ABC of Computing

JOHN FOX
PETER ALVEY

COMPUTER ASSISTED MEDICAL DECISION MAKING

In the early days of medical computing the computer was merely a means of organising and storing information, such as records, laboratory results, images, etc. Nowadays computers are being used to help present clinical data in a more meaningful form and even to help in clinical decision making itself.

There are two main approaches to computer assisted medical decision making. The first is a mathematical process in which calculated probabilities are used to describe a condition. In the second a system of rules known as the knowledge base is put together and forms the basis on which the computer can give an analysis.

The rationale for using computers is that it is impossible for one doctor to know everything or even to have complete and unbiased experience of everything in his specialty. Consider the case of a surgical registrar who is presented in the middle of the night with a very ill Chinese man with severe fluctuating pain in the left upper quadrant of the abdomen, associated with localised tenderness and guarding. The patient is restless, uncommunicative, and appears to have mild jaundice—but because of the artificial lighting this observation is not reliable. What is the registrar to do? The gravity of the case requires some action, but what is the diagnosis? Could it be a case of acute pancreatitis with pain in the wrong place? Could it be peritonitis without signs of generalisation? Could it, despite the patient's race, be splenic infarction? He will, of course, call on his consultant's more extensive experience of unusual cases; but how many consultants would recognise from a telephone description a case of Asiatic cholangiohepatitis with situs inversus?

This case (which is not fictitious: the diagnosis made at the time was acute cholecystitis with situs inversus, and the outcome was favourable) illustrates some of the difficulties of medical decision making. Problems of this type are usually approached by reference to various mental checklists followed by reasoning to assess the likelihood of each item on the lists. In general, doctors are good at making responsible decisions about specific items, but when it comes to recalling all of the causes of left upper quadrant pain, or all of the causes of jaundice in Chinamen, the human mind begins to show some of its limitations. Similarly, any senior medical student would be able to describe the effects of situs inversus on the presentation of an abdominal condition, but when one is under pressure in the middle of the night it is not the first thing that comes to mind.

This is where the powers of the computer might usefully supplement the decision making process. The computer can refer to a massive database of information—not just one consultant's lifetime experience—and it can be relied on to reiterate faithfully every item which has been stored in it. A simple information retrieval system for the relevant checklists could be very useful: not for making decisions, but for providing the doctor with the basic

information from which he would make the decision. The process may, however, be taken a step further. If the computer could be provided with a reliable method of reasoning about its stored facts its value would be considerably enhanced. Imagine the power of some future decision aiding program which could execute a rule such as:

IF: A likely diagnosis has not been found with the orthodox rules

THEN: Interchange RIGHT and LEFT and try again

It would never forget about situs inversus.

The main problem in the use of computers for decision making lies not in the storage of large numbers of facts but in the development of a reliable and acceptable method for reasoning about the stored facts. The first approach was to use statistical analysis, and such methods have been extensively developed and tested. More recently, attention has shifted to the use of logic for manipulating facts expressed in qualitative terms. These knowledge based techniques are still in their infancy but they are already showing signs of exciting potential.

Structure of a computer assisted aid

Individual decision making systems vary considerably but most have (a) an information base containing details of the things that happen—for example, diseases, side effects, the number of positive findings in tests, etc; (b) the information about that particular patient; (c) a method of applying the information base to the circumstances of the individual patient.

Mathematical method for using computers in planning investigations

It is often difficult to decide the most efficient way of investigating a patient's complaints. On the one hand, the cost and the invasiveness of the procedure have to be considered while, on the other, there is the problem of the expected information yield, its specificity, and its reliability. The computer can help by measuring the information yield of each test and, for a complete investigational strategy, by calculating the remaining uncertainty at the end of the process—before the patient is touched. Such foreknowledge can help a doctor optimise the plan of investigation. If, for example, he discovers that he will be still very unsure of the diagnosis at the end of the series of investigations he may be able to devise a better strategy. A further advantage is that such investigation planning may lead to a reduced use of investigations which are hazardous, uncomfortable, or have a low information yield. One extension of this basic concept is to incorporate into the decision system information on the cost-benefit balance for each investigation.

Mathematical method for using computers in diagnostic work

$$p[D_i:S] = \frac{p[S:D_i] \times p[D_i]}{\sum_i p[S:D_i] \times p[D_i]}$$

The mathematical probability of a disease D_i , given a symptom or sign S is the initial probability of D_i modified by probability of S if disease were actually present ($S:D_i$)

The most commonly used mathematical technique in diagnosis uses estimates of how frequently symptoms and signs occur in patients with particular diseases. If, for example, a symptom occurs more often in patients with disease A than in those with disease B, then it is more likely that a patient with the symptom has disease A rather than disease B. This concept can be expressed mathematically and a probability assigned to each situation, the numerical values ranging from 1, a certainty, to 0, never.

The doctor seeing his patient usually finds a complex condition with many symptoms and signs, not just one, and it is difficult to associate all the relevant information and arrive at an objective conclusion. A computer can take the basic probabilities from the database, combine them objectively using Bayes's rule and calculate the likelihood of each diagnosis given that particular set of symptoms.

Practical problems with probability methods

Computer-aided diagnosis program: acute abdominal pain

This is an experimental programme to help you make an accurate diagnosis.
CLINICAL JUDGEMENT MUST ALWAYS TAKE PRECEDENCE

Symptoms

Male	Severe
Age 28-29	Nausea
Pain onset LT loin	No vomiting
Pain now LT HF	Appetite OK
Pain now LT loin	No prev indigestion
Other aggrav	No jaundice
No rel factors	Bowels normal
Getting worse	Dysuria
Pain 12h	Prev similar pain
Steady at first	No prev abd surgery
Steady now	Not taking drugs

Examination

Clinician's prediction =

Investigations = Urine analysis plain x ray

Initial plan - (A & E) = Investigation / review

Computer prediction

	0%	100%
Appendicitis		
Diverticularitis		
Perforated ulcer		
Non-specific pain		
Cholecystitis		
Bowel obstruction		
Pancreatitis		
Renal colic		
Dyspepsia		

Probabilistic methods can be programmed on small computers, but a vast amount of information needs to be stored to make the system workable, as each disease may link with several indicants—that is symptoms, signs, and laboratory findings. Each set of indicants should be supported by statistically valid epidemiological data. Even for common diseases, prospective data on many thousands of patients would be necessary to allow valid probability calculations. When all this has been collected even then the data drawn from one population may be invalid elsewhere.

Studies have been done in which experienced clinicians have given estimates of the incidence of symptoms in various diseases; unfortunately these studies did not provide reliable information since it was difficult for the clinicians to be infinitely experienced and totally unbiased. The problem of information is even more complex because many indicants are themselves interlinked—for example, the bronchitic patient who has a productive cough is also likely to have wheezes in his chest just because of the nature of the disease. Unfortunately the work required to assess these higher order associations is so great that it is rarely attempted.

Having overcome the problems of the database, and actually fed the information into the computer, the doctor is still faced with a list of diagnoses each with a probability. This form of information is unfamiliar and open to misinterpretation. For example, the statement "Smith suffers from duodenal ulcer with a probability of 0.75" is either a comment about the expected frequency of duodenal ulcers in a group of patients exactly like Smith, or an index of the strength of the evidence in favour of the hypothesis "Smith has a duodenal ulcer." Either way there are problems in all but the most simple cases, when there is a large difference between the probabilities for each diagnosis. One school of thought suggests that when different diagnoses are close competitors in probability terms one should always select the numerically most favourable option, however marginal the figures. Another suggests that more tests should be done to try to increase the difference in the probabilities.

The probability based computer techniques have now been applied in many areas in medicine (diagnosis, investigation planning, treatment planning) and their potential to help in patient care has been amply demonstrated. Nevertheless, except in a few specialised disciplines these techniques have not become commonplace because they do not tell doctors what they want to know in a way that is useful. How many surgeons knowing that the probability of the patient's condition being caused by a perforated ulcer was 28.35%, would fail to operate? Who could set a probability boundary below which it would be unreasonable to operate? How could a doctor who is overreliant on computers be supported in court?

Knowledge based methods

STRONG ANALGESICS

value: (DEXTROMORAMIDE DIAMORPHINE DIPIPANONE
LEVORPHANOL METHADONE NEPENTHE OXYCODONE
PAPAVERTUM PHENAZOCINE)

RULE 034

If: 1) The current drug is one of: my list of strong analgesic drugs,
or
2) A: The current drug is pethidine, and
3) B: The route of administration of the current drug is one of:
im iv
Then: It is definite (1.0) that the current drug is a strong
analgesic drug

RULE 070

If: 1) The patient has uncontrolled pain, and
2) The frequency of the patient's analgesia is adequate
Then: It is definite (1.0) that the potency of the patient's
analgesia should be increased

RULE 190

If: 1) The potency of the patient's analgesia should be increased
2) The patient's most potent current analgesic is strong-oral,
and
3) The patient is not taking a psychoactive drug
Then: It is definite (1.0) that the following is one of the
recommendations for pain: For a greater analgesic effect you
should increase the dose of your strong analgesic or give it
intramuscularly or add DIAZEPAM or CHLORPROMAZINE

7) Does Fred Smith have dyspepsia ?

** WHY

[ie WHY is it important to determine whether PATIENT-1 has dyspepsia ?]

[1.0] This will aid in determining whether PATIENT-1 is suspected of
having a peptic ulcer.

If

[1.1] 1) Fred Smith has dyspepsia, or
2) Fred Smith has evidence of gastro-intestinal haemorrhage

then

it is definite (1.0) that Fred Smith is suspected of having a
peptic ulcer

[RULE 020]

[back to question 7.....]

** WHY

[ie WHY is it important to determine whether PATIENT-1 is suspected of
having a peptic ulcer ?]

[2.0] .....in order to determine the recommendations for pain.

It has already been established that

[2.1] an anti-inflammatory drug should be recommended for
Fred Smith, and

[2.2] oral drugs are suitable for Fred Smith

Therefore, if

[2.3] Fred Smith is not suspected of having a peptic ulcer

then

it is definite (1.0) that the following is one of the recommendations
for pain: Indomethacin orally

[RULE 081]

[back to question 7.....]

** NO

In the past few years non-mathematical methods have been developed to overcome the limitations of the probability based systems. Like the probability based systems, knowledge based systems have three major components.

The major feature of the knowledge base is that it is composed of small units of information, or rules written in a format which resembles English text, which are used in a way that is similar to human reasoning. A knowledge base may contain several hundred rules. The advantages of this form of structure are that qualitative and quantitative data can be used; the knowledge base is easily inspected and updated; and, perhaps most important, it is in a form familiar to humans.

The program which applies the knowledge base is often called the host program, and whereas the probability based program works by calculating the probabilities, the host works by making a series of logical connections based on the knowledge base. Sometimes even extensive systems cannot solve the problems and then probabilistic methods can be used to give degrees of probability to several possible answers.

At present most knowledge based systems work on a consultation basis and as such are often termed "expert systems." The system asks a series of questions to which the user responds with data describing the patient and his condition, and the system then formulates its suggestions using its inbuilt rules. One important feature of such systems is that the doctor can ask the system to display the reasons behind a question or a piece of advice and can satisfy himself of their validity.

Top: Examples of facts and rules from an advice system for terminal care which is being developed at the Imperial Cancer Research Fund using the host program EMYCIN.

Bottom: The explanation facility of EMYCIN. The user was puzzled when asked if his patient had dyspepsia so instead of answering immediately he entered the command WHY? The second WHY? produced an explanation at the next higher level in the logical process.

Two host strategies used in knowledge based systems

Knowledge based systems are based on either data driven strategies or goal driven strategies. In the "data driven" technique the host program examines the situation data provided by the user and operates all the rules which can be applied to it; this in turn increases the fund of situation data and may allow further rules to operate. The cycle may operate many times until a firm recommendation can be made. The "goal driven" system, on the other hand, starts at the opposite end of the logical process. It considers the various final conclusions that could be made to fulfil its main goals and then examines its rules to see if any intermediate information is necessary. This is then provided by using other rules or by asking the user. As in the data driven technique, the host may repeat the process many times, requiring many generations of subgoals before it collects all the necessary data to confirm or refute its main goals.

In both data and goal driven hosts the power of the total system lies in the comprehensiveness of the knowledge base, not in the sophistication of the host. The host merely organises the knowledge base by bringing into memory the right rules at the right time. One useful feature of goal and data driven systems is that they can operate, with minor modifications, on most types of knowledge base—for example, a terminal medical care advice system and a system for diagnosing rheumatism. The techniques of developing and refining knowledge bases are becoming known as knowledge engineering.

Examples of medical expert systems

1) Name
 ** FRED SMITH

2) What problems are troubling Fred Smith?
 ** DYSPNOEA

3) What is the severity of the dyspnoea?
 ** MILD

4) Does Fred Smith have a bronchospasm?
 ** YES

5) Is Fred Smith taking any medication?
 ** NO

6) Does Fred Smith have pleural effusion?
 ** NO

7) Does Fred Smith have anaemia?
 ** UNKNOWN

8) Does Fred Smith have a mediastinal tumour?
 ** NO

Advice for DYSPNOEA
 Steroids are the most effective treatment for bronchospasm (the long term effects are not important in terminal patients).
 Investigate the possibility of anaemia if its correction would improve Fred Smith's personal or family life.

General Advice for DYSPNOEA
 Oxygen is not recommended (because it does not usually ease the dyspnoea of terminal patients).

MYCIN, which gives comprehensive advice on the treatment of bacterial infections, is the first expert system to be developed, and its performance has been favourably evaluated. Subsequently its host program (EMYCIN) has been made available as an empty shell into which new knowledge bases can be inserted. The development of a system round EMYCIN requires a large computer at the outset, but it can then be rewritten for smaller machines. PUFF, an expert system for interpreting pulmonary function data, was developed in this way. PUFF, which agrees with the experts in 90% of cases, is used in teaching.

CASNET is a system which has been used to provide consultations for patients with glaucoma; it was judged to perform comparably with a panel of experts.

INTERNIST is an experimental system for making diagnoses in internal medicine, and its scale is noteworthy; over 600 diagnoses are stored. In its present state INTERNIST is still experimental and unsuitable for clinical use, but it has already provided an important contribution to medicine by systematising a large volume of medical data.

Conclusions

Artificial intelligence is in its infancy

The problem lies in codifying medical knowledge

Knowledge based systems may help to make a breakthrough in computer assisted decision making

The use of computers for assisting with medical decisions is still a novelty. Probabilistic systems have proved their worth but have not become accepted. Knowledge based systems have introduced a technical approach more natural to the medical decision maker, but it remains to be seen whether effective use can be made of their potential.

Dr John Fox, BSC, PHD, is head of the biomedical computing unit and Mr Peter Alvey, PHD, FRCS, is clinical research fellow, Imperial Cancer Research Fund Laboratories, London.

The illustration of a computer aided diagnosis program is reproduced, by permission, from *Practical Computing* September 1982.