Report 82-37
Stanford -- KSL

Scientific DataLink

Computer-Based Clinical Decision Aids:
Some Practical Considerations.
Edward H. Shortliffe,
May 1982

card 1 of 1

Heuristic Programming Project
Report No. HPP 82-37

May 1982

COMPUTER-BASED CLINICAL DECISION AIDS: SOME PRACTICAL CONSIDERATIONS

Edward H. Shortliffe, MD, PhD

Division of General Internal Medicine

Department of Medicine

Stanford University School of Medicine

Stanford, California 94305

To be presented at

AMIA Congress 82

Hyatt on Union Square

San Francisco, California

2-5 May 1982

*Dr. Shortliffe is recipient of Research Career Development Award LM00048 from the National Library of Medicine. Additional support provided by the National Library of Medicine under program project grant LM-03395 and by the Biotechnology Resources Program of the NIH under grant RR-07785.

ABSTRACT

Medical decision making research has tended to emphasize the generation of optimal decisions, an issue which is central to the development of clinically useful consultation programs. This paper stresses the need to consider other theoretical and practical issues that are pertinent if consultation systems are to be accepted by physicians. Since adequate decision making performance remains an essential component of acceptable systems, the paper suggests criteria for selecting clinical problems that may be amenable to short-term implementation using state-of-the-art techniques.

Introduction

At the beginning of a third decade of research into the development of computer-based diagnostic aids, it is appropriate for medical computer scientists to assess the strides that have been taken, the barriers that remain, and the optimal strategies for furthering the field in the years ahead. One purpose of this meeting is to take a thoughtful look at medical decision making research and to identify potential solutions to the theoretical and logistical problems that continue to abound [1],[2]. Professor Szolovits has surveyed some of the relevant past work in his introductory tutorial, and papers describing state-of-the-art research in the field have also been contributed to the program. A special workshop following the main program will allow us to take a closer look at theoretical issues that face us, and to seek new strategies for melding techniques into unified theories that will heighten the performance of decision making programs in the future.

Because these scientific issues have been addressed by others, I have chosen to concentrate here on some of the practical matters that must be considered in designing for the introduction of acceptable and useful consultation systems. In particular, my goal is to help identify the characteristics of performance systems that can be built now using state-of-the-art computational techniques.

Attitudes of Physicians Towards Computers

It has been claimed that physicians are inherently reluctant to use computers in their practice [3],[6], and some observers have even suggested that the psychological barriers are insurmountable. Recent advances in computing, in particular the advent of personal computers and the increasing use of machines for office management functions, may account for the observation that physician attitudes seem to be changing remarkably quickly. Our group recently studied the opinions of practicing physicians and medical school faculty towards the development and implementation of consultation systems [7]. We found that a significant segment of the medical community

believes that assistance from computer-based consultation systems will ultimately benefit medical practice. We also studied the physicians' demands regarding desirable features for such systems if they are to be useful and clinically accepted. The resulting design considerations highlight performance capabilities that are a challenge to medical computer scientists. Consider, for example, the six design features that physicians rated most important in future consultation systems:

1) They should be able to explain their diagnostic and treatment decisions to physician users;

2) They should be portable and flexible so that the MD can access them at any time and place;

3) They should display an understanding of their own medical knowledge;

4) They should improve the cost efficiency of tests and therapies;

5) They should automatically learn new information when interacting with medical experts; and

6) They should display common sense.

No current consultation system meets all these criteria, but the list does help identify both the barriers to successful implementations and the criteria for assessing new systems that may be introduced. The first, third, fifth, and sixth of these criteria are central issues being addressed by researchers in the field of artificial intelligence (AI) and thereby help emphasize the importance of AI as an ingredient in the development of clinically acceptable decision aids. But neither AI nor other methodologies have led to polished consultation tools in widespread use, and it is appropriate to ask whether any decision making tool of significant clinical utility can be anticipated in the near future.

Steps in Demonstrating the Effectiveness of a Consultation System

Diagnostic programs developed over the years have tended to be assessed on the basis of their decision making accuracy -- the issue that is usually central to the system's design and to the motivation of the system's developers. Other aspects warranting formal evaluation are often overlooked. Yet there are several additional components to the evaluation process when it is performed optimally. In order to control for confounding variables, we have suggested [4] that system evaluations should be undertaken in a series of steps as follows:

1) Demonstrate a need for the system. Are there data indicating that physicians need help with the task for which the consultation system is designed to assist?

2) Demonstrate that the system performs at the level of an expert. Can it be formally shown that the system reaches the same decisions as

experts who are presented with the same clinical decision tasks? If there are frequent disagreements, can it be shown that the system is "correct" at least as often as the experts are? Note that the determination of "correctness" thereby requires some gold standard against which the performance of both experts and the consultation system can be measured.

3) Demonstrate the system's useability. Can physicians easily learn to handle the mechanics of interacting with the consultation system? Is the response time adequate? Is the system's performance sufficiently transparent so that the clinician can obtain the information he or she needs in an efficient and straightforward manner?

4) Demonstrate acceptance of the system by physicians. Can it be shown that clinicians offered the decision tool will in fact return to use it, even when access to it is entirely optional.

5) Demonstrate an impact on the management of patients. If physicians use the system, can it be shown that they follow the advice that it offers? If not, has it changed their behavior in some other way?

6) Demonstrate an impact on the well-being of patients. If physicians are following the recommendations of the consultation system, can it be shown that patients are benefiting from its use? Are there objective measurements of patient-care quality that can be assessed before and after the decision aid has been introduced?

7) Demonstrate cost-effectiveness of the tool. If all the other validation criteria have been satisfied, can it be shown that there is a version of the consultation system that is cost-effective when both costs and benefits are assessed using some generally accepted criterion?

These seven steps for demonstrating the effectiveness of a medical consultation system are idealized and difficult to traverse. I know of no medical decision making system that has rigorously been shown to meet formal validation criteria at all seven stages of development. In fact, most systems have been assessed only at stage 2, and remarkably few have met even the criterion of "need" specified in step 1.

Some observers of the field may argue that the theoretical issues in the development of high performance consultation systems are still so great that it is folly to focus attention on stages 3 through 7 at this time. Yet I believe that many significant theoretical barriers to the successful implementation of consultation systems do not arise at stage 2 and will not be met until the subsequent stages are encountered. In order to learn more about the performance issues that must be addressed to develop programs that are useable and acceptable to physicians, it is appropriate to ask whether the current state-of-the-art will permit some researchers to experiment with implementation issues beyond stage 2 in the near future.

Characteristics of an Optimal Application Domain

Attitude surveys such as the one mentioned above [7] help delineate some of the issues that must be addressed by system builders if clinically acceptable decision tools are to be developed. However, since most of these issues are best studied and assessed at the later stages of system implementation, scientists who wish to address them in their current (i.e., early 1980's) research must select an appropriate clinical problem area. The following criteria for that selection seem to be particularly pertinent:

1) As indicated above, there must be a demonstrated need for help in the domain. A program that deals with an "interesting" problem, but one with which physicians already do rather well, will tend to generate little interest.

2) Equally as important, there must be a recognized need for help by the physicians themselves. Data showing poor performance by the overall population of physicians will not necessarily convince individual practitioners that they are among those needing help. Demand will come only from perceived need by the intended users.

3) The domain should ideally provide a core of formalized and readily available knowledge. We have learned that knowledge base development can be an arduous and time-consuming aspect of consultation system research. Theoretical issues regarding knowledge completeness, consistency, and acquisition must inevitably be faced when a complex system is built for a domain in which expert knowledge is poorly formalized.

4) The domain must provide a straightforward mechanism for introducing a computer-based tool into the daily routine of the physicians who use it. This point has several corollaries. First, use of the computer should ideally replace a task that is already being performed; this helps guarantee that the system will require a minimal additional time commitment. Second, the mechanical interface must be rapid, congenial, and easy for the user to learn. And third, the decision tool's design must demonstrate a respect for the physician's hectic schedule.

5) The program should maintain the physician's role as ultimate decision maker (e.g., by giving explanations for recommendations and allowing the user to override any advice that is offered).

6) The system developers must be able to identify highly motivated collaborators from the domain of expertise.

7) The problem area should allow the initial prototype system to avoid major theoretical barriers (e.g., the domain should not require solutions to problems such as the development of approaches to the management of inexact inference, generalized methods for the management of temporal reasoning, encoding of strategic knowledge for domain-specific problem solving, or generation of highly customized explanations that demonstrate "first principle" understanding of the clinical area).

Conclusions

Several criteria have been suggested for selecting a clinical domain if an investigator wishes to study acceptability issues for computer-based consultation systems. The criteria are idealized, and no clinical problem area may meet them perfectly. However, our own group has attempted to select its current area of research [5] largely on the basis of considerations such as those I have listed. The criteria are less pertinent for research in computer-based clinical decision making if the work is addressing theoretical issues in knowledge representation and control, or if direct physician interaction with the computer is not contemplated.

References

1. Croft, D.J. Is computerized diagnosis possible? Comput. Biomed. Res. 5:351-367 (1972).

2. Friedman, R.B. and Gustafson, D.H. Computers in clinical medicine: a critical review. Comput. Biomed. Res. 8:199-204 (1977).

3. Melhorn, J.M., Warren, K.L., and Clark, G.M. Current attitudes of medical personnel towards computers. Comput. Biomed. Res. 12:327 (1979).

4. Shortliffe, E.H. and Davis, R. Some considerations for the implementation of knowledge-based expert systems. SIGART Newsletter, 55:9-12 (1975).

5. Shortliffe, E.H., Scott, A.C., Bischoff, M.B., et al. ONCOCIN: An expert system for oncology protocol management. Proceedings of the 7th International Joint Conference on Artificial Intelligence, pp. 876-881, Vancouver, British Columbia, August 1981.

6. Startsman, T.S. and Robinson, R.E. The attitudes of medical and paramedical personnel towards computers. Comput. Biomed. Res. 5:218-227 (1972).

7. Teach, R.L. and Shortliffe, E.H. An analysis of physician attitudes regarding computer-based clinical consultation systems. Comput. Biomed. Res. 14:542-558 (1981).