1) M.S.H. の目標

M.S.H. の開発目標は次の通りである。1) ハノーバー医科大学付属病院の入院患者及びそのスタッフ及び施設に関する情報の収集、統合、検索及び分析を可能にし、患者の治療、病院の管理及び研究活動の合理化に資すること。2) ハノーバー医科大学における教育活動及び学校管理面での補助。3) 科学的データ分析のためのサービス及びシステムの提供。4) 医療情報科学の分野における研究、開発及び教育活動のための素材の提供。

(2) システムの概要

システムは7億バイトの外部記憶、32のCRTターミナル、22のTPターミナル、512kのCPU (IBM-360-67) から構成されていて、CPUのスーパーバイザーが、TSS全体をコントロールし、ソフトウェアとして、基礎科学プログラムの開発、統計プログラムがTSSで行われるほか、データベースのコントロール用にDL-1が用いられていて、TPターミナルを含むオペレーションシステムからデータベースを引くことができる(図-4)。

M.S.H. におけるシステムは、1) 病院管理に関するもの、2) 診療に関するもの、3) 医学研究に関するもの、

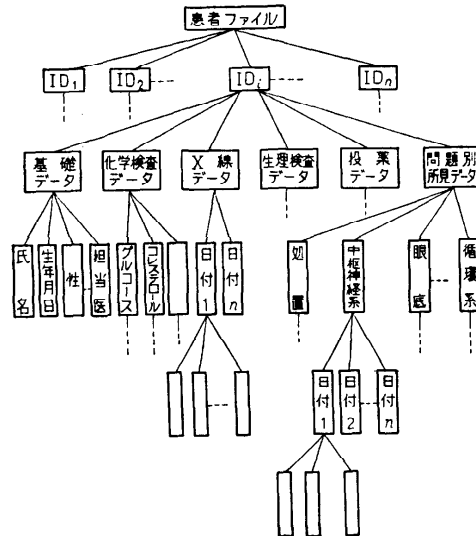


図-4 MUMPS における患者ファイルのデータ構造

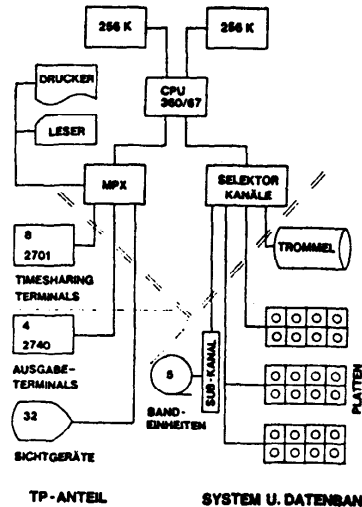


図-5 M.S.H. のシステム構成

の、の三つに大別され、これらがさらに細分されてモジュール化が行われている(図-5)。

入院患者のシステムを例にあげると、全ての入院患者は72の質問票が書記によって記入されることから始まり、この情報が分析、認識、検孔などの技法により分類チェックされ計算機に入力される。この入力には約10分間を要する。この他の入力、36の分野に分けられた入力用紙によりOCR、マニュアルの両方で可能である。患者IDはアルミニウムホイルを用いており、計算機への入力には、このアルミニウムホイ

ルからプリントされたマークを読みとる方式が使用されている。

M.S.H. のデータベースは、入力フォーマットが固定され、主としてコード化されたデータや数値データを対象としているので、MUMPS におけるような記述的データを扱うことには適していないが、統計的資料を迅速に得ることができるため、病院の管理や医学研究に適したシステムといえる。

#### 4.1.5 ストックホルム州地域医療コンピュータシステム

上述の二つの例が一つの病院を対象とした医療データベースであったのに対し、スウェーデンの例は、地域を対象としたデータベースの例として注目される。すなわち、スウェーデンでは、全国が7つの医療行政区に区分され、各々の地域で医療計画が立てられているが、その一つであるストックホルム地区では、150万の人口を対象にした医療用データベースが設けられている。その概略は以下の通りである。

##### (1) システムの基本構造

###### i) 地域全体を網羅

地域の全人口の健康状態を把握し、種々な形の医療資源をこれに対処し得るようにコントロールすることが主な目的である。

###### ii) リアルタイム処理

データの収集を、データが発生した場所でその問題に詳しい専門スタッフによって入力させるため、対話形式でリアルタイムに入力する方式をとっている。

###### iii) モジュール方式

MUMPS におけると同様に、最初から固定したシステムにするのではなく、将来の改良、拡大予測して柔軟性を持たせるためモジュール方式を採用している。

##### (2) ファイルの種類

ファイルは、1) 病歴に関するファイル、2) 医療資源に関するファイル、3) 主ファイルに大別され、このうち、主ファイルには、地域全人口について、出生、死亡に関する情報、医療機関受診情報、X線検査に関する情報と、検索のための個人識別情報が格納されていて、個人の一連の病歴を作成したり、種々の医療統計資料が計算されるようになっていく。

##### (3) ソフトウェア

このシステムは、MIDAS と称する制御システムによってファイルの処理が行われているが、MIDAS の言語は、高水準なプログラミング言語で、フォートラ

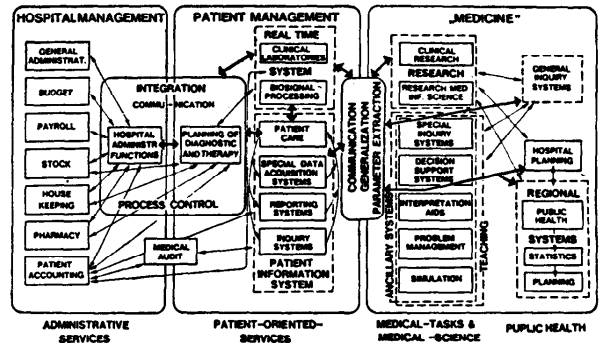


図-6 M.S.H. におけるシステムのモジュール化

に似た用語法をとっており、周辺装置との間のデータ転送のための特殊な命令やファイルの更新、検索、問合せの機能を備えている。

#### 4.2 その他の事例

以上あげた三つの例の他にも同様なデータベースの設置例がいくつかあるが、特殊なデータベースとしていわゆる医用データバンクがある。文献検索に普及して来た MEDLARS (medical literature analysis and retrieval system) や、それにデータ通信を組合せて広域化した MEDLINE は実用化段階に達した例の一つであるが、その他にも、薬剤副作用データバンク、中継コントロールセンタ、臓器移植データバンクなどが、次々と実用期に入りつつあり、今後の医療データバンク開発の重要なモチーフとなることは明らかである。

#### 5. 医療データベース設計上の問題点

以上、医療用データベースの事例をいくつか紹介したが、これらはいずれも特徴を有してはいるものの、なお、改良、拡充の余地を残しており最終的にどのようなデータベースを作ることが医療情報のシステム化と、それによる医療水準の向上に最も適しているかは、当分、実験と議論が繰り返される問題と思われる。そこで、最後に、医療用データベースを計画、設計する上で問題となる基本的な項目をあげておきたい。

##### 5.1 利用目的とデータファイルの対応関係を分析し明確にすること

前述した通り、医療情報はデータの種類、データの発生部位、データの量などにおいて複雑かつ膨大であり、その利用目的、利用者、利用場所も多岐にわたっているから、医療情報を十分に分析して利用目的とデータファイルとの対応関係を整理しておく必要がある。そして、互いにあまり重複のないサブグループに

分けてモジュラー化を行ない、一つ一つを完成して行くことが得策と思われる。

#### 5.2 データの標準化、コード化及びデータの収集法

データベース以前の問題であるが、医療データの標準化は最も重要な課題である。医療用データベースが医学、医療上にどれだけのメリットをもたらすかは、医療情報のどの範囲まで標準化とコード化が拡大し、充実した内容の医療データがデータベースで管理されるかによって左右される。また、異なった場所の医療施設からのデータ収集のための伝送方法や、さまざまな医療用検査機器からのデータ入力方法も並行して研究されねばならない。

#### 5.3 データベース言語とソフトウェア

医療用データベースは、その性格上、ソフトウェアに特別に要求されることがいくつかある。特に、一般医師やパラメディカルスタッフが容易に使用し得るように「会話型」であること、記述的データに適するように「ストリング処理機能」を有すること、同時に、医学研究用に複雑な統計計算が可能ないように「マルチプログラミング機能」を有することが要求される。し

たがってこれらの機能を満足させる言語方式としては、親言語と専用言語の両機能を有することが望ましいと考えられる。

#### 5.4 機密の保持

医療データベースを計画する際に最も重要なことは、機密をどのようにして守るかということである。ソフトウェアにおいて強力な機密保持機能を備えておくことは勿論重要であるが、データベースを実際に管理運営する人間の資格や、機密保持に関する法的規制の問題など十分に検討される必要がある。

#### 6. おわりに

以上、医療用データベースの概要を述べたが、医療情報が内包する特殊問題、技術上の問題、本文では触れなかったが経費とその負担の問題、プライバシーの問題など今後、検討を要する課題が残されている。しかし、現状の医療危機を回避し、社会的ニーズに対応した医療に改革するには、医療情報のシステム化は不可欠な対策と考えられ、近い将来、わが国においても優れた医療用データベースが設置されることを期待したい。