AIDS^2: A Decision-Support Tool for Decreasing Physicians' Uncertainty Regarding Patient Eligibility for HIV Treatment Protocols

Lucila Ohno-Machado,[a] Eduardo Parra,[b] Suzanne B. Henry,[c]
Samson W. Tu,[a] and Mark A. Musen[a]
[a]Section on Medical Informatics, Stanford University School of Medicine
[b]Department of Computer Science, Stanford University, Stanford, California
[c]School of Nursing, University of California, San Francisco, California

We have developed a decision-support tool, the AIDS Intervention Decision-Support System (AIDS^2), to assist in the task of matching patients to therapy-related research protocols. The purposes of AIDS^2 are to determine the initial eligibility status of HIV-infected patients for therapy-related research protocols, and to suggest additional data-gathering activities that will decrease uncertainty related to the eligibility status. AIDS^2 operates in either a patient-driven or protocol-driven mode. We represent the system knowledge in three combined levels: a classification level, where deterministic knowledge is represented; a belief-network level, where probabilistic knowledge is represented; and a control level, where knowledge about the system's operation is stored. To determine whether the design specifications were met, we presented a series of 10 clinical cases based on actual patients to the system. AIDS^2 provided meaningful advice in all cases.

INTRODUCTION

As the AIDS epidemic enters its second decade, advances in treatment are occurring rapidly, resulting in numerous therapy-related research protocols for which patients who have HIV infection may be eligible. Matching the HIV-infected patient with the appropriate protocol and ensuring adequate numbers of subjects for each study are essential tasks in the testing of new therapeutic regimens. The eligibility-screening process can be inefficient and extremely time consuming. Radcliffe and associates estimated that they screened and evaluated 10 patients for every one subject whom they enrolled in a research protocol [1]. We have developed a decision-support tool, the AIDS Intervention Decision-Support System (AIDS^2), to assist in the task of matching patients to therapy-related research protocols. The purposes of AIDS^2 are to determine the initial eligibility status of the HIV-infected patient for a protocol, and to suggest additional data-gathering activities that will decrease uncertainty related to the eligibility status.

Background

Resource allocation under a limited budget is an important concern for the success of any health-care procedure; therapy-related research protocols are no exception. The protocols usually present intricate inclusion and exclusion criteria by which patients are judged "eligible" or "not eligible." These judgments reflect data gathered in the process of history taking, physical and special examinations, and the interpretation of test results. By making the eligibility process more reliable and cost-effective, a computer-based system can indirectly enhance patient care, since resources can be saved and then used more efficiently for patients who in fact are the most likely to benefit from them. Although many computer-based approaches to data analyses of the outcome of clinical protocols exist [2, 3], only a few systems directly provide decision support for protocol-eligibility screening [1, 4]. Radcliffe and associates described the implementation and evaluation of computer-assisted telephone-based eligibility screening for research protocols [1]. Although no increase in the number of subjects enrolled in protocols was reported, the system decreased the number of patients requiring clinical evaluation to establish protocol-eligibility status. Tu and associates studied the problem of automatic eligibility determination for HIV clinical trials [4], and showed that a computer-based system can assist researchers in performing this task. These studies support the need for and feasibility of computer-based approaches to the task of screening for protocol eligibility.

Although Tu and associates discussed the use of Bayesian belief networks for representing uncertainty related to missing values in the medical record [4], they implemented their system based on a qualitative heuristic approach. In contrast, the work presented in this paper systematically uses Bayesian belief networks to manage uncertainty, and represents domain knowledge in a *classification hierarchy*. Belief networks are directed acyclic graphs in which nodes represent random variables, and the connections between nodes represent probabilistic dependencies [5].

Design Considerations

After performing a careful analysis of the task of protocol eligibility screening, we generated design goals related to the user-interface and decision-support capabilities of AIDS^2. To support the different needs of the primary-care provider who is screening an individual patient and of the principal investigator or research nurse who is attempting to achieve full enrollment for a protocol, the system must operate in two modes. In the patient-driven mode, the system must determine whether a particular patient is

0195-4210/92/$5.00

eligible, is not eligible, or is potentially eligible for a protocol. This procedure is currently done every time a new patient comes to the clinic, and every time an old patient demonstrates a significant change in his clinical condition. The system must also select the best way to pursue further investigation for this patient, by suggesting adequate tests or examinations. In the protocol-driven mode, the system must query a database of patient descriptions and rank-order the patients based on their probability of eligibility for a particular protocol. This procedure is currently done every time a new protocol is introduced in the clinic. This approach allows focused data-gathering activities for the patients most likely to be eligible for a protocol. Suggestions for specific data acquisition for each patient must then proceed from the patient-driven mode. The system must also be able to deal with missing values. Additional design goals for the user interface included minimal data entry from the keyboard and extensive use of graphical displays, to facilitate physicians' acceptance of the system.

SYSTEM DESCRIPTION

In the following sections, we provide a detailed description of the function, knowledge representation, and inference mechanisms of AIDS^2.

System Function

We determine patient-eligibility status by examining the inclusion and exclusion criteria for enrollment in a specific therapy-related research protocol. If the status of a patient relative to a criterion can be determined with the available information, then we directly determine the criterion-related eligibility by performing a query to a database of patient information. If we do not have the specific information for a particular criterion, then we can only indirectly estimate the eligibility by obtaining information from the database that influences the belief that the patient satisfies the criterion-related eligibility. For example, if one of the protocol's inclusion criterion is "ALT < 6 times the normal level," and we have no current entry for the patient's alanine aminotransferase value, the system will estimate the value using available information on related data, such as current values of aspartate aminotransferase, or any other indications of liver damage. Using either the direct or indirect method of eligibility determination, we obtain an initial patient eligibility estimate for a specific protocol. The indirect method uses a Bayesian belief network to represent the dependencies among related patient states and observations.

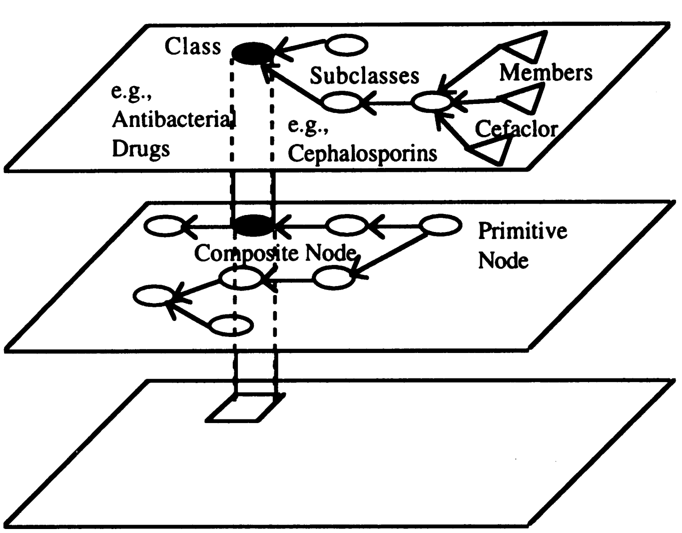
We use one of two user-interface modes, as described previously. In the protocol-driven mode, we obtain a ranking of patients according to their eligibility status for a given protocol. If we want suggestions for data-acquisition tasks to decrease uncertainty related to the eligibility status for a particular patient, however, we proceed in the patient-driven mode. This mode guides the user in selecting the pieces of information (generally laboratory-test results) that are most relevant in determining the eligibility status of a given patient for the selected protocol. We divide the inclusion and exclusion criteria into three *categories* ("history," "examination," and "tests") based on importance in determining eligibility status for the protocol, on risk to the patient, and on cost, which includes the cost of the procedure and the cost of the clinician's time. We divide the test categories into three subcategories based on the cost of the tests, as estimated by Detmer and associates [6]. We use the categories and subcategories to reason about the order in which missing data related to the protocol criteria should be pursued, and as a key to group all output information.

We envision AIDS^2 as part of an embedded computational environment that actively interacts with other sources of information, such as the laboratory database and the electronic medical record. We implemented an initial version of the system on the Macintosh™, using Ergo™ and Hypercard™. We have now fully implemented AIDS^2, except for the direct link between the system and a database for information retrieval. We are currently simulating the queries to the database, by loading all relevant information for a protocol directly into Hypercard™.

Knowledge Representation

There is an extensive literature related to the use of multiple levels to represent medical knowledge for solving a determined problem [7]. This approach derives from the fact that a detailed representation of all concepts and relations that are potentially involved in solving a specific problem may be computationally intractable or unnecessary. For example, it would be difficult to obtain probability distributions for every test conditioned on every disease to obtain a model that can help in determining patients' eligibility. We can, however, estimate the values of certain tests given that the patient has a certain type of disease (e.g., aminotransferases, given that the patient has active viral hepatitis). The degree of detail and the inference mechanisms applied at different levels vary according to the researcher's goal and available resources.

To represent the knowledge contained in protocols, we used three different levels, as shown in Figure 1. The belief-network level models uncertainty related to missing values. The medical concepts are modeled in the classification level. The control level stores information regarding the connections between the other two levels, predefined information-retrieval priorities, and certain protocol-specific information (which nodes represent criteria, what the cut-off values for criteria are, and so on).



LEVEL	WHAT IS REPRESENTED
Classification Level	Medical Concepts
Belief-Network Level	Uncertainty about Criteria Eligibility
Control Level	Procedural Knowledge

Figure 1 Knowledge representation. Three different levels are used to represent the knowledge contained in the protocols: classification level, belief-network level, and control level.

Modeling of Uncertainty

The need to represent the uncertainty involved in estimating criterion-specific eligibility status, the complex relations among the variables of interest, and the need to provide a model that could be reused in different protocols led us to represent some data relationships in a Bayesian belief network. Every node in the belief network either represents a clinical parameter (e.g., "Hemoglobin") that is used in a criterion (e.g., Hemoglobin > 11"), which we call *criterion node*, or represents a clinical data element that influences other nodes (e.g., "Anemia"). These influences are expressed as probabilistic dependencies (e.g., probability that hemoglobin > 11 given that the patient has clinical signs of anemia). State values of nodes in the network can be either unknown or set to a particular predefined state. If the state of a node is known, the probability of that state is 1.0, whereas the probabilities of all other states are 0.0. If the node value is unknown, the network is used to compute a posterior distribution conditioned on all nodes that have a relationship with the node being observed. AIDS^2 determines the patient eligibility status based on each criterion by examining the value of the corresponding belief-network criterion node. The patient eligibility for each criterion is represented by a probability. This probability is computed as the addition of the probabilities of all mutually exclusive states of the criterion node that are considered to be eligible. For example, if we are looking for "Hemoglobin > 11," we add the probabilities for all intervals where the hemoglobin value is over 11 for that patient (there are often more than two intervals, because the network was designed to be valid across different protocols, and each of the protocols may define a different cut-off value for its inclusion and exclusion criteria).

Modeling of Medical Concepts

The classification level stores a set of hierarchical models to represent generic medical data and abstractions (diseases and classes of diseases, drugs and classes of drugs, and so on). The purpose of these *classification hierarchies* is to simplify the definition of the belief network. A belief-network node can represent either a single element (*primitive node*, e.g., "penicillin"), or a group of elements (*composite node*, e.g., "antibacterial drugs"). For example, if a protocol criterion requires that the patient does not take any antibacterial drugs, the composite node "antibacterial drugs" in the belief network must be checked. The latter is linked to a *class* at the classification level. Each class is composed of *subclasses* or *members* (the class "antibacterial," for example, contains the subclass "cephalosporin," which contains the member "cefaclor"). In this case, we need only to define a relation between the composite node at the belief-network level and the class node at the classification level. We do not have to include different sets of arcs (dependencies) in the belief network for every member already represented in the classification hierarchy. The classification level enables the use of the simplified belief network. No belief-network modification is needed when a new member of the class is added, or when an existing class is extended at the classification level. We need to modify only the information at the classification level. This classification level can be reused in other models.

Control of Inference

The control level contains all annotations to the belief network and classification levels, which are necessary to control the inference done by the system. It includes information such as which nodes in the belief network are primitive and which are composite (and therefore are represented as classes in the classification level), which states of a belief-network criterion node represent eligibility in the specific protocol, and what information is needed to guide the user. The system computes an initial eligibility status for all the patients being considered for the particular protocol as follows:

```
for all criterion C in protocol P, do
  if there is database information for criterion C,
    then set the appropriate state for criterion C in the corresponding belief-network criterion node
  else
    acquire all possible information of belief-network nodes related to the criteria node
    propagate probabilities and obtain an estimate for the criterion state distribution
```

For each criterion node, we obtain the database information by executing the appropriate query stored in the control level. If there is sufficient information to determine the state of the belief-network node, we set the criterion state correspondingly. When the database does not have direct information to determine the patient eligibility based on the target criterion, we need to access the information of related belief-network nodes, to obtain a probabilistic state distribution for the criterion node of interest. Once all related node information has been recorded, we use the belief-network probability-propagation inference to estimate the criterion state distribution.

Any node in the belief network can be composite (i.e., can be linked to a class at the classification level). When the control mechanism encounters a composite node, the system performs a series of database queries by navigating through the classification hierarchy until sufficient data are retrieved to set a state in the belief-network node. Once all nodes that represent criteria in the belief network have an associated state or a distribution over the states, we can determine the patient eligibility. We do so by examining all the states of the belief-network criterion node that are consistent with patient eligibility, by adding the probabilities of all mutually exclusive states of the criterion node that are considered to be eligible. The user is presented with both a graphical and a textual representation of the eligibility status of the patients, ordered by the three categories ("history," "examination," and "tests"). If the user wants to acquire further information to determine the patient eligibility status precisely, the system separates known from estimated eligibility status. A known eligibility status for a criterion is either "not eligible" or "eligible" (eligibility = 0.0 or eligibility = 1.0, respectively). Unknown values will lie between these limits. The user can assert new values for a criterion at any time (e.g., "ALT = 360"). The system suggests information acquisition based on the three categories, but the user is allowed to enter data in any order. Once the user asserts a new state for a criterion, the system updates all information and presents the changes in eligibility status for the patient. At any point, the user can retract a previously asserted value for a criterion, or can restart with the initial eligibility status. The system also presents a summary report to the user. If a patient is not eligible — that is, if there are one or more criteria that are not satisfied by the patient for the specific protocol — then the system reports the eligibility status to the user, highlighting the criteria that were not satisfied. If the patient is eligible — that is, if all criteria are met by the patient for the protocol — then the system presents a message to inform the user that all criteria have been satisfied by the patient. If further data are needed to determine accurately the eligibility status of a patient, the system presents a list of data to be collected, ordered by the defined three categories.

When the system is in the protocol-driven mode, the user can request a listing of patients ordered by potential eligibility for the particular protocol being considered. To rank the patients, we use an unique measure of eligibility for each of the three categories, rather than eligibility status for every criterion in the protocol. We obtain this measure by multiplying the probabilities that a patient is eligible for each of the criteria in that category. Each category has an associated utility value. Using the categories as keys, the system sorts all patients currently being considered. The primary key is the history category (highest utility value); the secondary key is the examination category; and the last sorting key is the test category.

DISCUSSION

During our selection process, we gained insight about the variation in language precision among therapy-related research protocols and among criteria within a protocol. Certain criteria were specified precisely (e.g., "creatinine <= 2 mg/dl"), whereas others were vague (e.g., "therapy with drugs with potential to cause neuropathy"). This problem in protocol specification has been observed previously in the domain of oncology clinical trials [8]. Other researchers dealing with protocol ambiguity have suggested a combined qualitative and quantitative approach. In our own work, we implemented a combined approach by modeling knowledge representation in three levels.

We needed a classification level that could manage the large number of diseases and drugs. Without this level, the belief network would be unnecessarily complex, and probability assessment would be impossible. Using this representation, we were able to deal with abstractions efficiently (e.g., "use of antihypertensive drugs"), without having to include every member or subclass as a belief-network node (e.g., "use of captopril"). The choice of the belief network as a representation for the pathophysiological dependencies was based on the notion that the system should provide the ability to score the eligibility of a given patient according to prespecified groups of criteria, even if complete data were not available. The formal probabilistic foundations of Bayesian networks and the explicit causal modeling make belief networks suitable for this type of task [9]. Since the classification was independent of the protocol, only small portions of the belief network and the control level have to be modified for the system to accommodate large numbers of protocols. The control level was necessary to establish the links between the classification and the belief-network level, and to determine the information that would have the highest overall value. The function that calculates the overall value of the information was chosen arbitrarily, but we believe that it represents a clinically relevant approximation. We had difficulty assessing the probabilities of the belief network, even though the published medical literature related to AIDS and HIV-infection is extensive. Several conditional probabilities are not explicitly available, and we estimated them in a superficial way, based on clinical experience. Our aim, however, is to generalize the design of the belief network and the classification levels so that, each time a new protocol is entered into the system, fewer probabilities require assessment and data entry at all levels is facilitated.

Wyatt used Donabedian's framework to categorize the attributes of a medical decision aid that can be measured as structure, function, and impact [10]. We undertook a Type I laboratory study, focused on the system's performance for evaluation for AIDS^2. To determine whether the design specifications were met, we presented a series of 10 clinical cases based on actual patients to the system. The system demonstrated full functionality related to the design goals in all cases. Future evaluation plans include a Type II laboratory study in which primary-care providers and research nurse coordinators screen patients for protocol eligibility, with and without the system, in a laboratory setting. Variables of interest will include number of patients identified as eligible for the protocol, time taken to decrease the uncertainty related to patient eligibility status, and user satisfaction with the interface and with the advice given by the system.

AIDS^2 demonstrates that a combined approach of qualitative and quantitative information, in the presence of missing values, can help physicians to determine patient eligibility for a treatment protocol. This approach makes certain levels of the system, such as the classification level and the belief-network level, reusable in different protocols, and permits more efficient resource allocation in terms of the designer's and the health-care worker's time.

Acknowledgments

This work was supported by computing resources provided by the Stanford University CAMIS project, which is funded under grant LM 05305 from the National Library of Medicine (NLM) of the National Institutes of Health. Dr. Ohno-Machado was supported by a grant from CNPq, Brazilian Ministry of Education, and in part by grant HS06330 from the Agency for Health Care and Policy and Research.

Dr. Henry was supported by NLM training grant LM 07033. The authors thank Michael Walker and Lyn Dupré for their assistance.

Reference

[1] Radcliffe, J.M., Latham, G.S., Sunderland, T., and Lawlor, B.A. Computer-assisted telephone screening: A new system for patient evaluation and recruitment. In Miller, R.A. (ed.), *Proceedings of the Fourteenth SCAMC,* Los Alamitos, CA: IEEE Computer Society Press, 1990, pp.831–834.

[2] Shortliffe, E.H. Medical expert systems: Knowledge tools for physicians. *Western Journal of Medicine,* 145, 1986, 830–839.

[3] Elting, L.S., and Bodey, G.P. Evaluation of benefits derived from a computerized data management system for clinical trials data. In Clayton, P.D. (ed.), *Proceedings of the Fifteenth SCAMC,* New York: McGraw-Hill, 1991, pp.48–52.

[4] Tu, S.W., Kemper, C.A., Lane, N.M., Carlson, R.W., and Musen, M.A. A methodology for determining patients' eligibility for clinical trials *Methods of Information in Medicine,* 32, 1993, 317–25.

[5] Pearl, J. Fusion, propagation, and structure in belief networks. *Artificial Intelligence,* 29, 1986, 241–288.

[6] Detmer, W.M., McPhee, S.J., Nicoll, N., and Chou, T.M. *Pocket Guide to Diagnostic Tests,* Norwalk, CT: Appleton and Lange, 1992.

[7] Patil, R.S. Efficient structuring of composite causal hypothesis in medical diagnosis. In Stead, W.W. (ed.), *Proceedings of the Eleventh SCAMC,* New York, 1987, pp.23–29.

[8] Musen, M.A., Rohn, J.A., Fagan, L.M., and Shortliffe, E.H. Knowledge engineering for a clinical trial advice system: Uncovering errors in protocol specification. *Proceedings of the AAMSI Congress,* Anaheim, 1986, pp.24–27.

[9] Neapolitan, R.E. *Probabilistic Reasoning in Expert Systems: Theory and Algorithms.* New York: John Wiley and Sons, 1990.

[10] Wyatt, J., and Spiegelhalter, D. Evaluating medical expert systems: What to test, and how? *Medical Informatics* 15, 1990, 205–217.