EMERGING TECHNOLOGY

Michael Anbar
Tore Scherstén

COMPUTER ASSISTED CLINICAL DECISIONS: Present Scope, Limitations, and Future

Michael Anbar
School of Medicine, State University of New York, Buffalo, N.Y.

Can computers provide substantial support for the daily decision processes in medical practice? If they can do this, to what extent might they, given appropriate hardware and programs, take over what is generally considered the most important function of the medical practitioner, that of a decision maker? The answer to the first question is a qualified *Yes*. Computers can assist in clinical decision making; however, only very recently has this application of computers reached a practical level, and only in limited areas of medical practice. The answer to the second question is a definite *No*; judging from what has been achieved to date in the use of computer assisted medical decisions (CAMD), the notion of autonomous computerized medical decision making is quite unlikely. In this short report we will review some of these most interesting medical applications of computers, discuss the requirements to be met by an effective computerized medical decision support system and thus try to substantiate these two answers.

From the standpoint of computer programming, there are trivial and nontrivial applications of computers in clinical decision making. The former include the retrieval of specified information from the literature or of relevant data from patient's records, the calculation of drug doses based on pharmacokinetic data or the retrieval of facts on drug interactions. They also include the immediate interpretation of data provided by clinical laboratory analyses on the basis of a fixed

Based in part on "Your Humble Consultant—Computer Assisted Medical Decisions," M. Anbar, Chapter 10 in *Computers in Medicine*, M. Anbar, Editor, Rockville, Maryland: Computer Science Press, 1986 (in press).

set of criteria. Other examples are the prescription of the cancer screening tests appropriate for a given patient based on epidemiological data, or the statistical analysis of clinical data to identify characteristic features or patterns. Another group of computational applications uses conditional probability to assess the gain in diagnostic information or evaluates the utility of a particular intervention compared with alternative approaches.

Computerized clinical protocols constitute a separate type of simple applications of computers. Such programs tell the user to take a certain course of action, following a fixed set of branched logic; computerized protocols are quite helpful in instructing paramedical personnel what to do in a given common emergency, but they are limited in scope. In spite of seemingly providing expert advice, these are little more than data banks which retrieve from their memory a fixed set of instructions following a fixed set of conditional selection criteria.

Although any of these computer applications can be highly useful in providing the user with a recommended rational clinical decision, none of them emulates the human decision process, even in part; they do not prompt the user to improve the quality of the decision by suggesting alternative solutions, or point out gaps in the available information. Most of these computer applications handle data rather than knowledge; even if they use knowledge, such as that contained in the medical literature or in clinical protocols, this knowledge is treated essentially as data without any inference.

More sophisticated computer programs that may have a more profound effect on medical practice are those that handle primarily medical knowledge and directly *infer* from it a prefered diagnosis or a recommended action, such as a particular treatment or the acquisition of additional specific data. Acting in this manner the computer emulates the behavior of a human *expert*—a highly knowledgeable medical specialist. These computer applications can be classified as *artificial intelligence* (AI), the domain in computer science which includes robotics, artificial vision, the understanding of natural language, the use of knowledge bases derived primarily from the input of experts (computerized expert systems), and programs able to learn autonomously and improve performance based on "experience."

There are essentially two elements in any computerized expert system—a knowledge base, and an "inference engine" that uses the knowledge base to provide the user with "expert advise" in response to a given input. The separation between knowledge and inference is essential in any system with a broad range of clinical applications (handling of many diseases or of complex treatment procedures). One may compare the knowledge base of an expert system to human *memory*, and the inference algorithms to human *reasoning*. However, this analogy is far from perfect and its validity cannot be tested in any case, since we know very little objectively about how the human mind makes decisions.

Medical expert systems developed so far can be classified into two different types by the way their knowledge bases are structured and the way their output is inferred. The first type includes *rule-based systems* that follow a declarative grammar-like model whereas the other type follows a hypothesis-forming-and-testing *cognitive model*. The first type can be represented by MYCIN and ONCOCIN, developed at Stanford University, while the best known example of the second type is INTERNIST-I developed in a collaborative effort between the

Universities of Pittsburgh and Stanford. We shall describe first the rule-based approach.

In this approach all the available knowledge is presented in the form of elementary rules, each making a single statement (e.g., if A is true then B is blue, if B is blue than C is red, if C is red than D is open). This rule system allows, therefore, chained inferences (e.g., taking the previous example, if A is true then D is open). It can also be consequent driven rather than antecedent driven (e.g., if D is closed C cannot be red and therefore B cannot be blue and A must be false). If we find a conflict (say, we independently know, in the last example that B *is* blue) then the program questions the user for the evidence ("Is D actually closed or is B actually yellow?"), otherwise the rule must be altered.

The advantage of this declarative approach is that it is simple and easy to understand, it is modular and therefore easy to expand or modify, it is also easy to trace back any conclusion. However, it must be remembered that most of medical knowledge is not presented as categorical rules (try to translate this short paragraph into a set of rules). Also, a rule-based knowledge base has minimal educational value since knowledge is not learned this way. Consequently, this approach is best utilized when dealing with categorical information that contains minimal contexual content, like the grammar of a language. Medical problems that lend themselves to this approach include alternative treatments, risk factors, or drug interaction, which can be readily presented in a "if . . . then . . ." framework. Since this approach is relatively easy to program, it has been used, in spite of its conceptual limitations, in most medical expert systems developed so far.

The cognitive approach, on the other hand, tries to emulate the way physicians think and make their decisions. Essentially it is based on two premises: (1) That knowledge is stored in an associative fashion, e.g., when you remember a certain disease you also remember some of its major symptoms and signs as well as other associated disorders, which manifestation is most frequently associated with this disease, and what are its most common treatments. Likewise, if you remember a certain symptom you remember several diseases which manifest it, and if you recall a certain treatment you associate it in your memory with several diseases. (2) That a decision is the end result of many iterative hypothesis making and testing cycles. For example, you observe a certain symptom, associate it with a certain disease, and then realize that this disease is also associated with several other symptoms and test your tentative hypothesis by looking for them. If you find them, you remember the associated signs and test for them, and so on; if you do not confirm your hypothesis you consider another disease with the same symptom or other major manifestations, and so on.

INTERNIST-I, a computerized diagnostic program developed at the University of Pittsburgh, uses a diagnostic knowledge base of about 550 diseases and syndromes in internal medicine and about 4000 manifestations (history facts, symptoms, and signs). Each of the diseases is associated with 20 to 200 manifestation. The diseases may be linked associatively with other disorders by causual links, predispositions (e.g., genetic), coincidental links (where the causuality is not understood) and organ-system-related links (e.g., detecting an infection in one organ may be related to the same infection in another organ or to a systemic infection). In performing a differential diagnosis, by an associative reiterative process similar to the one described above, the program puts weighting factors

(1 to 5) on the frequency of occurance of a given manifestation in a particular disease, and on the prevalence of diseases with a given manifestation.

The feature of this program is not that it comes up eventually with a differential diagnosis, but that in this process it asks the user for the additional information it needs to reach a satisfactory conclusion. These question and response cycles are not only useful for reaching a correct diagnosis, but they are highly instructive as well, reminding the user of the different characteristics of the disease process. When requesting additional tests, the program selects them in an order that reduces the cost of reaching a satisfactory differential diagnosis, i.e., minimizing side effects, discomfort to the patient, waiting time, as well as monetary costs. A satisfactory diagnosis is defined by the program as one where a certain disease is substantially more likely than any other.

Testing of this program retrospectively by comparison with the complicated clinical cases presented in the MGH clinicopathological conferences, gave reasonably good results. The limitations of this unique system (the only medical expert system with such a wide scope) include, however, the necessity to communicate with the program using a rather arbitrary and restrictive medical vocabulary, the arbitrariness of the knowledge base produced by a group dominated essentially by single expert, the inability to find out readily the reasoning of the program in case the user disagrees with its recommended differential diagnosis, the inability to handle more than one suspected disease at a time even in cases where the patient is known to have multiple diseases, the lack of anatomical, physiological and biochemical associations as well of the temporal development of the diseases, which may provide important diagnostic clues. Another limitation is its accessibility—it is being run exclusively on a computer at Stanford. An improved version of this program, named CADUCEUS, is being developed that promises to remedy some of these shortcomings. Thus, in spite of its impressive history which culminates a 20 year-long effort, this ambitious program is still far from being ready for routine use.

Considering the current performance and limitations of both the cognitive model and the rule-based approaches to medical decision making, we conclude that artificial intelligence might at best provide auxiliary tools for but not replace the human decision maker. Computers function very differently from human minds. While humans are comparatively very slow in retrieving memorized facts and in calculations, they readily outdo computers in image recognition, knowledge acquisition, and reasoning. The use of computers to emulate human experts must, therefore, best utilize the features of computers rather than try to imitate or model human behavior. This may mean that neither the rule-based nor the cognitive model programs will be the programmer's method of choice in the development of medical expert systems in the future.

Any practical expert system that might be adopted by the medical community for routine use must meet the following basic requirements:

1. *Provide rapidly unambiguous response to queries regarding* ***complex*** *clinical situations.* (Straightforward situations do not necessitate computer assistance, except in learning settings.) Such response may be a suggested diagnosis or differential diagnosis, a recommended intervention, or a request for additional information about the patient. The consultation time

regarding a complex clinical problem may take an hour or more, which adds a substantial cost to the use of computer-assisted decision making, and virtually excludes its applicability to emergency medicine or operating room situations, where prompt decisions are required. This time requirement is not likely to diminish significantly in the future, since most of it consists of user input time.

2. *Accept input in a variety of formats and linguistic expressions.* The current restrictions on the semantics of interaction with the computer require the buildup of a user's expertise in the effective use of a particular expert system. A human expert would be much more forgiving and much less likely to give utterly wrong recommendations due to semantic ambiguities.
3. *Have a well defined domain of expertise and goals.* These may be as wide as that of INTERNIST-I or its successor CADUCEUS, which promise to encompass diagnostic problems in all of internal medicine, or be far more limited as is ONCOCIN, which deals only with the treatment regimens of certain common types of cancer. It may be purely diagnostic like INTERNIST-I or AI/RHEUM, which is limited to rheumatological diagnosis, purely therapeutics-oriented like ONCOCIN, or management oriented like ATTENDING, which critiques methods and procedures in anaesthesia, or TIA, which offers consultation on the management of transient ischemic attacks in the brain. To avoid pitfalls the user must realize beforehand the scope and limitations of the program.
4. *Be comprehensive and self-consistent.* This criterion is not readily met. Since the inclusion of knowledge in the knowledge base is subjectively decided on by the designers of the system, any computerized expert system can be only as good as the experts who created it. Any evaluation of the performance of such a system is also of limited validity depending on the criteria used and the experts who evaluated it. Thus the reported "sensitivity" and "specificity" of a diagnostic expert system are not as "hard" statistical data as those of a laboratory test, since it is often difficult to define the "true" diagnosis. In the case of intervention, there may be no "best" method of treatment, since the quality of treatments depends on the values assigned to the different outcomes. To use the rate of agreement with the judgment of different experts or of a panel of experts as a measure of "accuracy" of the system is a recourse open to criticism. In the real world there is room for differences between expert opinions; this option is generally not available in computerized systems, which provide an (arbitrary) "consensus" that is absolutely necessary to keep the system self-consistent.
5. *Provide, on demand, an explanation for the computer's response to a given query.* The latter requirement, which is not readily met by many of the existing clinical expert systems, is not only important from a didactic standpoint, but is essential for establishing the credibility of the computer's suggestions, especially when they seem counterintuitive to the user. An exception to the lack of adequate explanation in most medical expert programs is ATTENDING and its analogs, developed at Yale University, which provide a narrative critique of the user's decision, listing the alternative options, which might have been missed by the user, and spelling out the rationale of the critique.

6. *Have the intrinsic capability to "learn."* Any static expert system is doomed in face of the rapidly changing medical knowledge. This requirement has two different aspects: The updating of the system's current knowledge base, on one hand, and its extention to include more diseases and manifestations or interventions, on the other. Many of the expert systems have "user friendly" "toolboxes" which allow non-programmers to extend their knowledge base. However, whereas the extention must be done by experts, preferably by the designers of the system, the updating to make the existing knowledge base more precise should be also achieved by the program itself during routine use. This should happen whenever the program analysis detects inconsistencies (which have not been resolved by interaction with the user), i.e., the program must learn from its own mistakes. In computerized diagnosis a common mistake is coming up with a wrong diagnosis because the disease of the patient under consideration does not exist in the knowledge base. Ideally, in such cases the program should state its limitations to the user, who might then endorse the expansion of the knowledge base. Without such self-improving routines, medical expert systems may remain research or teaching tools in major medical research centers, rather than widely used clinical instruments.
7. *Be transferable.* In view of the cost and relative slowness of long-distance communication with remote mainframe computers, practical expert systems must be installed in the medical center that uses it, preferably in each clinical department or even on each user's desk; this is possible with the currently available sophisticated microcomputers. However, many of the medical expert systems developed in the last 20 years used older hardware and are not readily transportable. This limitation may be overcome as the demand for such systems makes the time investment in reprogramming worthwhile. However, the decentralization of such systems requires the development of comprehensive documentation plus the built-in "learning" ability of the program, discussed above; to develop these may take several years of substantial effort.

On the other hand, it is likely that expert system "generators," or "shell" programs, recently developed for business and industry and which run on up-to-date computers, will be used by medical experts and will allow a much faster development of new effective medical expert systems. This approach has been successfully applied in the development of the TIA program at the University of Maryland based on a locally developed generic expert system KMS, and in the development of AI/RHEUM at the University of Missouri, based on a production rule program named EXPERT developed at Rutgers University.

It is interesting to note that the knowledge base but not the inference algorithms of INTERNIST-I have been recently transfered to a microcomputer format and became available as an "electronic textbook" of internal medicine, named QUICK.

Another expert system which has been designed from the start to be run on a microcomputer is the Stroke Consultant developed by a collaboration between the Illinois Institute of Technology and the Michael Reese Hospital; this may be a forerunner of a new generation of medical decision aids which take full advantage

https://doi.org/10.1017/S0266462300002890 Published online by Cambridge University Press

of current technology and may eventually bring these tools into the offices of practitioners.

Once these seven criteria are met, there still remains the question of cost effectiveness of consultation with a computerized expert system. The cost of an unsubsidized single consultation is estimated at over $300. This includes the cost of creating and maintaining an updated knowledge base and inference system, and of maintaining and upgrading of the hardware, as well as the time of the user. In comparison with the costs of unnecessary diagnostic tests or of ineffective treatments that might be avoided, the cost of computerized consultation does not seem excessive. This appears to have been recognized by insurance carriers in the case of planning and monitoring cancer therapy with the help of ONCOCIN. This program enables practitioners to provide ambulatory patients effective up-to-date treatment outside the framework of a major research hospital.

The situation may be different in the case of diagnosis. If computers will be used in the diagnosis of only the most complex cases, which are generally limited to tertiary care teaching hospitals, the cost per consultation may increase to a prohibitive level. Part of this cost could be offset by using the same up-to-date knowledge base and possibly similar inference algorithms for continuing medical education. In any case, the economics require the use of each such system by a large number of clinical centers.

A different strategy may overcome some of the economic constraints on the implementation of medical expert systems—making such a system a subset of a clinical data base management system; in other words, incorporating logic analysis capability into the patient record and clinical data base management system that currently exists in most hospitals. This has been done at the University of Utah Medical Center with their program HELP, and may become soon commercially available in several hospital-wide data management systems. The large overhead of the data management system may thus cover the maintenance costs of the decision-making-assist sub-system, and make it available to the hospital's staff.

It must be concluded that in spite of the substantial effort invested over the last 20 years in medical expert systems, and the high visibility of some of these programs, this aspect of computer application to medicine is not yet ready for wide range routine use. The availability of more versatile and less expensive hardware, and of "user friendly" expert system generators now allows the development of numerous diagnostic and management decision-assist systems which are expected to meet most of the criteria outlined above. Some of the knowledge accumulation, sorting, and evaluation necessary for such systems will also be done by computers; this will substantially expedite the process. The intergration of decision assisting systems into commercial clinical data management systems may become the best mechanism for attaining a wide use of these tools. Only then will we find out for sure if these systems are cost effective and fully acceptable by the practicing physician.

BIBLIOGRAPHY

Toward the Simulation of Clinical Cognition: Taking a Present Illness by Computer. S. G. Pauker, G. A. Gorry, J. P. Kassirer and W. B. Schwartz, *Am. J. Med.* 60:981–996 (1976).

Production Rules as a Representation for a Knowledge-based Consultation Program. R. Davis, B. G. Buchanan and E. H. Shortliffe, *Artificial Intelligence* 8:15–45 (1977).

BLAH: A System which Explains its Reasoning. J. L. Weiner, *Artificial Intelligence* 15:19–48 (1980).

INTERNIST-I an Experimental Computer-based Diagnostic Consultant for General Internal Medicine. R. A. Miller, H. E. Pople and J. D. Myers, *N. Engl. J. Med.* 307:468–476 (1982).

A Heuristic Approach to Risk Analysis in Computer Assisted Medical Management. P. L. Miller, Proc. 6th Ann. Symp. Computer Applications in Medical Care (SCAMC). IEEE Comp. Soc., Computer Society Press, Washington DC, 1982 pp 697–701.

Explanatory Power for Medical Expert Systems: Studies in the Representation of Causal Relationships for Clinical Consultations. J. W. Wallis and E. H. Shortliffe, *Meth. Info. Med.* 21:127–136 (1982).

Transferability of Medical Decision Support Based on Bayesian Classification. R. J. Zagoria and J. A. Reggia, *Med. Decision Making* 3:501–509 (1983).

Diagnostic Expert Systems Based on a Set Covering Model. J. A. Reggia, D. S. Nau and P. Y. Yang, *Internat. J. Man-Machine Studies* 19:437–460 (1983).

Expert Systems Research. R. O. Duda and E. H. Shortliffe, *Science 220:261–268 (1983*).

The Epistemology of a Rule-Based Expert System: A Framework for Explanation. W. J. Clancey, *Artificial Intelligence* 20:215–251 (1983).

Adapting a Consultation System to Critique User Plans. C. P. Langlotz and E. H. Shortliffe, *Internat. J. Man-Machine Studies* 19:479–496 (1983).

A Critiquing Approach to Expert Computer Advice; ATTENDING. P. L. Miller, Pitman, 1984.

An Expert System for Oncology Protocol Management. E. H. Shortliffe, A. C. Scott, M. B. Bischoff et al. in *Rule-Based Expert Systems*, B. G. Buchanan and E. H. Shortliffe Editors, Addison-Wesley 1984 pp 653–665.

Critiquing: a Different Approach to Expert Computer Advice in Medicine. P. L. Miller. Proc. 9th Ann. Symp. Computer Applications in Medical Care (SCAMC), IEEE Comp. Soc., G. S. Cohen Editor, Computer Society Press, Washington DC, 1984 pp 17–23.

Introduction: Medical Artificial Intelligence Programs. W. J. Clancey and E. H. Shortliffe in *Readings in Medical Artificial Intelligence: The First Decade.* W. J. Clancey and E. H. Shortliffe, Editors, Addison-Wesley 1984 pp 1–17.

Anticipating the Second Decade. E. H. Shortliffe and W. J. Clancey, in *Readings in Medical Artificial Intelligence: The First Decade.* W. J. Clancey and E. H. Shortliffe, Editors, Addison-Wesley 1984 pp 463–472.

Explanation Capabilities for Medical Consultation Systems. E. H. Shortliffe. *AAMSI Congress 1984*, D. A. B. Lindberg and M. F. Collen Editors, Am. Assoc. Med. Systems and Informatics, Publ. 1984 pp 193–196.

QUICK (Quick Index to Caduceus Knowledge): Using the Internist-I/Caduceus Knowledge Base as an Electronic Textbook of Medicine. M. B. First, L. J. Soffer and R. A. Miller, *Computers Biomed. Res.* 18:137–143 (1985).

The Architecture of the ITT—RH Stroke Consultant. H. Hill, C. L. Curt, B. K. Kozar et al. *Proc. 9th Ann. Symp. Computer Applications in Medical Care (SCAMC),* IEEE Comp. Soc., M. J. Ackerman Editor, Computer Society Press, Washington DC, 1985 pp. 314–319.

An Overview of Methods for Computer Assisted Medical Decision Making. J. A. Reggia and S. Tuhrim in *Computer-Assisted Medical Decision Making*, J. A. Reggia and S. Tuhrim, Editors, Springer-Verlag 1985 pp 3–43.

Maintaining a Knowledge Base using the MEDAS Knowledge Engineering Tools. F. Naeymi-Rad, M. Evans, T. Koschmann et al. *Proc. 9th Ann. Symp. Computer Applications in Medical Care (SCAMC),* IEEE Comp. Soc., M. J. Ackerman Editor, Computer Society Press, Washington DC, 1985 pp 298–303.

A New, Efficient Version of HELP. M. R. Ostler, J. D. Stansfield and A. T. Pryor, *Proc. 9th Ann. Symp. Computer Applications in Medical Care (SCAMC)*, IEEE Comp. Soc., M. J. Ackerman Editor, Computer Society Press, Washington DC, 1985 pp. 296–297.

Evaluation of Medical Expert Systems. J. A. Reggia. *Proc 9th Ann. Symp. Computer Applications in Medical Care (SCAMC)*, IEEE Comp. Soc., M. J. Ackerman Editor, Computer Society Press, Washington DC, 1985 pp 287–291.

Issues in the Evaluation of Artificial Intelligence Systems in Medicine. P. L. Miller. *Proc. 9th Ann. Symp Computer Applications in Medical Care (SCAMC)*, IEEE Comp. Soc., M. J. Ackerman Editor, Computer Society Press, Washington DC, 1985, pp 281–286.

New Directions for Expert Systems: Examples from the ONCOCIN project. L. Fagan. AAMSI Congress 1985, A. H. Levy and B. T. Williams Editors, Am. Assoc. Med. Systems and Informatics, Publ. 1985, pp 183–186.

Knowledge Representation, Clinical Action and Expert Systems. L. J. Kohout and W. Bandler in *Knowledge Representation in Medical and Clinical Behavioral Science*. L. J. Kohout and W. Bandler, Editors, Abacus Press 1985.