AD_____

Award Number: DAMD17-00-1-0244

TITLE: Improving Breast Cancer Research Through Automated Matching of Patients to Clinical Trials

PRINCIPAL INVESTIGATOR: Lawrence O. Hall, Ph.D.
Dmitry B. Goldgof, Ph.D.
Jeffrey Krischer, Ph.D.

CONTRACTING ORGANIZATION: University of South Florida
Tampa, Florida 33620-2897

REPORT DATE: August 2004

TYPE OF REPORT: Final Addendum

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	August 2004	Final Addendum (3 Jul 03-2 Jul 04)

4. TITLE AND SUBTITLE
Improving Breast Cancer Research Through Automated Matching of Patients to Clinical Trials

5. FUNDING NUMBERS
DAMD17-00-1-0244

6. AUTHOR(S)
Lawrence O. Hall, Ph.D.
Dmitry B. Goldgof, Ph.D.
Jeffrey Krischer, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of South Florida
Tampa, Florida 33620-2897

E-Mail: hall@csee.usf.edu

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 Words)

An enhanced Web based prototype intelligent agent/expert system for matching breast cancer patients to clinical trials has been built. It allows for cost preferences to be entered. Therefore, the system user can choose to rule patients out of trials as quickly as possible without regard to the cost of tests necessary to do this. A user can choose to have questions appear so that the patient is ruled out of the trial with the minimal set of costs (tests) or can choose some combination of approaches. The system has been tested with 15 protocols and designed for maximal responsiveness and scalability as new protocols are added. The files of 178 former patients have been used to test the accuracy of the system. Additionally, the files of 213 current patients have been tested for eligibility. Patients for each of the protocols were correctly found eligible for one or more trials. We found 240 new matching clinical trials for the 213 current patients. A probabilistic prototype system has been developed to reorder questions based on the probability they will determine the patient is ineligible for trial and preliminary experiments have shown up to 13% less questions will be required on average. It can also indicate the probability of patients being eligible for protocols. We have also developed a prototype system to quickly add new clinical trials. This has been successfully used by novices to enter new trials.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Cancer control, clinical trials, expert system	46
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

Table of Contents

3 Introduction

Increasing the enrollment of patients in clinical trials is important to making progress towards finding more effective treatments for breast cancer. Accrual is complicated by a large number potential studies and the cost and complexity of determining whether a patient meets the necessary eligibility criteria. Under this proposal, we are developing a Web based expert system which can determine the patients eligibility for clinical trials. The expert system is designed to take into account the cost of tests which are required to meet inclusion criteria and acquire information in the most cost-effective way possible.

Additionally, it is important to be able to easily add and remove clinical trials to the system. Trials are continually becoming available, going on suspension or being closed to accrual. Towards this end, we have developed a companion Web based system that enables anyone to simply enter the information required to describe the eligibility/ineligibility criteria for a clinical trial. A newly entered trial/protocol can then be directly included in the Clinical trial assignment expert system with no expert intervention.

Finally, we have worked on methods of utilizing probabilities to order questions so that those most likely to rule patient out of a protocol are first. Recent testing has shown this is effective.

4 Body

In the extension year, fourth year, we have done the following. We have entered approximately 44 new patients into our system to check them for eligibility in available breast cancer protocols. In total, they were found eligible for 78 trials to which they did not get assigned. We have added several new protocols. We have revised and published a journal paper (in the Artificial Intelligence in Medicine journal) that shows significant potential for the increase in accruals to clinical trials using our system.

A new version of the software has been developed that keeps all of the eligibility criteria in memory after the user provides some information. This significantly speeds up the response of the system and makes it more usable clinically.

The system has been tested by both research nurses and physicians. One of the concerns they have expressed is that some questions get asked after they have given a previous answer which implies the answer for the latter question. It is possible to develop implication rules which take care of this problem, but that requires interaction with very busy experts in the medical field. Hence, a study been undertaken into how to learn these implications as patient data is entered.

Using Association Rules, we have been able to find all of the expert derived implications. Further, we have discovered some new ones in which when information about a particular test is available, this indicates that biopsy has been done, for example.

Also, a new method of minimizing the number of questions needed or the amount of data needed to determine if the patient is ineligible has been developed. A slight modification of the same approach, provides the user of the system with an indication of how likely it is that the patient will be eligible for any particular protocol they are exploring. This will enable them to focus on a particular protocol or protocols as information is entered, if they wish. The results of this work were codified in a conference paper in the IEEE Computer-based medical systems conference.

Table 1: Results of selecting clinical trials for the 187 past patients and 169 current patients. We give the number of trial participants, selected by both the system and Moffitt clinicians, and the number of the other eligible patients, identified by the system.

(a) Results for the 187 past patients.

Clinical Trial	Participants	Other Eligible
10822	10	5
10840	0	19
11072	48	26
11378	4	19
11992	5	6
12100	8	20
12101	20	30

(b) Results for the 169 current patients.

Clinical Trial	Participants	Other Eligible
11132	4	1
11931	2	26
11971	4	0
12100	0	5
12101	11	52
12385	0	19
12601	0	1
12643	16	36
12757	1	3
12775	23	17

For completeness, we repeat what was done in the third year below. In the third-year, we have refined the original prototype to produce version 1.4. We have tested it with data from 187 retrospective patients and 169 more recent patients including some who are currently undergoing treatment. We have extensively tested its ability to order questions associated with tests to save dollar costs on over 300 patients. Table 1 summarizes our matching results on the past and current patients. Patients are only evaluated for trials that are currently enrolling patients. The trial status can change when a trial is put on suspension, closed, brought off suspension, or initiated. It can be seen that the system finds all matches that correspond to trials in which patients have been enrolled. For the current 169 patients for which extensive tests have been done, we found 160 new matches to protocols! This is quite promising for increasing accrual.

The cost savings are shown in Table 2. We show the mean test costs with and without the ordering heuristics. Six clinical trials have incurred selection costs; the heuristics have reduced

Table 2: Cost savings by test reordering.

(a) Results for the 187 past patients.

Clinical Trial	Mean Cost	
	W/O Test Reordering	With Test Reordering
10822	$70	$11
10840	$0	$0
11072	$209	$60
11378	$35	$19
11992	$0	$0
12100	$0	$0
12101	$0	$0

(b) Results for the 169 current patients.

Clinical Trial	Mean Cost	
	W/O Test Reordering	With Test Reordering
11132	$0	$0
11931	$0	$0
11971	$192	$192
12100	$0	$0
12101	$0	$0
12385	$0	$0
12601	$36	$3
12643	$0	$0
12757	$107	$107
12775	$0	$0

the costs for four of these trials, and have not affected the costs for the other two trials.

There are now 15 protocols available in the system. At the present time, all breast cancer protocols at the Moffitt Cancer Center which are accruing at least two patients a month are available through our system. Our automated clinical trial updating system continues to allow us to easily add trials to the system [1, 2].

We have created a question ordering system that uses a crude probabilistic heuristic. As patients are tested against the system over time, we can keep a record of how many times each question causes a patient to be classified as ineligible for a protocol. These results will take the form of x out of y times that a question was asked it directly caused a patient to be determined ineligible for protocol z. The value $(\frac{x}{y})_{qp}$ can be treated as the probability that question q will cause a patient to be declared ineligible for protocol p. This value will be reasonably reliable after the question has been asked more than 30 times. At that point, it can be used to reorder questions. The question with the highest probability of making a patient ineligible for a trial can be displayed first. By doing this patients will be quickly determined ineligible with a minimum

number of questions.

Preliminary experiments have been done which indicate this approach does in fact reduce the number of questions necessary to determine eligibility. Results are shown in Table 3.

We selected 90 patients at random from our list of patients and used their data in experiments. A ten-fold cross validation was carried out, so that the system was trained on 81 patients and the remaining 9 patients were tested using the system. The test was done on six protocols for which 90 patients had been tested. As can be seen in Table 3, the probabilistic system allows approximately 13% less questions to be answered to determine eligibility, on average.

Table 3: Probabilistic question ordering vs. analytic question ordering.

Ten-fold cross validation				
	Average number of questions			
Protocol	Probabilistic System	Analytical System	Difference	Difference %
11931	15.35	18.90	3.55	18.78
12100	13.85	13.95	0.10	0.72
12101	21.65	24.75	3.10	12.53
12521	14.75	19.05	4.30	22.57
12601	13.90	15.70	1.80	11.46
12777	14.40	16.10	1.70	10.56
Average	15.65	18.08	2.43	13.42

Key Research Accomplishments:

- We have enhanced our prototype system to very stable version 1.4. We have corrected cost functionality (mostly by getting the costs of tests correct and determining all tests that are done in the routine care) and tested this successfully.
- Utilizing retrospective patient data and current patient data, it has been found that patients are eligible for multiple protocols/trials. Further, with current patient data we find patients eligible for trials and not put on any trial.
- Extensive testing of cost functionality has been done. We have determined that in many cases there is no possibility of saving costs. However, when it is possible the cost mechanism recommends questions in order that will always allow eligibility to be determined in the minimal cost fashion.
- We have developed a method of determining probabilities that questions will show patients are ineligible for trials. The probabilities are determined empirically while the system is in use. We showed that the use of these probabilities to order questions on a query page will result in the need to answer less questions at all times.
- We have applied data mining, via association rules, to 100 cases that have been put through the system. This has enabled us to find what we call fact implications rules. For example,

- if a question on a biopsy is answered as yes, it is clear that some surgery has been done and it is not necessary to ask questions about whether the patient has ever had any surgery. The ability to recover such rules will streamline the system from a usability point of view and allow to improve with time without requiring interviews with physicians and nurses.

Reportable Outcomes: We have had a paper [3], which is attached, published the 2003 IEEE International Conference on Systems, Man, and Cybernetics. We have a paper about to be published in the Artificial Intelligence in Medicine journal [4]. It is attached. A paper has been just recently published at the IEEE Computerized Medical Based Systems conference. It is attached.

We have submitted a revised proposal to the National Institutes of Health to take this system into clinical operation at the Moffit Cancer Center. The previous proposal was well-received, but there were concerns about the level of cooperation with the Cancer Center.

A complete bibliography of papers from this grant is:

- Savvas Nikiforou, Eugene Fink, Lawrence O. Hall, Dmitry B. Goldgof, and Jeffry P. Krischer, Knowledge Acquisition for Clinical-Trial Selection, IEEE International Conference on Systems, Man and Cybernetics, October 2002.
- Princeton K. Kokku, Lawrence O. Hall, Dmitry B. Goldgof, Eugene Fink, and Jeffry P. Krischer, A Cost-effective Agent for Clinical Trial Assignment, IEEE International Conference on Systems, Man and Cybernetics, October 2002.
- E. Fink, L. O. Hall, D. B. Goldgof, B. Goswami, M. Boonstra, J. P. Krischer, Experiments on the Automated Selection of Patients for Clinical Trials, IEEE International Conf. on Systems, Man and Cybernetics, pp. 4541-4545, Oct. 2003.
- E. Fink, P.K. Kokku, S. Nikiforou, L.O. Hall, D.B. Goldgof, J.P. Krischer, Selection of Patients for Clinical Trials: An Interactive Web-Based System, Artificial Intelligence in Medicine, To Appear 2004.
- Bhavesh D. Goswami and Lawrence O. Hall and Dmitry B. Goldgof and Eugene Fink and Jeffrey P. Krischer, Using Probabilistic Methods to Optimize Data Entry in Accrual of Patients to Clinical Trials, The 17th IEEE Symposium on Computer-Based Medical Systems, 2004.

A web prototype of the clinical trial assignment system is available at
http://morden.csee.usf.edu/moffit with password available from the principal investigator.

M.S. Theses:
Bhavesh Goswami, Computer Science, May 2004.
Princeton Kokku, Computer Science, August 2003.
Savvas Nikiforou, Computer Science, May 2002.

5 Conclusions

We have developed a scalable prototype which currently can determine eligibility for sixteen breast cancer clinical trials. The system has been tested using retrospective data from 201

patients who are assigned to some clinical trial and more recently active patients numbering 213. The system correctly finds cases in which a patient is eligible for multiple clinical trials. It has found 240 matching trials for the 213 current patients. This indicates that it is quite likely they use of the system will significantly increase accruals to trials.

The system is able to utilize monetary cost in requesting tests to rule in/rule out a patient from the set of available clinical trials. The default ordering of questions allows the system user to rapidly determine the eligibility or ineligibility of a patient for any subset of the available clinical trials entered into the system. We have been able to show a good average cost saving by using the cost feature to order questions. Of course, there is no guarantee that a clinician would order tests as suggested by the question ordering of our system. However, the potential for cost savings is significant.

The system is Web based and password protected. It provides rapid response when a person enters answers to one or more questions on a page of system selected questions. It can be used from any computer on the World Wide Web. Hence, community physicians will be able to determine the potential eligibility (they may not wish to run all tests) of the patient for clinical trials at cancer centers in their region.

A prototype to enable physicians, nurses or technicians to enter new protocols has been completed. The system is now in use. It reduces the time required to add a new trial or protocol to approximately 1 hour. It enables non-computer scientists to add trial/protocols to the system. This knowledge acquisition tool has been designed to minimize/eliminate the cases where similar questions acquiring essentially the same information would have to be asked. This feature has the potential to cause slight changes to the wording of inclusion/exclusion criteria. We believe that this change is minor and will have no effect on IRB approval.

Last year, we intended to evaluate whether IRB approval would be affected. However, institutional issues prevented this. However, this year we will have new protocols entered using existing questions and plan to go back to the IRB board to discuss any changes in criteria wording to fit existing questions within the system. An example would be a protocol in which there are two questions which ask "is a test value is greater than some threshold" and then a separate question that asks if it is less than some threshold, versus a single question which asks if a test is in some range. We believe that such a change is trivial, but this must be addressed in practice and we will evaluate whether it causes review board decisions to potentially change.

We have utilized Bayes rule to provide a likelihood prediction for a patient being eligible for particular trials at any given step after enough patients have been put through the system for that trial. It appears quite useful. A version of it has been used to minimize the number of questions required to be answered before patient is ruled ineligible for a protocol. There was between a 13 and 22 percent reduction in the number of questions needed.

5.1 The future

The prototype system shows the potential for allowing community physicians, as well as cancer center physicians, to quickly and cost effectively determine for which clinical trials a patient may be eligible. It holds the promise of enabling greater patient accrual for trials by increasing the awareness of each trial for treating physicians throughout a region. In this future, we would like to change our IRB to allow evaluation of how many patients not eligible for clinical trials were

actually missed by clinical practitioners vs. excluded for a particular reason (e.g. it was clear they would not agree) or were offered a trial and declined to enter it.

References

[1] Savvas Nikiforou. Selection of clinical trials: Knowledge representation and acquisition. Master's thesis, University of South Florida, 2002.

[2] Savvas Nikiforou, Eugene Fink, Lawrence O. Hall, Dmitry B. Goldgof, and Jeffrey P. Krischer. Knowledge acquisition for clinical-trial selection. In *IEEE International Conference on Systems, Man, and Cybernetics*, pages 66–71, 2002.

[3] E. Fink, L.O. Hall, D.B. Goldgof, B.D. Goswami, M. Boonstra, and J.P. Krischer. Experiments on the automated selection of patients for clinical trials. In *IEEE International Conference on Systems, Man,& Cybernetics*, pages 4541–4545, 2003.

[4] E. Fink, P.K. Kokku, S. Nikiforou, L.O. Hall, D.B. Goldgof, and J.P. Krischer. Selection of patients for clinical trials: An interactive web-based system. *Artificial Intelligence in Medicine*, 2004. Available on ScienceDirect Articles in Press.

- List of people paid off the grant: Lawrence O. Hall
  Dmitry Goldgof
  Princeton Kokku
  Eugene Fink
  Prodip Hore
  Cindy Bethel
  Matt Boonstra
  Bhavesh Goswami
  Rochelle Harris
  Hong Zhang

6 Papers

Artificial Intelligence in Medicine (2004) xxx, xxx–xxx

ELSEVIER

Artificial Intelligence in Medicine

http://www.intl.elsevierhealth.com/journals/aiim

Selection of patients for clinical trials: an interactive web-based system

Eugene Fink*, Princeton K. Kokku, Savvas Nikiforou, Lawrence O. Hall, Dmitry B. Goldgof, Jeffrey P. Krischer

Computer Science and Engineering, University of South Florida, Tampa, FL 33620, USA

Received 6 February 2003; received in revised form 19 August 2003; accepted 17 January 2004

KEYWORDS
Clinical trials; Expert system; Cost reduction; Breast cancer

Summary The purpose of a clinical trial is to evaluate a new treatment procedure. When medical researchers conduct a trial, they recruit participants with appropriate health problems and medical histories. To select participants, they analyze medical records of the available patients, which has traditionally been a manual procedure.

We describe an expert system that helps to select patients for clinical trials. If the available data are insufficient for choosing patients, the system suggests additional medical tests and finds an ordering of the tests that reduces their total cost. Experiments show that the system can increase the number of selected patients. We also present an interface that enables a medical researcher to add clinical trials and selection criteria without the help of a programmer. The addition of a new trial takes 10–20 min, and novice users learn the functionality of the interface in about an hour.

1. Introduction

Cancer causes 550,000 deaths in the United States every year [1,2], and the treatment of cancer is an active research area. Medical researchers explore new treatment methods, such as drugs, surgery techniques, and radiation therapies. An experiment with a new treatment procedure is called a *clinical trial*. When researchers conduct a trial, they recruit patients with appropriate cancer types and medical histories. The selection of patients has traditionally been a manual procedure, and studies have shown that clinicians can miss up to 60% of the eligible patients [3–8].

If the available records do not provide enough data, clinicians perform medical tests as part of the selection process. The costs of most tests have declined over the last decade, but the number of tests has increased [9,10], which is partially due to inappropriate ordering of tests [11,12]. Clinicians can reduce the cost by first requiring inexpensive tests and then using their results to avoid some expensive tests; however, finding the right ordering may be a complex optimization problem.

The purpose of the described work is to automate the selection of patients for clinical trials and minimize the cost of related tests. We have developed an expert system that identifies appropriate trials for eligible cancer patients, designed a web-based interface that enables a clinician to enter new trials without the help of a programmer, and built a knowledge base for trials at the Moffitt Cancer Center, located at the University of South Florida.

We begin with a review of the previous work on medical expert systems (Section 2). We then explain the design of the developed system and present empirical confirmation of its effectiveness (Section

*Corresponding author.
E-mail addresses: e.fink@cs.cmu.edu (E. Fink), kokku@csee.usf.edu (P.K. Kokku), savvasn@ucy.ac.cy (S. Nikiforou), hall@csee.usf.edu (L.O. Hall), goldgof@csee.usf.edu (D.B. Goldgof), jpkrischer@moffitt.usf.edu (J.P. Krischer).

0933-3657/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.artmed.2004.01.017

3). We also describe the interface for adding new knowledge (Section 4). In conclusion, we point out some limitations of the developed system and compare it with other trial-selection systems (Section 5).

2. Previous work

Researchers have developed several expert systems that help to select clinical trials for cancer and AIDS patients. In particular, Musen et al. built a rule-based system, called EON, that matched AIDS patients to clinical trials [13]. Ohno-Machado et al. developed the AIDS^2 system, which also assigned AIDS patients to clinical trials [14]. They integrated logical rules with Bayesian networks, which helped to make decisions based on incomplete data and to quantify the decision certainty.

Bouaud et al. created a cancer expert system, called ONCODOC, that suggested alternative clinical trials for each patient and allowed a physician to choose among them [15,16]. Séroussi et al. used ONCODOC to select participants for clinical trials at two hospitals, which helped to increase the number of selected patients by a factor of 3 [17,18].

Hammond and Sergot created the OaSiS architecture [19], which had a graphical interface for entering patients' data and extending the knowledge base. Smith et al. built a qualitative system that assisted a clinician in selecting medical tests, interpreting their results, and reducing the number and cost of tests [9,20].

Theocharous developed a Bayesian system that selected clinical trials for cancer patients [21,22]. It learned conditional probabilities of medical-test outcomes and evaluated the probability of a patient's eligibility for each trial. On the negative side, the available medical records were often insufficient for learning accurate probabilities. Furthermore, when adding a new trial, the user had to change the structure of the underlying Bayesian network. To address these problems, Bhanja et al. built a qualitative rule-based system for the same task [23].

Breitfeld et al. built a system that pre-selected potential participants for three clinical trials related to a specific cancer, called rhabdomyosarcoma [24]. Their system asked eight questions about a patient, and used a decision tree to determine a patient's eligibility. The questions did not cover some relevant factors, and a physician had to make a final eligibility decision for pre-selected patients. The authors used trial-specific information in building their system, and they pointed out that extending the system to include other trials would require a major effort.

Fallowfield et al. studied how physicians selected cancer patients for clinical trials, and compared manual and automatic selection [25]. They showed that expert systems could improve the selection accuracy; however, their study also revealed that physicians were reluctant to use these systems. Carlson et al. conducted similar studies with AIDS trials, and also concluded that expert systems could lead to a more accurate selection [26].

Researchers have also investigated various representations of medical knowledge. In particular, Ohno-Machado et al. proposed the GuideLine Interchange Format for medical knowledge [27]. Lindberg et al. considered an alternative format, called the Unified Medical Language System, and developed tools for converting various databases into this format [28]. Rubin et al. analyzed selection criteria for cancer clinical trials and proposed a format for these criteria [29,30]. Wang et al. compared eight previously developed formats and identified main elements of medical knowledge, which included patient data, treatment decisions, and related actions [31].

Eriksson pointed out the need for general-purpose tools that would allow efficient knowledge acquisition, and described a system for building such tools [32]. Tallis et al. developed a library of scripts for modifying knowledge bases, which helped to enforce the consistency of the modified knowledge [33–35]. Blythe et al. designed a general knowledge-acquisition interface based on previous techniques [36]. Musen developed the PROTÉGÉ environment for creating knowledge-acquisition tools [37]; later, researchers used it in the work on AIDS expert systems [38,39], and on an asthma treatment-selection system [40]. Musen et al. extended PROTÉGÉ and built a new version, called PROTÉGÉ-2000 [41].

3. Selection of clinical trials

Physicians at the Moffitt Cancer Center have about 150 clinical trials available for cancer patients. They have identified criteria that determine a patient's eligibility for each trial, and they use these criteria to select trials for eligible patients. Traditionally, physicians have selected trials by a manual analysis of patients' data. The review of resulting selections has shown that they usually do not check all clinical trials and occasionally miss an appropriate trial.

To address this problem, we have developed an expert system that helps to select trials for each

	General information What is the patient's sex? What is the patient's age?
1. The patient is female. 2. She is at most 45 years old. 3. Her cancer stage is II or III. 4. Her cancer is not invasive. 5. At most 3 lymph nodes have cancer. 6. Either • there are no cardiac arrhythmias, • or all tumors are at most 2.5 cm.	*Mammogram*, Cost is \$150 What is the cancer stage? Is the cancer invasive? *Biopsy*, Cost is \$400 What is the cancer stage? How many lymph nodes have cancer? What is the greatest tumor diameter? *Electrocardiogram*, Cost is \$200 Are there cardiac arrhythmias?
(a) Eligibility criteria.	(b) Tests and questions.

Figure 1 Example of eligibility criteria (a), and tests and questions (b).

patient. It prompts a clinician to enter the results of medical tests, and uses them to identify appropriate trials. If the available records do not provide enough data, the system suggests additional tests. We give an example of the selection process, describe the main elements of the knowledge base, and outline the system's web-based interface. We then give experimental results, which confirm that the system helps to find eligible patients and to reduce the related costs.

3.1. Example

In Fig. 1(a), we give a simplified example of eligibility criteria for a clinical trial. This trial is for young and middle-aged women with a noninvasive cancer at stage II or III. When testing a patient's eligibility, a clinician has to order three medical tests (Fig. 1(b)).

The system first prompts a clinician to enter the patient's sex and age. If the patient satisfies the corresponding conditions, the system asks for the mammogram results and verifies Conditions 3 and 4; then, it requests the biopsy and electrocardiogram data. The ordering of tests depends on their costs and on the amount of information provided by test results. The system begins with the mammogram because it is cheaper than the other tests and provides data for two eligibility criteria.

If the patient's records already include some test results, the clinician can answer the corresponding questions while entering the personal data, before the system selects tests. For example, if the records indicate that the cancer stage is IV, the clinician can enter the stage along with the sex and age, and the system immediately determines that the patient is ineligible for this trial.

3.2. Knowledge base

The knowledge base includes questions, medical tests, and logical expressions that represent eligibility criteria for each trial. Since clinicians specify eligibility criteria as hard constraints, without priorities or soft constraints, we allow only hard-constraint logical expressions. The system does not prioritize eligibility criteria, and it treats the results of medical tests in the same way as other data, such as sex, age, and medical history. We give a simplified example of tests and questions in Fig. 1(b), and logical expressions in Fig. 2.

The system supports three types of questions; the first type takes a yes/no response, the second is multiple choice, and the third requires a numeric answer. For example, the cancer stage is a multiple-choice question, and the tumor diameter is a numeric question. The description of a medical test includes the test name, dollar cost, and list of questions that can be answered based on the test results. For instance, the mammogram in Fig. 1 has a cost of US\$ 150, and it allows the answering of two questions. Different tests may answer the same

sex = FEMALE and *age* $\leq$ 45 and *cancer-stage* $\in$ {II, III} and *invasive-cancer* = NO and *lymph-nodes* $\leq$ 3 and (*arrhythmias* = NO or *tumor-diameter* $\leq$ 2.5)	*sex* = MALE or *age* $>$ 45 or *cancer-stage* $\in$ {I, IV} or *invasive-cancer* = YES or *lymph-nodes* $>$ 3 or (*arrhythmias* = YES and *tumor-diameter* $>$ 2.5)
(a) Acceptance expression.	(b) Rejection expression.

Figure 2 Logical expressions for the criteria in Fig. 1(a). The acceptance expression (a) represents the eligibility conditions, whereas the rejection expression (b) describes ineligible patients.

question; for example, both mammogram and biopsy show the cancer stage.

We encode the eligibility for a trial by a logical expression that does not have negations, called the *acceptance expression*. It includes variables that represent the available medical data, as well as equalities, inequalities, "set-element" relations, conjunctions, and disjunctions. For example, we encode the criteria in Fig. 1(a) by the expression in Fig. 2(a). In addition, the system uses the logical complement of the eligibility criteria, called the *rejection expression*, which also does not have negations (Fig. 2(b)); it describes the conditions that make a patient ineligible for the trial.

The system collects data until it can determine which of the two expressions is TRUE. For instance, if a patient's sex is MALE, then the rejection expression in Fig. 2(b) is TRUE, and the system immediately rejects this trial. If the sex is FEMALE, and the other values are unknown, then neither acceptance nor rejection expression is TRUE, and the system asks more questions.

If the knowledge base includes multiple clinical trials, the system checks a patient's eligibility for each of them. It first prompts the clinician to enter the personal data for a patient, then asks for the tests related to multiple trials, and finally requests additional tests for specific trials. After getting each new answer, the system re-evaluates the patient's eligibility for each trial. It displays the list of matching trials, rejected trials, and trials that require additional information.

3.3. Order of tests

If a patient's medical records do not include enough data, the system asks for additional tests; for example, if the records do not provide data for the eligibility criteria in Fig. 1, the system asks for the mammogram, biopsy, and electrocardiogram. The total cost of tests may depend on their order; for instance, if we begin with the mammogram, and it shows that the cancer stage is IV, then we can immediately reject the trial in Fig. 1 and avoid the more expensive tests.

We have explored heuristics for ordering the tests based on the test costs and the structure of acceptance and rejection expressions. The heuristics use a disjunctive normal form of these expressions; that is, each expression must be a disjunction of conjunctions. For example, the rejection expression in Fig. 2(b) is in disjunctive normal form, whereas the acceptance expression in Fig. 2(a) is not. If the system uses ordering heuristics, it has to convert this acceptance expression into the disjunctive normal form shown in Fig. 3.

$$\begin{pmatrix} sex = \text{FEMALE and} \\ age \leq 45 \text{ and} \\ cancer\text{-}stage \in \{\text{II, III}\} \text{ and} \\ invasive\text{-}cancer = \text{NO and} \\ lymph\text{-}nodes \leq 3 \text{ and} \\ arrhythmias = \text{NO} \end{pmatrix} \text{or} \begin{pmatrix} sex = \text{FEMALE and} \\ age \leq 45 \text{ and} \\ cancer\text{-}stage \in \{\text{II, III}\} \text{ and} \\ invasive\text{-}cancer = \text{NO and} \\ lymph\text{-}nodes \leq 3 \text{ and} \\ tumor\text{-}diameter \leq 2.5 \end{pmatrix}$$

Figure 3 Disjunctive normal form of the acceptance expression in Fig. 2.

The system chooses the order of tests that reduces their expected cost. After getting the results of the first test, it re-evaluates the need for the other tests and revises their ordering. The choice of the first test is based on three criteria. The system scores all required tests according to these criteria, computes a linear combination of the three scores for every test, and chooses the test with the highest score.

(1) *Cost of the test*: The system gives preference to cheaper tests. For instance, it may start with the mammogram, which is cheaper than the other two tests in Fig. 1.
(2) *Number of clinical trials that require the test*: When the system checks a patient's eligibility for several trials, it prefers tests that provide data for the largest number of trials. For example, if the electrocardiogram gives data for two different trials, whereas the mammogram provides data for only one trial, the system may prefer the electrocardiogram despite its higher cost.
(3) *Number of clauses that include the test results*: The system prefers the tests that provide data for the largest number of clauses in the acceptance and rejection expressions. For example, the mammogram data affect both clauses of the acceptance expression in Fig. 3 and two clauses of the rejection expression in Fig. 1(b). On the other hand, the electrocardiogram affects only one clause of the acceptance expression and one clause of the rejection expression; thus, the system should order it after the mammogram.

The system disregards the costs of tests performed in the normal course of treatment, and accounts only for the costs related to the selection of clinical trials. For example, if a patient needs the mammogram regardless of trial participation, the system views it as a zero-cost test. On the other hand, if the only purpose of the biopsy and electrocardiogram is to select trials, the system uses heuristics to order these tests.

Although the system suggests the single most effective test, it allows a clinician to order multiple tests at once. For instance, if it indicates that the

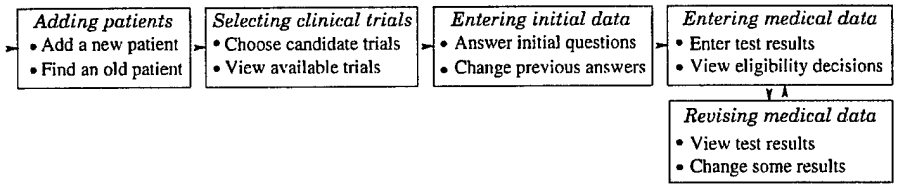


Figure 4 Entering a patient's data. The web-based interface for the data entry consists of five screens. We show these screens by rectangles and the transitions between them by arrows.

mammogram is the best test, the clinician can determine that the electrocardiogram is also an effective test, and order both tests at the same time.

3.4. User interface

The system includes a web-based interface that allows clinicians to enter patients' data through remote computers; the interface consists of five screens (Fig. 4).

The start screen is for adding new patients and retrieving old patients (Fig. 5). After a user enters a patient's name, the system displays a list of the available trials (Fig. 6). The user can choose a subset of these trials, and then the system checks eligibility only for the selected trials. The next screen is for basic personal and medical data, such as sex, age, and cancer stage (Fig. 7).

After the system gets these basic data, it prompts the user for medical information related to specific trials (Fig. 8). When the user enters medical data,

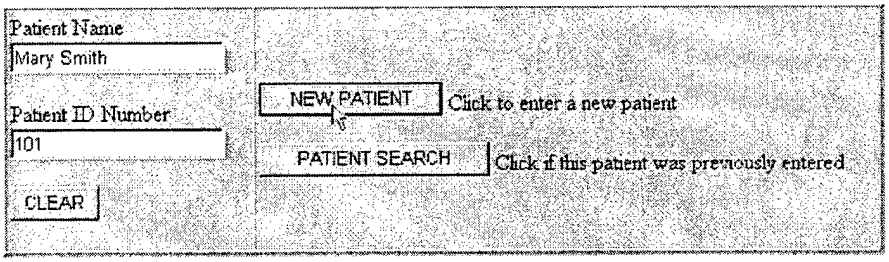


Figure 5 Adding new patients and retrieving existing patients.

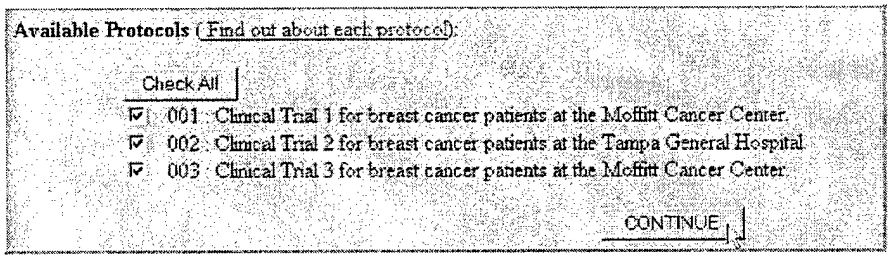


Figure 6 Selecting clinical trials.

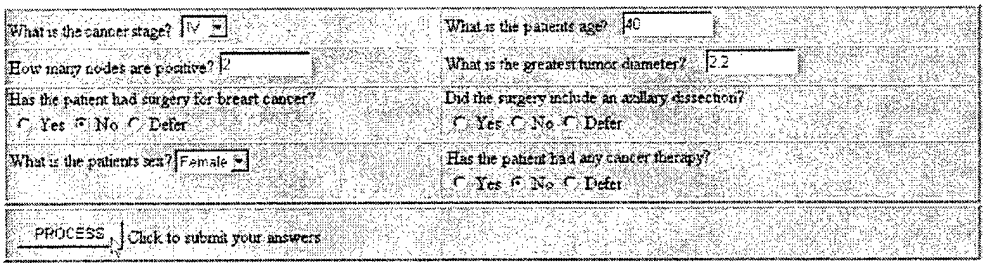


Figure 7 Entering basic information for a patient.

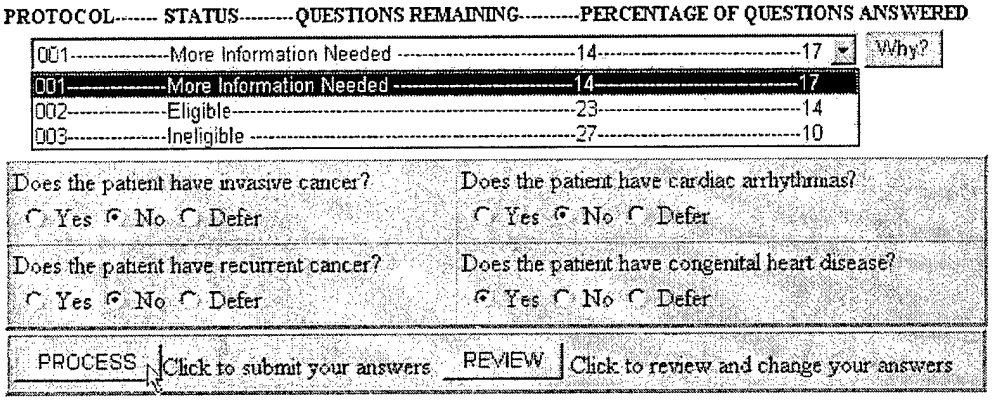


Figure 8 Entering medical data.

the system continuously re-evaluates the patient's eligibility and shows the decision for each trial. If the patient is ineligible for some trials, the user can find out the reasons by clicking the *"Why"* button. The interface also includes a screen for the review and modification of the previous answers, similar to the screen in Fig. 8.

3.5. Experiments

We have built a knowledge base for the breast-cancer clinical trials at the Moffitt Cancer Center, applied the system to the retrospective data from 187 past patients and 74 current patients, and compared the results with manual selection by Moffitt clinicians. The number of matching trials for a patient has ranged from zero to three. For most patients, the system rejects most trials during the initial entry of basic data, such as sex, age, and cancer stage. It usually identifies two to five potential matches based on these basic data, and narrows the selection down to one or two trials based on the following trial-specific questions.

We summarize the results for the past patients in Table 1(a), and the results for the current patients in Table 1(b). The "same matches" column includes the number of patients who have been selected by both human clinicians and the expert system. The "new matches" column gives the number of patients who have been matched by the system but missed by human clinicians. Finally, the last column shows the number of matching patients whose available records are incomplete. Clinicians have found trials for these patients, but the system cannot identify these matches because of insufficient data. Since these patients are no longer at Moffitt, we cannot obtain the missing data; note that this problem is due to the use of retrospective data, and it does not arise when clinicians select trials for new patients.

The system has identified a number of situations when patients were eligible for clinical trials, but did not participate in these trials. We have checked these results with Moffitt clinicians, and they have confirmed that all matches are correct. In most cases, patients did not participate in the matching trials because clinicians missed these matches; however, for some of the past cases, we have been unable to verify that physicians actually missed the matches, rather than having undocumented reasons for omitting them.

We show the mean test costs with and without the ordering heuristics in Table 2, and give a

Table 1 Results of matching 187 past patients and 74 current patients

Clinical trial	Same matches	New matches	Missing data
(a) Results for the 187 past patients			
10822	10	5	0
10840	0	19	3
11072	48	26	19
11378	4	19	3
11992	5	6	0
12100	8	20	13
12101	20	30	0
(b) Results for the 74 current patients			
11132	4	1	1
11931	1	12	0
11971	3	0	0
12100	0	3	0
12101	6	26	1
12601	0	1	2
12775	4	5	1

We give the number of matches found by both the expert system and human clinicians, as well as the number of new matches identified by the system. We also show the number of matches missed by the system because of insufficient data.

Table 2 Cost savings by test reordering

Clinical trial	Average dollar cost	
	Without test reordering	With test reordering
(a) Results for the 187 past patients		
10822	$70	$11
10840	$0	$0
11072	$209	$60
11378	$35	$19
11992	$0	$0
12100	$0	$0
12101	$0	$0
(b) Results for the 74 current patients		
11132	$0	$0
11931	$0	$0
11971	$314	$314
12100	$0	$0
12101	$0	$0
12601	$64	$6
12775	$0	$0

graphical view of the cost savings in Fig. 9. The results confirm that the heuristics reduce the cost of the selection process. Five clinical trials have incurred selection costs; the heuristics have significantly reduced the costs for four of these trials, and have not affected the cost for the fifth trial. The other trials have not incurred costs because all related tests were performed in the normal course of treatment before the trial selection.

3.6. Scalability

The time complexity of evaluating the acceptance and rejection expressions is linear in their total size. Experiments on a Sun Ultra 10 have shown that the evaluation takes about 0.02 s per question, and the time is linear in the number of questions. Typical eligibility conditions for a clinical trial include 10–30 questions; thus, the evaluation time is 0.2–0.6 s per trial.

The linear scalability is an advantage over Bayesian systems, which usually do not scale to a large number of clinical trials [14,42,43]. The authors of these systems have reported that the sizes of the underlying networks are superlinear in the number of trials [44,45], and that the training time is superlinear in the network size [21,22].

If the system uses the cost-reduction heuristics, it converts the acceptance and rejection expressions into disjunctive normal form, which can potentially lead to an explosion in their size. For example, if eligibility conditions are as shown in Fig. 10(a), the system initially generates the expression in Fig. 10(b). If the system converts it to disjunctive normal form, the resulting expression consists of eight clauses.

Although the conversion may result in impractically large expressions, experiments with cancer trials have shown that this problem does not arise in practice because the number of nested disjunctions is usually small. Furthermore, we can eliminate some disjunctions by combining their elements into longer questions. For instance, we can represent Condition 3 in Fig. 10(a) by a single question: "Does the patient have both invasive and recurrent cancer?" If we apply this modification to Conditions 3 and 5, then we obtain the expression in Fig. 10(c), and its conversion to disjunctive normal form results in an expression with two clauses.

4. Entering eligibility criteria

We describe a web-based interface for adding new clinical trials [46], which consists of two main parts; the first part is for information about medical tests (Fig. 11), and the second is for eligibility criteria (Fig. 12). The interface includes fifteen screens; three of them are "start screens," which can be reached from any other screen. We give an example of entering eligibility criteria, describe the two parts of the interface, and present experiments to illustrate its effectiveness.

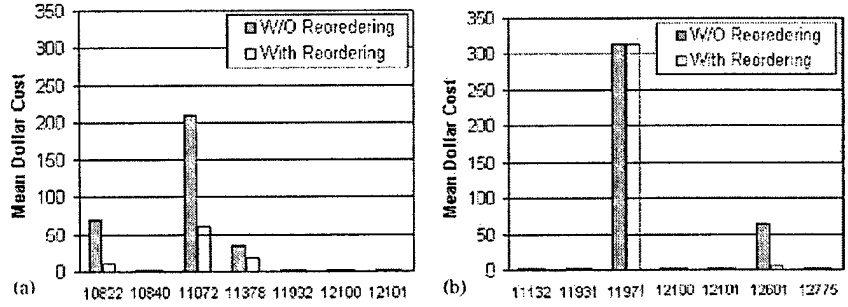


Figure 9 Costs with and without test reordering, for 187 past patients (a) and 74 current patients (b).

1. The patient is female.
2. She is at most 45 years old.
3. Either
   - her cancer is not invasive, or
   - her cancer is not recurrent.
4. Either
   - at most 3 lymph nodes have cancer,
   - or all tumors are at most 2.5 cm.
5. Either
   - there are no cardiac arrhythmias,
   - or there is no congenital heart disease.

(a) Eligibility criteria.

sex = FEMALE and
age $\leq$ 45 and
(*invasive* = NO or
recurrent = NO) and
(*lymph-nodes* $\leq$ 3 or
tumor-size $\leq$ 2.5) and
(*arrhythmias* = NO or
congenital = NO)

(b) Acceptance expression.

sex = FEMALE and
age $\leq$ 45 and
invasive-and-recurrent = NO and
(*lymph-nodes* $\leq$ 3 or
tumor-size $\leq$ 2.5) and
arrhythmias-and-congenital = NO

(c) Reduced expression.

Figure 10 Reducing the number of disjunctions. The conversion of the eligibility criteria (a) into a logical expression (b) leads to an explosion in the size of the corresponding disjunctive normal form. We may prevent the explosion by replacing some disjunctions with single questions (c).

4.1. Example

Suppose that a user needs to enter the criteria shown in Fig. 1, and the system initially has no data about the related tests. The user has to describe the tests and questions, and specify the eligibility conditions.

First, the user utilizes the "*Adding tests*" screen to enter the new tests; we illustrate the entry of a test in Fig. 13. Then, the user adds the related questions; to enter questions for a specific test, the user selects the test and clicks "*Modify*" (Fig. 14), and the system displays the "*Modifying a test*"screen (Fig. 15). To add a question, the user clicks the appropriate button at the bottom (Fig. 15) and then types the question (Fig. 16).

After adding the questions for all tests, the user goes to the "*Adding clinical trials*" screen and initializes a new trial (Fig. 17). The user gets the "*Selecting tests*" screen and chooses the tests related to the current trial (Fig. 18). Then, the user marks relevant questions and the answers that make

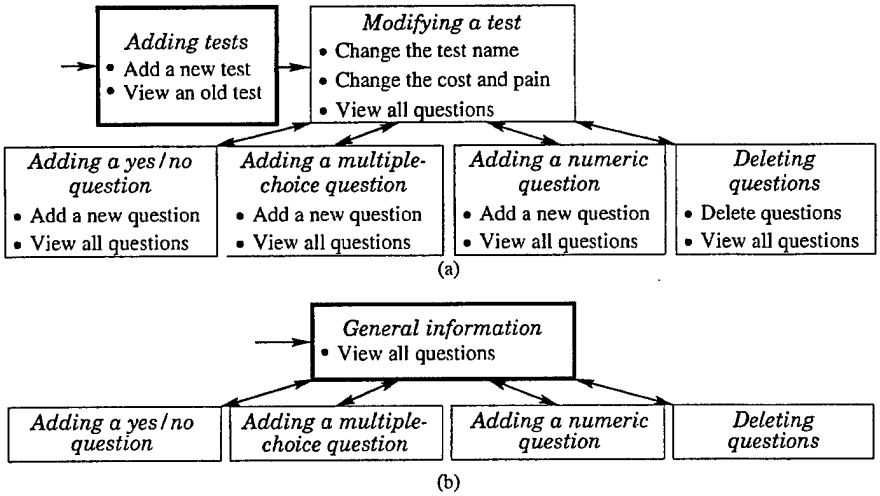


Figure 11 Entering tests and questions (a), and general questions (b). We show the screens by rectangles and the transitions between them by arrows; the bold rectangles are the start screens.

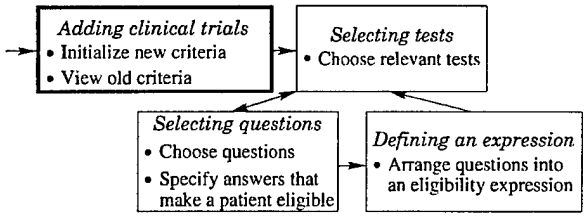


Figure 12 Entering eligibility criteria.

a patient eligible (Fig. 19). If the eligibility criteria include disjunctions, the user has to utilize the screen for composing logical expressions (Fig. 20).

4.2. Tests and questions

We now describe the six-screen interface for adding tests and questions (Fig. 11a). The start screen allows viewing the available tests and defining

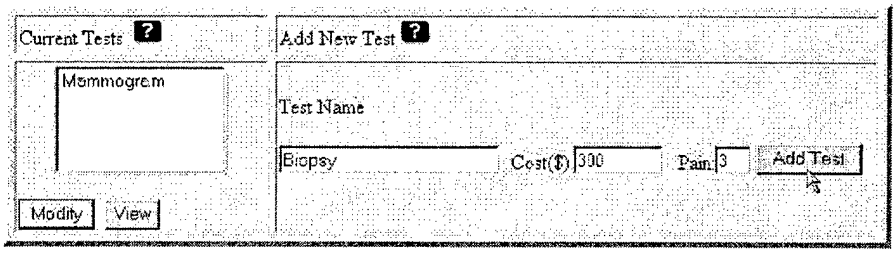


Figure 13 Adding a new test.

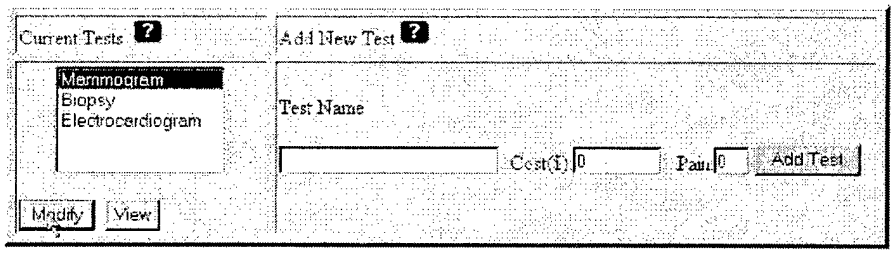


Figure 14 Selecting a test for entering the related questions.

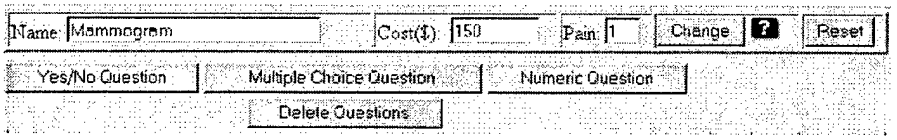
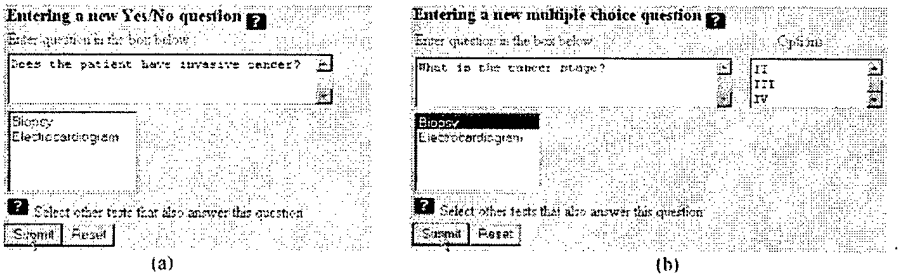


Figure 15 Modifying a test; the bottom buttons are to move to question-entry screens.



(a) (b)

Figure 16 Adding yes/no questions (a) and multiple-choice questions (b); the user types a question and the answer options.

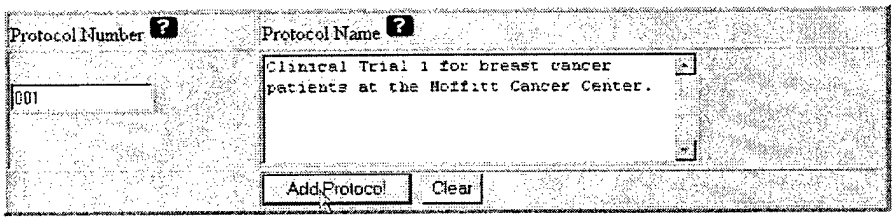


Figure 17 Adding a new clinical trial.

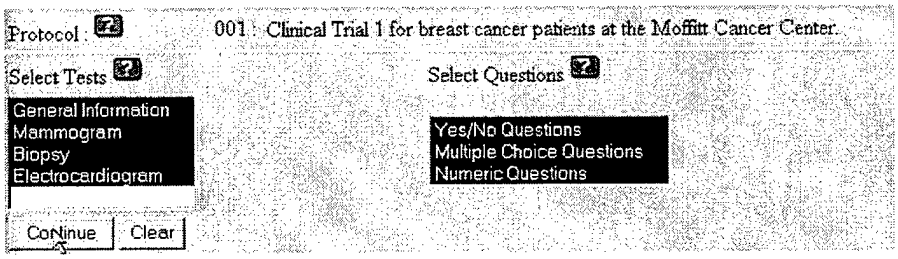


Figure 18 Choosing tests and question types.

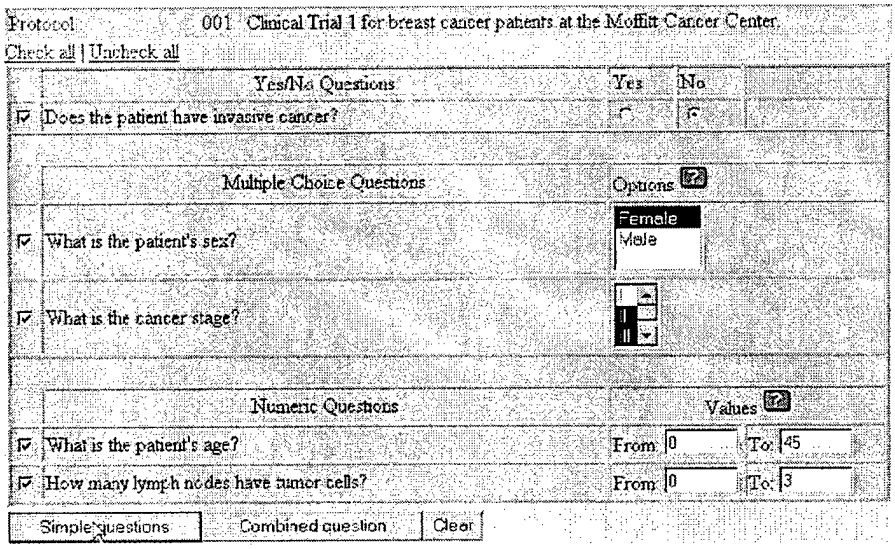


Figure 19 Selecting questions and answers. The user checks the questions for the current clinical trial and marks the answers that satisfy the eligibility criteria.

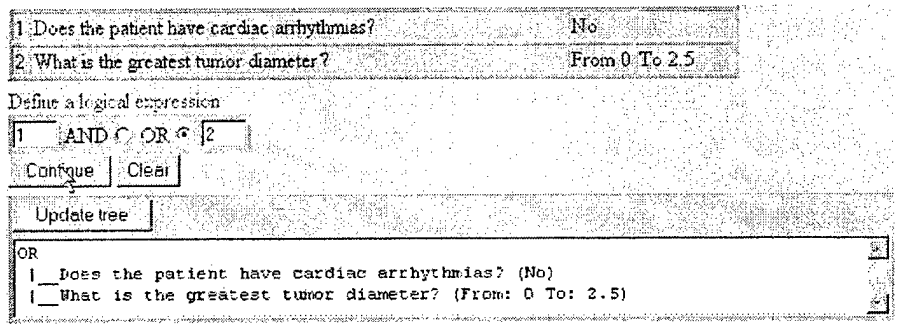


Figure 20 Combining questions into a logical expression.

new ones, whereas the other screens are for modifying tests and adding questions.

We show the start screen in Fig. 13; its left-hand side allows viewing questions and going to a modification screen. If the user selects a test and clicks *"View,"* the system shows the questions related to this test at the bottom of the same screen. If the user clicks *"Modify,"* it displays the *"Modifying a test"* screen (Fig. 15). The right-hand side of the start screen allows adding a new test by specifying its name, cost, and pain level.

The *"Modifying a test"* screen shows the information about a specific test, which includes the test name, cost, pain level, and related questions. The user can change the test name, cost, and pain level; the four bottom buttons allow moving to the screens for adding and deleting questions.

We show the screen for adding yes/no questions in Fig. 16(a) and multiple-choice questions in Fig. 16(b); the screen for numeric questions is similar. The user can enter a new question for the current test, along with a set of allowed answers. If the question is also related to other tests, the user has to mark them in the lower box. The *"Deleting questions"* screen is for removing old questions, which allows modification of old eligibility criteria.

The mechanism for adding general questions, such as sex and age, consists of five screens (Fig. 11b), and the user adds general questions in the same way as test-related questions.

4.3. Eligibility conditions

We next describe the mechanism for entering eligibility criteria, which consists of four screens (Fig. 12). The start screen allows the user to initialize a new clinical trial and view the criteria for old trials. If the user needs to modify a clinical trial, the system first displays the test-selection screen (Fig. 18). The user then chooses related tests and question types, and clicks *"Continue"* to get the question list.

The next screen (Fig. 19) allows the user to select specific questions and mark answers that make a patient eligible. For a multiple-choice question, the user may specify several eligibility options; for example, a patient may be eligible if her cancer stage is II or III. For a numeric question, the user has to specify a range of values; for instance, a patient may be eligible if her age is between 0 and 45 years. If the user clicks *"Simple questions,"* the system generates a conjunction of the selected criteria. If the eligibility conditions involve a more complex expression, the user has to click *"Combined question"* and then use the screen for composing logical expressions (Fig. 20).

The system combines the eligibility criteria into an acceptance expression, and then generates the corresponding rejection expression by recursive application of DeMorgan's laws. If the system uses the cost-reduction heuristics, it converts these expressions into disjunctive normal form using a standard conversion algorithm [47,48].

4.4. Entry time

We have run experiments with 16 novice users, who had no prior experience with the interface. First, every user has entered four sets of medical tests; each set has included three tests and 10 questions. Then, each user has added eligibility expressions for 10 clinical trials used at the Moffitt Cancer Center; the number of questions in an eligibility expression has varied from 10 to 35.

We have measured the entry time for each test set and each clinical trial. In Fig. 21, we show the mean time for every test set and the time per question for the same sets. All users have entered the test sets in the same order; since they had no prior experience, their performance has improved

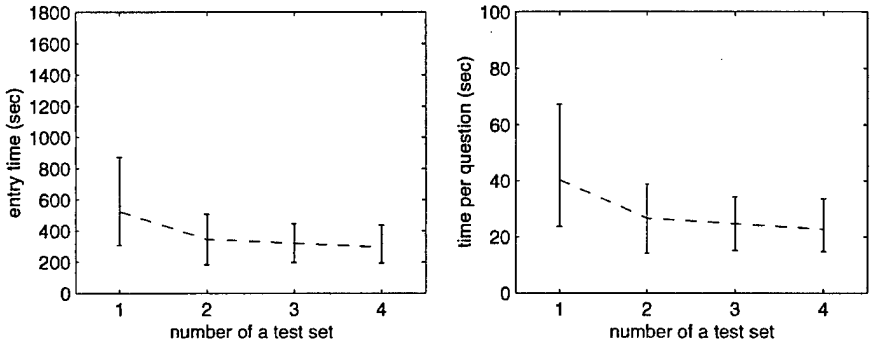


Figure 21 Entry time for test sets (left) and the mean time per question for each set (right). We plot the average time (dashed lines) and the time of the fastest and slowest users (vertical bars).

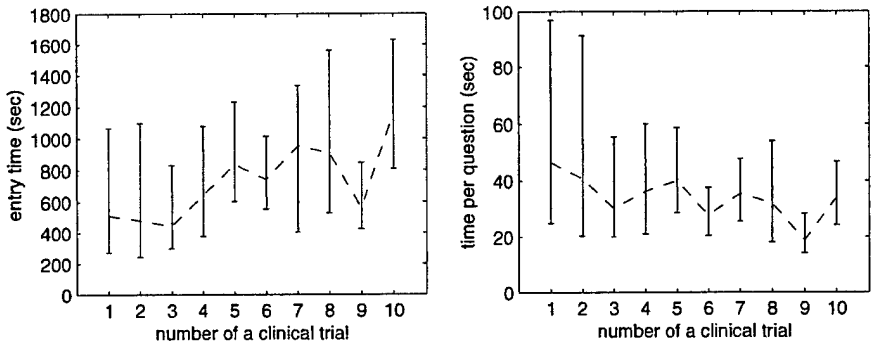


Figure 22 Entry time for eligibility criteria. We show the average time for each clinical trial and the time per question (dashed lines), along with the performance of the fastest and slowest users (vertical bars).

during the experiment. In Fig. 22, we give similar graphs for the entry of trials.

The experiments have shown that novices can efficiently use the interface; they quickly learn its full functionality, and their learning curve reaches a plateau after about an hour. The average time per question is 31 s for the entry of medical tests and 37 s for eligibility criteria, which means that a user can enter all 150 cancer trials used at Moffitt in about 2 weeks.

5. Conclusions

We have developed an expert system that assigns cancer patients to clinical trials. We have described the representation of selection criteria, heuristics for ordering of tests, and a web-based interface for entering patients' data, which will enable physicians across the country to access a central repository of clinical trials. The system also includes an interface for extending its knowledge base, which allows a user to enter a new trial in 10–20 min. Novices can use the interface without prior instructions, and they reach their full speed after about an hour. Although cancer research has provided the motivation for this work, the system is not limited to cancer, and we can use it for trials related to other diseases.

The system uses logical eligibility expressions, similar to those in EON [13] and ONCODOC [15,16]; this approach is different from AIDS^2 [14] and Theocharous's system [21,22], which are based on Bayesian networks. The use of logical expressions ensures scalability and ease of adding new trials, but it does not allow probabilistic decisions based on incomplete data.

We have applied the system to the data from 261 breast-cancer patients admitted to the Moffitt Cancer Center in the last 3 years. The experiments have confirmed that the system can improve the speed and accuracy of selecting trial participants. The results suggest that physicians miss about 60% of matching trials, which means that the system can increase the number of participants by a factor of 2.5. These results are consistent with the studies of the manual trial selection [3–8], which confirm that clinicians miss up to 60% of matches. They are also consistent with the experiment on using ONCODOC at two French hospitals, which has increased the number of selected matches by a factor of 3 [17,18]. We have been unable to compare the results with those of AIDS^2, EON, and Theocharous's system, because the authors of these systems have not reported large-scale clinical experiments.

The developed system includes heuristics for the ordering of medical tests, which is an advantage over the other trial-selection systems. The experiments have shown that the ordering of tests affects their overall cost, and the implemented heuristics reduce this cost.

We now point out some limitations of the developed system and related future challenges. First, the system does not access patient data in the Moffitt clinical database, and nurses have to enter all relevant information through the system's interface. The data in the clinical database are mostly in natural language, as dictated by physicians; we plan to develop a mechanism for transferring these data into the trial-selection system, which will require domain-specific tools for natural-language processing. Second, the system does not keep track of temporal changes in the data; for example, it does not update a patient's age, and does not flag the out-of-date test results. We have recently designed a mechanism for temporal reasoning, and we plan to integrate it with the system. Third, the test-ordering heuristics do not account for the probabilities of possible test results, and we are presently working on the integration of the current heuristics with probabilistic reasoning.

Acknowledgements

We are grateful to Alison Syme for her detailed comments and suggestions. This work has been partially supported by the Breast Cancer Research Program of the U.S. Army Medical Research and Materiel Command under contract DAMD17-00-1-0244, and by the Moffitt Cancer Center.

References

[1] Keppel KG, Pearcy JN, Wagener DK. Trends in racial and ethnic-specific rates for the health status indicators: United States, 1990–98. Healthy people 2000. Statistical notes, vol. 23. 2002.

[2] Cancer facts and figures 2002. American Cancer Society; 2002.

[3] Gennari JH, Reddy M. Participatory design and an eligibility screening tool. In: Proceedings of the American Medical Informatics Association Annual Fall Symposium, 2000. p. 290–4.

[4] Gotay CC. Accrual to cancer clinical trials: directions from the research literature. Soc Sci Med 1991;33(5):569–77.

[5] Kotwall C, Mahoney LJ, Myers RE, Decoste L. Reasons for non-entry in randomized clinical trials for breast cancer: a single institutional study. J Surg Oncol 1992;50:125–9.

[6] Séroussi B, Bouaud J, Antoine É-C. Enhancing clinical practice guideline compliance by involving physicians in the decision process. In Horn W, Shahar Y, Lindberg G, Andreassen S, Wyatt JC, editors. Artificial intelligence in medicine. Berlin, Germany: Springer-Verlag; 1999. p. 76–85.

[7] Tu SW, Kemper CA, Lane NM, Carlson RW, Musen MA. A methodology for determining patients' eligibility for clinical trials. J Methods Inform Med 1993;32(4):317–25.

[8] Yusuf S, Held P, Teo KK, Toretsky ER. Selection of patients for randomized controlled trials: implications of wide or narrow eligibility criteria. Stat Med 1990;9:73–86.

[9] Smith BJ, McNeely MDD. The influence of an expert system for test ordering and interpretation on laboratory investigations. Clin Chem 1999;45(8):1168–75.

[10] Van Walraven C, Naylor CD. Do we know what inappropriate laboratory utilization is? A systematic review of laboratory clinical audits. J Am Med Assoc 1998;280(6):550–8.

[11] Bareford D, Hayling A. Inappropriate use of laboratory services: long term combined approach to modify request patterns. Br Med J 1990;301(6764):1305–7.

[12] Perraro F, Rossi P, Liva C, Bulfoni A, Ganzini G, Giustinelli A. Inappropriate emergency test ordering in a general hospital: preliminary reports. Quality Assurance Health Care 1992;4:77–81.

[13] Musen MA, Tu SW, Das AK, Shahar Y. EON: A component-based approach to automation of protocol-directed therapy. J Am Med Inform Assoc 1996;3(6):367–88.

[14] Ohno-Machado L, Parra E, Henry SB, Tu SW, Musen MA. AIDS^2: A decision-support tool for decreasing physicians' uncertainty regarding patient eligibility for HIV treatment protocols. In: Proceedings of the Seventeenth Annual Symposium on Computer Applications in Medical Care, 1993. p. 429–33.

[15] Bouaud J, Séroussi B, Antoine É-C, Gozy M, Khayat D, Boisvieux J-F. Hypertextual navigation operationalizing generic clinical practice guidelines for patient-specific therapeutic decisions. J Am Med Inform Assoc 1998;5(Suppl):488–92.

[16] Bouaud J, Séroussi B, Antoine É-C, Zelek L, Spielmann M. Reusing ONCODOC, a guideline-based decision support system, across institutions: A successful experiment in sharing medical knowledge. In: Proceedings of the American Medical Informatics Association Annual Symposium, vol. 7, 2000.

[17] Séroussi B, Bouaud J, Antoine É-C. Users' evaluation of ONCODOC, a breast cancer therapeutic guideline delivered at the point of care. J Am Med Inform Assoc 1999;6(5): 384–9.

[18] Séroussi B, Bouaud J, Antoine É-C. ONCODOC: A successful experiment of computer-supported guideline development and implementation in the treatment of breast cancer. Artif Intell Med 2001;22(1):43–64.

[19] Hammond P, Sergot MJ. Computer support for protocol-based treatment of cancer. J Logic Program 1996;26(2):93–111.

[20] McNeely MDD, Smith BJ. An interactive expert system for the ordering and interpretation of laboratory tests to enhance diagnosis and control utilization. Can Med Inform 1995;2(3):16–9.

[21] Papaconstantinou C, Theocharous G, Mahadevan S. An expert system for assigning patients into clinical trials based on Bayesian networks. J Med Syst 1998;22(3): 189–202.

[22] Theocharous G. An expert system for assigning patients into clinical trials based on Bayesian networks. Master's thesis. Computer Science and Engineering Department, University of South Florida; 1996.

[23] Bhanja S, Fletcher LM, Hall LO, Goldgof DB, Krischer JP. A qualitative expert system for clinical trial assignment. In: Proceedings of the Eleventh International Florida Artificial Intelligence Research Society Conference, 1998. p. 84–8.

[24] Breitfeld PP, Ullrich F, Anderson J, Crist WM. Web-based decision support for clinical trial eligibility determination in an international clinical trials network. Controlled Clin Trials 2003;24(6):702–10.

[25] Fallowfield L, Ratcliffe D, Souhami R. Clinicians' attitudes to clinical trials of cancer therapy. Eur J Cancer 1997;33(13):2221–9.

[26] Carlson RW, Tu SW, Lane NM, Lai TL, Kemper CA, Musen MA, Shortliffe EH. Computer-based screening of patients with HIV/AIDS for clinical trial eligibility. Online J Curr Clin Trials 1995;4(179).

[27] Ohno-Machado L, Gennari JH, Murphy S, Jain NL, Tu SW, Oliver DE, Pattison-Gordon E. The guideline interchange format: a model for representing guidelines. J Am Med Inform Assoc 1998;5(4):357–72.

[28] Lindberg DAB, Humphreys BL, McCray AT. The unified medical language system. J Methods Inform Med 1993;4(32):281–91.

[29] Rubin DL, Gennari JH, Musen MA. Knowledge representation and tool support for critiquing clinical trial protocols, In: Proceedings of the American Medical Informatics Association Fall Symposium, 2000. p. 724–8.

[30] Rubin DL, Gennari JH, Srinivas S, Yuen A, Kaizer H, Musen MA, Silva JS. Tool support for authoring eligibility criteria for cancer trials. In: Proceedings of the American Medical Informatics Association Annual Symposium, 1999. p. 369–73.

[31] Wang D, Peleg M, Tu SW, Shortliffe EH, Greenes RA. Representation of clinical practice guidelines for computer-based implementations. Med Inform 2001;10:285–9.

[32] Eriksson H. Specification and generation of custom-tailored knowledge-acquisition tools. In: Proceedings of the

Thirteenth International Joint Conference on Artificial Intelligence, vol. 1, 1993. p. 510–8.
[33] Gil Y, Tallis M. A script-based approach to modifying knowledge bases. In: Proceedings of the Fourteenth National Conference on Artificial Intelligence, 1997. p. 377–83.
[34] Tallis M. A script-based approach to modifying knowledge-based systems. In: Proceedings of the Tenth Innovative Applications of Artificial Intelligence Conference, 1998. p. 1183–95.
[35] Tallis M, Gil Y. Designing scripts to guide users in modifying knowledge-based systems. In: Proceedings of the Sixteenth National Conference on Artificial Intelligence, 1999. p. 242–9.
[36] Blythe J, Kim J, Ramachandran S, Gil Y. An integrated environment for knowledge acquisition. In: Proceedings of the International Conference on Intelligent User Interfaces, 2001. p. 13–20.
[37] Musen MA. Automated generation of model-based knowledge acquisition tools. San Mateo, CA: Morgan Kaufmann; 1989.
[38] Puerta AR, Egar JW, Tu SW, Musen MA. A multiple-method knowledge-acquisition shell for the automatic generation of knowledge-acquisition tools. Knowledge Acquisition 1992;4(2):171–96.
[39] Puerta AR, Eriksson H, Egar JW, Musen MA. Generation of knowledge-acquisition tools from reusable domain ontologies. Technical Report KSL-92-81. Computer Science Department, Stanford University; 1992.
[40] Johnson PD, Musen MA. Development of a guideline authoring tool with PROTÉGÉ-II based on the DILEMMA generic protocol and guideline model. Technical Report SMI-96-0620. Stanford Medical Informatics Group, Stanford School of Medicine; 1996.
[41] Musen MA, Fergerson RW, Grosso WE, Noy NF, Crubezy M, Gennari JH. Component-based support for building knowledge-acquisition systems. In: Proceedings of the Conference on Intelligent Information Processing of the International Federation for Information Processing World Computer Congress, 2000.
[42] Díez FJ, Mira J, Iturralde E, Zubillaga S. DIAVAL, a Bayesian expert system for echocardiography. Artif Intell Med 1997;10(1):59–73.
[43] Onisko A, Druzdzel MJ, Wasyluk H. Application of Bayesian belief networks to diagnosis of liver disorders. In: Proceedings of the Third Conference on Neural Networks and Their Applications, 1997. p. 730–6.
[44] Oniśko A, Druzdzel MJ, Wasyluk H. Learning Bayesian network parameters from small data sets: Application of noisy-Or gates. In: Proceedings of the Workshop on Bayesian and Causal Networks: From Inference to Data Mining, 2000.
[45] Wang H, Druzdzel MJ. User interface tools for navigation in conditional probability tables and elicitation of probabilities in Bayesian networks. In: Proceedings of the Sixteenth Conference on Uncertainty in Artificial Intelligence, 2000. p. 617–25.
[46] Nikiforou S. Selection of clinical trials: Knowledge representation and acquisition. Master's thesis, Department of Computer Science and Engineering, University of South Florida, 2002.
[47] Crama Y, Hammer PL. Boolean functions. Berlin, Germany: Springer-Verlag; 2001.
[48] Kenneth RH. Discrete mathematics and its applications. 3rd ed. New York, NY: McGraw-Hill; 1988.

Experiments on the Automated Selection of Patients for Clinical Trials

Eugene Fink eugene@csee.usf.edu
Lawrence O. Hall hall@csee.usf.edu
Dmitry B. Goldgof goldgof@csee.usf.edu
Bhavesh D. Goswami bgoswami@csee.usf.edu
Matthew Boonstra boonstra@csee.usf.edu
Jeffrey P. Krischer jpkrischer@moffitt.usf.edu

Computer Science and Engineering, University of South Florida, Tampa, Florida 33620

Abstract – *When clinicians test a new treatment procedure, they need to identify and recruit patients with appropriate medical conditions. We have developed an expert system that helps clinicians select patients for experimental treatments, and to reduce the number and overall cost of related medical tests. We describe experiments on selecting patients for new treatments at the Moffitt Cancer Center. The experiments have shown that the system can increase the number of selected patients by a factor of three, and that it can also reduce the cost of the selection process.*

Keywords: Medical expert systems, breast cancer, cost reduction.

1 Introduction

When clinicians conduct treatment experiments, called *clinical trials,* they have to recruit participants from current patients. To select prospective participants, clinicians analyze the data of available patients and identify patients with appropriate medical conditions. This analysis has traditionally been a manual process, and studies have shown that clinicians miss up to 60% of the matching patients, which delays the completion of clinical trials [7, 17].

To address this problem, several researchers built expert systems to help clinicians select trial participants. Ohno-Machado *et al.* developed the AIDS^2 system, which selected AIDS patients for clinical trials [11]. Musen *et al.* built a rule-based system, called EON, that also selected AIDS trial participants [8]. Bouaud *et al.* created the ONCODOC system, which suggested trials for cancer patients [2, 3]. Séroussi *et al.* used ONCODOC to identify trial participants at two hospitals, which helped increase the number of selected patients by a factor of three [13, 14, 15].

The National Cancer Institute created a search engine for selecting clinical trials, available through the Internet at *www.cancer.gov/search/clinical_trials.* It prompts a user to answer several questions about a patient, and gives a list of potentially matching trials; however, it does not determine whether the patient satisfies all of the requirements of these trials.

Fallowfield *et al.* studied how physicians selected cancer patients for clinical trials, and compared manual and automated selection [5]. They showed that expert systems could improve the selection accuracy, but physicians were reluctant to use these systems. Carlson *et al.* conducted similar studies with AIDS trials, and also concluded that expert systems could lead to a more accurate selection [4].

A recent project at the University of South Florida has also been aimed at automated identification of prospective trial participants. Theocharous developed a Bayesian system that selected clinical trials for cancer patients [12, 16], and Bhanja *et al.* built a qualitative rule-based system for the same task [1].

We have continued their work, built a new version of the rule-based system [6, 9, 10], and applied it to selecting patients for breast-cancer trials at the Moffitt Cancer Center, located on campus of the University of South Florida. We outline the design of this system and present an empirical evaluation of its effectiveness.

2 Knowledge base

Physicians at the Moffitt Cancer Center currently conduct about 150 clinical trials. We have developed an expert system to help physicians select trials for eligible patients; it consists of a knowledge base and a web-based interface for entering patient data. The knowledge base contains information about related medical

0-7803-7952-7/03/$17.00 © 2003 IEEE.

(a) MEDICAL TESTS

General information
What is the patient's sex?
What is the patient's age?

Mammogram, Cost is $150
What is the cancer stage?
Does the patient have invasive cancer?

Biopsy, Cost is $400
How many lymph nodes have tumor cells?
What is the greatest tumor diameter?

Electrocardiogram, Cost is $200
Does the patient have cardiac arrhythmias?

(b) ELIGIBILITY CRITERIA

sex = FEMALE and
age $\leq$ 45 and
cancer-stage $\in$ {II, III} and
invasive-cancer = NO and
lymph-nodes $\leq$ 3 and
(*arrhythmias* = NO or
tumor-diameter $\leq$ 2.5)

Figure 1: Description of medical tests and trial-eligibility criteria in the trial-selection system.

tests, as well as logical expressions that determine a patient's eligibility for each trial. The description of a medical test includes its dollar cost and list of questions that can be answered based on the test results (Figure 1a). The trial-eligibility criteria are represented by a logical expression, which includes variables that represent the patient data, as well as equalities, inequalities, "set-element" relations, conjunctions, and disjunctions (Figure 1b).

The system collects data until it can determine whether the eligibility expression is TRUE or FALSE. For example, if a clinician uses the system to determine a patient's eligibility for the trial in Figure 1(b), it first asks about the patient's sex and age. If the patient satisfies the corresponding conditions, it asks for the mammogram results, and then requests the biopsy and electrocardiogram data. The ordering of tests depends on their costs and on the amount of information provided by test results. The system begins with the mammogram because it is cheaper than the other tests and provides data for two clauses of the eligibility expression.

3 Selection of participants

We have built a knowledge base for the breast-cancer trials at the Moffitt Cancer Center, including five completed trials and ten current trials, and applied the system to retrospective data from the Moffitt patients who have had a breast-cancer surgery in the last three years. We have discarded the patients whose available records are incomplete, and used all remaining patients, which include 187 past patients and 169 current patients.

We have compared the results of automated trial selection for these patients with the manual selection by Moffitt clinicians. The system has identified all eligible patients for each trial, whereas the clinicians have selected about one-third of the eligible patients. We summarize the results for the past patients in Table 1(a), and the results for the current patients in Table 1(b). The "participants" column shows the number of actual participants of each trial; the "other eligible" column gives the number of the other eligible patients identified by the system.

For every current patient who did not participate in a matching trial, we have checked whether she participated in any other trial, and we show the results in Table 2. We have not done a similar analysis for the past patients due to insufficient data. The "incompatible" column in Table 2 includes the number of eligible patients who did not participate in a specified trial because of participation in another incompatible trial. The "compatible" column shows the number of patients who participated in another compatible trial, and could also have participated in the specified trial. Finally, the "no other trial" column gives the number of eligible patients who have not participated in any trial.

The results show that the system can identify eligible patients who have not been selected by clinicians; thus, it can increase the number of trial participants. For the patients in the reported experiments, it could increase the overall number of participants by a factor of three. In particular, it has found prospective participants for some trials with a very small number of manually selected patients. For example, it has found nineteen matching patients for trial 12385, which currently has no participants, and twenty-six patients for trial 11931, which has only two participants.

4 Cost reduction

If the available patient records do not provide enough data for trial selection, clinicians perform medical tests as part of the selection process. They can reduce the overall test cost by first ordering inexpensive tests, and then using their results to avoid some expensive tests.

The system suggests the ordering of tests that reduces their expected cost. After getting the results of the first test, it re-evaluates the need for the other tests and revises their ordering. The choice of the first test is based on three criteria. The system scores all required tests according to these criteria, computes a linear combination of the three scores for every test, and chooses the test with the highest score.

Table 1: Results of selecting clinical trials for the 187 past patients and 169 current patients. We give the number of trial participants, selected by both the system and Moffitt clinicians, and the number of the other eligible patients, identified by the system.

(a) Results for the 187 past patients.

Clinical Trial	Participants	Other Eligible
10822	10	5
10840	0	19
11072	48	26
11378	4	19
11992	5	6
12100	8	20
12101	20	30

(b) Results for the 169 current patients.

Clinical Trial	Participants	Other Eligible
11132	4	1
11931	2	26
11971	4	0
12100	0	5
12101	11	52
12385	0	19
12601	0	1
12643	16	36
12757	1	3
12775	23	17

Table 2: Participation of the patients who skipped a matching clinical trial in other trials. We show the number of patients who skipped the trial because of participation in another incompatible trial; the number of patients who were on another trial compatible with the skipped trial; and the number of eligible patients who were not on any trial.

Clinical Trial	Incompatible	Compatible	No Other Trial
11132	0	1	0
11931	0	11	15
11971	0	0	0
12100	0	1	4
12101	13	6	33
12385	8	2	9
12601	0	0	1
12643	0	10	26
12757	0	1	2
12775	3	3	11

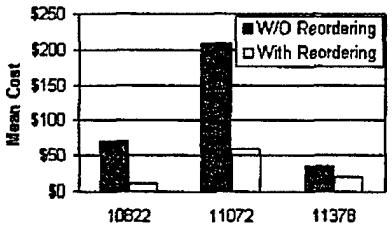
Table 3: Cost savings by test reordering.

(a) Results for the 187 past patients.

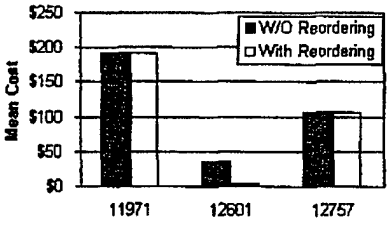
Clinical Trial	Mean Cost	
	W/O Test Reordering	With Test Reordering
10822	$70	$11
10840	$0	$0
11072	$209	$60
11378	$35	$19
11992	$0	$0
12100	$0	$0
12101	$0	$0

(b) Results for the 169 current patients.

Clinical Trial	Mean Cost	
	W/O Test Reordering	With Test Reordering
11132	$0	$0
11931	$0	$0
11971	$192	$192
12100	$0	$0
12101	$0	$0
12385	$0	$0
12601	$36	$3
12643	$0	$0
12757	$107	$107
12775	$0	$0



(a) Results for the 187 past patients.



(b) Results for the 169 current patients.

Figure 2: Costs with and without test reordering. We plot the results for the six clinical trials that have incurred nonzero selection costs.

1. *Cost of a test.* The system gives preference to less expensive tests.
2. *Immediate decision.* If a test can lead to an immediate acceptance or rejection of the trial, the system prefers it to other tests.
3. *Number of related clauses.* The system prefers the tests that provide data for large number of clauses in the eligibility expression.

The system disregards the costs of tests performed in the normal course of treatment, and accounts only for the costs related to the trial selection. For example, if a patient needs a mammogram regardless of trial participation, the system views it as a zero-cost test. On the other hand, if the only purpose of the biopsy and eléctrocardiogram is to select trials, the system uses the heuristics to order these tests.

We show the mean test costs with and without the ordering heuristics in Table 3, and give a graphical view of the cost savings in Figure 2. The results confirm that the heuristics reduce the cost of the selection process. Six clinical trials have incurred selection costs; the heuristics have reduced the costs for four of these trials, and have not affected the costs for the other two trials.

The results in Table 3(a) differ from similar experiments with an earlier version of the system [6], because of two changes to the system. First, the current version disregards the costs of the tests required for the regular treatment, which do not affect the trial-selection expenses, whereas the earlier version counted all costs. Second, some costs in the old system were out-of-date, and we have corrected them based on the data from the Moffitt accounting department.

5 Reduction of data entry

The system tries to minimize not only the overall cost of medical tests, but also the amount of data entry, that is, the number of questions asked about a patient. For each question, it estimates the probability that the answer will lead to an immediate acceptance or rejection of the trial, and it gives preference to the questions with the highest probability of an immediate decision. Thus, when a clinician enters the available data, the system asks the related questions in the decreasing order of the immediate-decision probabilities. It estimates these probabilities from past experience with other patients. For each question, it determines the percentage of past answers that have led to immediate decisions, and uses this percentage as the probability estimate.

We have evaluated the effectiveness of this ordering heuristic for six clinical trials, using the data from the 169 current patients. We have performed ten-fold cross-validation; that is, we have used 90% of the patients to compute the related probabilities, and then measured the mean number of questions for the other 10%.

We show the results with and without the ordering heuristic in Table 4, and give a graphical view of the same results in Figure 3. The heuristic has reduced the number of questions for all six trials; the reduction ranges from 1% to 29%, and its mean is 15%. The results confirm that the accumulated statistical data help reduce the number of questions.

Table 4: Reduction of data entry by the reordering of questions, for the 169 current patients.

Clinical Trial	Mean Number of Questions	
	W/O Question Reordering	With Question Reordering
11931	18.9	15.4
12100	14.0	13.9
12101	24.8	21.7
12385	19.1	14.8
12601	15.7	13.9
12775	16.1	14.4

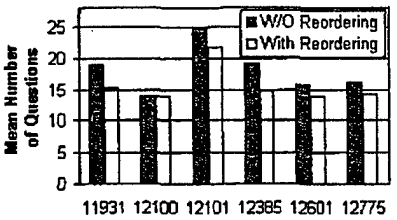


Figure 3: Number of questions with and without the reordering heuristic, for the 169 current patients.

6 Concluding remarks

We have developed an expert system that selects clinical trials for eligible patients. Experiments have confirmed that the system can increase the number of clinical-trial participants. They have also shown that the ordering of related medical tests affects the overall test cost, and the implemented heuristics can reduce this cost.

Acknowledgments: This work has been partially supported by the Moffitt Cancer Center and by the Breast Cancer Research Program of the U.S. Army Medical Research and Materiel Command under contract DAMD17-00-1-0244.

References

[1] Sanjukta Bhanja, Lynn M. Fletcher, Lawrence O. Hall, Dmitry B. Goldgof, and Jeffrey P. Krischer. A qualitative expert system for clinical trial assignment. In *Proceedings of the Eleventh International Florida Artificial Intelligence Research Society Conference*, pages 84–88, 1998.

[2] Jacques Bouaud, Briggite Séroussi, Éric-Charles Antoine, Mary Gozy, David Khayat, and Jean-

François Boisvieux. Hypertextual navigation operationalizing generic clinical practice guidelines for patient-specific therapeutic decisions. *Journal of the American Medical Informatics Association*, 5(suppl.):488–492, 1998.

[3] Jacques Bouaud, Briggite Séroussi, Éric-Charles Antoine, Laurent Zelek, and Marc Spielmann. Reusing ONCODOC, a guideline-based decision support system, across institutions: A successful experiment in sharing medical knowledge. In *Proceedings of the American Medical Informatics Association Annual Symposium*, volume 7, 2000.

[4] Robert W. Carlson, Samson W. Tu, Nancy M. Lane, Tze L. Lai, Carol A. Kemper, Mark A. Musen, and Edward H. Shortliffe. Computer-based screening of patients with HIV/AIDS for clinical trial eligibility. *Online Journal of Current Clinical Trials*, 4(179), 1995.

[5] Lesley Fallowfield, D. Ratcliffe, and Robert Souhami. Clinicians' attitudes to clinical trials of cancer therapy. *European Journal of Cancer*, 33(13):2221–2229, 1997.

[6] Princeton K. Kokku, Lawrence O. Hall, Dmitry B. Goldgof, Eugene Fink, and Jeffrey P. Krischer. A cost-effective agent for clinical trial assignment. In *Proceedings of the* IEEE *International Conference on Systems, Man, and Cybernetics*, 2002.

[7] Cyrus Kotwall, Leo J. Mahoney, Robert E. Myers, and Linda Decoste. Reasons for non-entry in randomized clinical trials for breast cancer: A single institutional study. *Journal of Surgical Oncology*, 50:125–129, 1992.

[8] Mark A. Musen, Samson W. Tu, Amar K. Das, and Yuval Shahar. EON: A component-based approach to automation of protocol-directed therapy. *Journal of the American Medical Informatics Association*, 3(6):367–388, 1996.

[9] Savvas Nikiforou. Selection of clinical trials: Knowledge representation and acquisition. Master's thesis, Department of Computer Science and Engineering, University of South Florida, 2002.

[10] Savvas Nikiforou, Eugene Fink, Lawrence O. Hall, Dmitry B. Goldgof, and Jeffrey P. Krischer. Knowledge acquisition for clinical-trial selection. In *Proceedings of the* IEEE *International Conference on Systems, Man, and Cybernetics*, 2002.

[11] Lucila Ohno-Machado, Eduardo Parra, Suzanne B. Henry, Samson W. Tu, and Mark A. Musen. AIDS^2: A decision-support tool for decreasing physicians' uncertainty regarding patient eligibility for HIV treatment protocols. In *Proceedings of the Seventeenth Annual Symposium on Computer Applications in Medical Care*, pages 429–433, 1993.

[12] Constantinos Papaconstantinou, Georgios Theocharous, and Sridhar Mahadevan. An expert system for assigning patients into clinical trials based on Bayesian networks. *Journal of Medical Systems*, 22(3):189–202, 1998.

[13] Briggite Séroussi, Jacques Bouaud, and Éric-Charles Antoine. Users' evaluation of ONCODOC, a breast cancer therapeutic guideline delivered at the point of care. *Journal of the American Medical Informatics Association*, 6(5):384–389, 1999.

[14] Briggite Séroussi, Jacques Bouaud, and Éric-Charles Antoine. ONCODOC: A successful experiment of computer-supported guideline development and implementation in the treatment of breast cancer. *Artificial Intelligence in Medicine*, 22(1):43–64, 2001.

[15] Briggite Séroussi, Jacques Bouaud, Éric-Charles Antoine, Laurent Zelek, and Marc Spielmann. Using ONCODOC as a computer-based eligibility screening system to improve accrual onto breast cancer clinical trials. In Silvana Quaglini, Pedro Barahona, and Steen Andreassen, editors, *Artificial Intelligence in Medicine*, pages 421–430. Springer-Verlag, Berlin, Germany, 2001.

[16] Georgios Theocharous. An expert system for assigning patients into clinical trials based on Bayesian networks. Master's thesis, Computer Science and Engineering Department, University of South Florida, 1996.

[17] Samson W. Tu, Carol A. Kemper, Nancy M. Lane, Robert W. Carlson, and Mark A. Musen. A methodology for determining patients' eligibility for clinical trials. *Journal of Methods of Information in Medicine*, 32(4):317–325, 1993.

A Cost-Effective Agent for Clinical Trial Assignment

Princeton K. Kokku, Lawrence O. Hall, Dmitry B. Goldgof, Eugene Fink, and Jeffrey P. Krischer

kokku@csee.usf.edu, hall@csee.usf.edu, goldgof@csee.usf.edu,
eugene@csee.usf.edu, jpkrischer@moffitt.usf.edu

Computer Science and Engineering, University of South Florida, Tampa, Florida 33620

Abstract— **The purpose of a clinical trial is to evaluate a new treatment procedure. When medical researchers conduct a trial, they recruit participants with appropriate medical histories. To select participants, the researchers analyze medical records of the available patients, which has traditionally been a manual procedure. We describe an intelligent agent that helps to select patients for clinical trials. If the available data are insufficient for choosing patients, the agent suggests additional medical tests and finds an ordering of the tests that reduces their total cost.**

Keywords— **Medical expert systems, automated diagnosis, clinical trials.**

I. Introduction

A *clinical trial* is an experiment with a new treatment procedure. When medical researchers test a new treatment, they recruit patients with appropriate health problems and medical histories. The selection of patients has traditionally been a manual procedure, and recent studies have shown that clinicians can miss up to 60% of the eligible patients [9, 10, 14, 26, 35, 38].

If the available records do not provide enough data, clinicians perform medical tests as part of the selection process. The costs of most tests have declined over the last decade, but the number of tests has significantly increased [33, 36], which is partially due to inappropriate ordering of tests [1, 25]. Clinicians can reduce the cost by first requiring inexpensive tests and then using their results to avoid some expensive tests; however, finding the right ordering may be a complex problem.

The purpose of the described work is to automate the selection of patients for clinical trials and minimize the cost of related tests. We have developed an agent that identifies appropriate trials for each patient, and built a knowledge base for breast-cancer trials.

II. Previous Work

Researchers began to work on medical expert systems in the early seventies. Shortliffe *et al.* developed the MYCIN system, which diagnosed bacterial diseases [5, 30, 31]. Its knowledge base consisted of *if-then* rules, which allowed for the analysis of symptoms and evaluation of the certainty of the diagnosis. Experiments showed that MYCIN correctly diagnosed common diseases, which led to the development of other medical systems [5, 19], such as NEOMYCIN, PUFF, CENTAUR, and VM. Shortliffe *et al.* created a system for selecting chemotherapy treatments, called ONCOCIN [32], which also evolved from MYCIN.

Lucas *et al.* constructed a rule-based system for diagnosing liver and biliary-tract diseases [16], but it often gave an incorrect diagnosis [12, 23]. Korver and Lucas converted the initial system into a Bayesian network, which improved its performance [13, 15].

Musen *et al.* built a rule-based system, called EON, that selected AIDS patents for clinical trials [20]. Ohno-Machado *et al.* developed the AIDS^{2}sy stem, which also assigned AIDS patients to clinical trials [21]. They integrated logical rules with Bayesian networks, which helped to make decisions in the absence of some data.

Bouaud *et al.* created a cancer expert system, called ONCODOC, that suggested alternative clinical trials for each patient and allowed a physician to choose among them [3, 4]. Séroussi *et al.* used ONCODOC to select participants for clinical trials at two hospitals, which helped to increase the number of selected patients by a factor of three [27, 28, 29].

Hammond and Sergot created the OaSiS architecture [11], which combined the techniques from earlier systems, including EON and ONCOCIN. Smith *et al.* built a system that assisted a clinician in selecting medical tests and reducing their number and cost [17, 18, 33].

Fallowfield *et al.* studied how physicians selected cancer patients for clinical trials, and compared manual and automatic selection [8]. They showed that expert systems could improve the selection accuracy; however, their study also revealed that physicians were reluctant to use these systems. Carlson *et al.* conducted similar studies with AIDS trials, and also concluded that expert systems could lead to a more accurate selection [6].

Theocharous developed a Bayesian system that selected clinical trials for cancer patients [24, 34]. It learned conditional probabilities of medical-test outcomes and evaluated the probability of a patient's eligibility for each trial. On the negative side, the available medical records were often insufficient for learning accurate probabilities. Furthermore, when adding a new clinical trial, the user had to change the structure of the underlying Bayesian network.

To address these problems, Bhanja *et al.* built a rule-based system for the same task [2]. We have continued that work, extended the system, and added a mechanism for reducing costs involved in patient selection.

III. Example

We have developed an intelligent agent that helps to select clinical trials for eligible patients. It prompts a clinician to enter the results of medical tests, and identifies appropriate trials. If the available records do not provide enough data, the agent suggests additional tests.

In Figure 1(a), we give a simplified example of eligibil-

(a) Eligibility criteria

1. The patient is female.
2. She is at most forty-five years old.
3. Her cancer stage is II or III.
4. Her cancer is not invasive.
5. At most three lymph nodes have tumor cells.
6. Either
   - the patient has no cardiac arrhythmias, or
   - all tumors are smaller than 2.5 centimeters.

(b) Tests and questions

General information
What is the patient's sex?
What is the patient's age?

Mammogram, Cost is $150
What is the cancer stage?
Does the patient have invasive cancer?

Biopsy, Cost is $300
What is the cancer stage?
How many lymph nodes have tumor cells?
What is the greatest tumor size?

Electrocardiogram, Cost is $200
Does the patient have cardiac arrhythmias?

Fig. 1. Example of eligibility criteria, tests, and questions.

(a) Acceptance	(b) Rejection
$sex = \text{FEMALE}$ and $age \leq 45$ and $stage \in \{\text{II}, \text{III}\}$ and $invasive = \text{NO}$ and $lymph\text{-}nodes \leq 3$ and ($arrhythmias = \text{NO}$ or $tumor\text{-}size \leq 2.5$)	$sex = \text{MALE}$ or $age > 45$ or $cancer \in \{\text{I}, \text{IV}\}$ or $invasive = \text{YES}$ or $lymph\text{-}nodes > 3$ or ($arrhythmias = \text{YES}$ and $tumor\text{-}size > 2.5$)

Fig. 2. Logical expressions for the criteria in Figure 1(a).

ity criteria for a clinical trial. This trial is for young and middle-aged women with a noninvasive cancer at stage II or III. When testing a patient's eligibility, a clinician has to order three medical tests (Figure 1b).

The agent first prompts a clinician to enter the patient's sex and age. If the patient satisfies the corresponding conditions, the agent asks for the mammogram results and verifies Conditions 3 and 4; then, it requests the biopsy and electrocardiogram data. If the patient's records already include some test results, the clinician can answer the corresponding questions while entering the personal data, before the agent selects test procedures. For example, if the records indicate that the cancer stage is IV, the clinician can enter the stage along with sex and age, and then the agent immediately determines that the patient is ineligible for this trial.

IV. Knowledge Base

The agent's knowledge base includes questions, medical tests, and logical expressions that represent eligibility criteria for each trial. We give a simplified example of tests and questions in Figure 1(b), and logical expressions in Figure 2.

$$\left(\begin{array}{l} sex = \text{FEMALE and} \\ age \leq 45 \text{ and} \\ stage \in \{\text{II}, \text{III}\} \text{ and} \\ invasive = \text{NO and} \\ lymph\text{-}nodes \leq 3 \text{ and} \\ arrhythmias = \text{NO} \end{array}\right) \text{or} \left(\begin{array}{l} sex = \text{FEMALE and} \\ age \leq 45 \text{ and} \\ stage \in \{\text{II}, \text{III}\} \text{ and} \\ invasive = \text{NO and} \\ lymph\text{-}nodes \leq 3 \text{ and} \\ tumor\text{-}size \leq 2.5 \end{array}\right)$$

Fig. 3. Disjunctive normal form of the acceptance expression.

The agent supports three types of questions; the first type takes a yes/no response, the second is multiple choice, and the third requires a numeric answer. For example, the cancer stage is a multiple-choice question, and the tumor size is a numeric question. The description of a medical test includes the test name, dollar cost, and list of questions that can be answered based on the test results (Figure 1).

We encode the eligibility for a clinical trial by a logical expression that does not have negations, called the *acceptance expression*. It includes variables that represent medical data, as well as equalities, inequalities, "set-element" relations, conjunctions, and disjunctions (Figure 2a). In addition, the agent uses the logical complement of the eligibility criteria, called the *rejection expression*, which also does not have negations (Figure 2b). It describes the conditions that make a patient ineligible for the trial.

The agent collects data until it can determine which of the two expressions is TRUE. For instance, if a patient's sex is MALE, then the rejection expression in Figure 2(b) is TRUE, and the agent immediately determines that this trial is inappropriate. If the sex is FEMALE, the agent asks more questions.

If the knowledge base includes multiple clinical trials, the agent checks a patient's eligibility for each of them. It first asks for the tests related to multiple trials, and then requests additional tests for specific trials. After getting each new answer, the agent re-evaluates the patient's eligibility for each trial.

V. Order of Tests

If a patient's records do not include enough data, the agent asks for additional tests; for example, if the records do not provide data for the eligibility criteria in Figure 1, the agent asks for the mammogram, biopsy, and electrocardiogram. The total cost of tests may depend on their order; for instance, if we begin with the mammogram, and it shows that the cancer stage is IV, then we can immediately reject the trial in Figure 1 and avoid the more expensive tests.

We have explored heuristics for ordering the tests, based on the test costs and the structure of acceptance and rejection expressions. The heuristics use a disjunctive normal form of these expressions: that is, each expression must be a disjunction of conjunctions. For example, the rejection expression in Figure 2(b) is in disjunctive normal form, whereas the acceptance expression in Figure 2(a) is not. If the system uses ordering heuristics, it converts this acceptance expression into the disjunctive normal form shown in Figure 3.

The agent chooses the order of tests that reduces their expected cost. After getting the results of the first test, it re-evaluates the need for the other tests and revises their ordering. The choice of the first test is based on three criteria. The agent scores all required tests according to these criteria, computes a linear combination of the three scores for every test, and chooses the test with the highest score.

1. Cost of the test. The agent prefers cheaper tests. For instance, it may start with the mammogram, which is cheaper than the other two tests in Figure 1.

2. Number of clinical trials that require the test. When the agent checks a patient's eligibility for several trials, it prefers tests that provide data for the largest number of trials. For example, if the electrocardiogram gives data for two different trials, the agent may prefer it to the mammogram despite its higher cost.

3. Number of clauses that include the test results. The agent prefers the tests that provide data for the largest number of clauses in the acceptance and rejection expressions. For example, the mammogram data affect both clauses of the acceptance expression in Figure 3 and two clauses of the rejection expression in Figure 1(b). On the other hand, the electrocardiogram affects only one clause of the acceptance expression and one clause of the rejection expression; thus, the agent should order it after the mammogram.

VI. User Interface

The agent includes a web-based interface that allows clinicians to enter patients' data through remote computers; the interface consists of five screens (Figure 4).

The start screen is for adding and retrieving patients (Figure 5). After a user enters a patient's name, the agent displays a list of the available trials (Figure 6). The user can choose a subset of these trials, and then the agent checks eligibility only for the selected trials. The next screen is for basic personal and medical data, such as sex, age, and cancer stage (Figure 7).

After the agent gets the basic data, it prompts the user for medical information related to specific trials (Figure 8). When the user enters medical data, the agent continuously re-evaluates the patient's eligibility and shows the decision for each trial. If the patient is ineligible for some trials, the user can find out the reasons by clicking the *"Why"* button. The interface also includes a screen for the review and modification of the previous answers, similar to the screen in Figure 8.

VII. Experiments

We have built a knowledge base for the breast-cancer clinical trials at the H. Lee Moffitt Cancer Center, applied the agent to retrospective data from 187 past patients and 57 current patients, and compared the results with manual selection by clinicians at the cancer center.

We summarize the results for the past patients in Table I, and the results for the current patients in Table II. The "same matches" column includes the number of patients who have been selected by both human clinicians and the automated agent. The "new matches" column gives the number of patients who have been matched by the agent but potentially missed by human clinicians. Finally, the last column shows the number of patients whose available records are incomplete. Clinicians have found trials for these patients, but the agent cannot identify these matches because of missing data. The agent has found a number of matches potentially missed by human clinicians; thus, it can help to recruit more patients for clinical trials.

TABLE I
RESULTS OF MATCHING 187 PAST PATIENTS.

Clinical Trial	Same Matches	New Matches	Missing Data
10822	10	5	0
10840	0	19	3
11072	48	26	19
11378	4	19	3
11992	5	6	0
12100	8	20	13
12101	20	30	0

TABLE II
RESULTS OF MATCHING 57 CURRENT PATIENTS.

Clinical Trial	Same Matches	New Matches	Missing Data
11132	4	1	1
11971	3	0	0
12100	0	2	0
12101	4	21	0
12601	0	1	0
11931	1	8	0
12775	1	4	0

In Table III, we give the mean test costs with and without the ordering heuristics for the 187 past patients. The results show that the implemented heuristics reduce the costs by more than a factor of two.

VIII. Scalability

The time complexity of evaluating the acceptance and rejection expressions is linear in their size. Experiments on a Sun Ultra 10 have shown that the evaluation takes about 0.02 seconds per question, and the time is linear in the number of questions. Typical eligibility conditions for a clinical trial include ten to thirty questions; thus, the evaluation time is 0.2 to 0.6 seconds per trial.

TABLE III
COST SAVINGS BY TEST REORDERING.

Clinical Trial	Average Dollar Cost	
	Without Test Reordering	With Test Reordering
10822	$20	$8
10840	$0	$0
11072	$556	$194
11378	$34	$0
11992	$87	$34
12100	$0	$0
12101	$24	$22

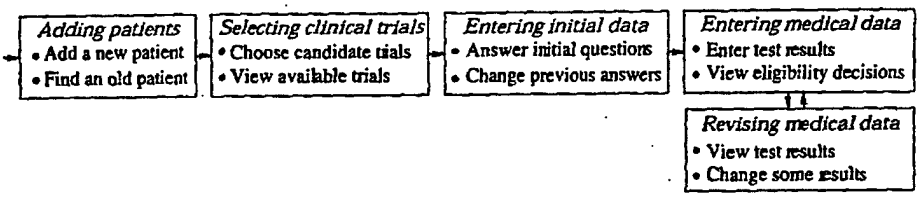


Fig. 4. Entering a patient's data. The web-based interface for data entry consists of five screens. We show these screens by rectangles and the transitions between them by arrows.

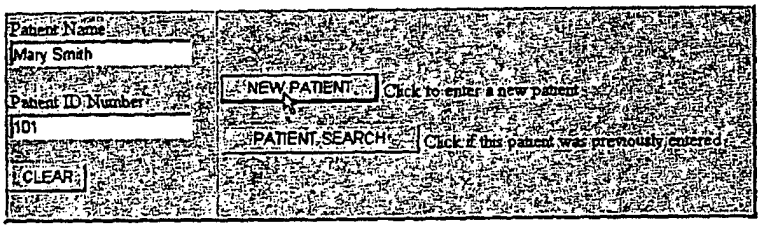


Fig. 5. Adding new patients and retrieving existing patients.

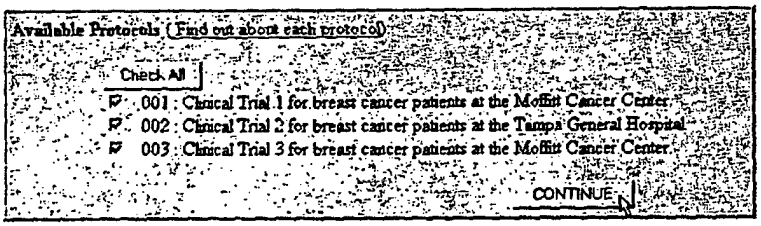


Fig. 6. Selecting clinical trials.

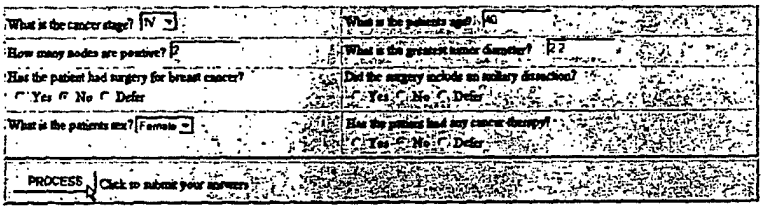


Fig. 7. Entering basic information for a patient.

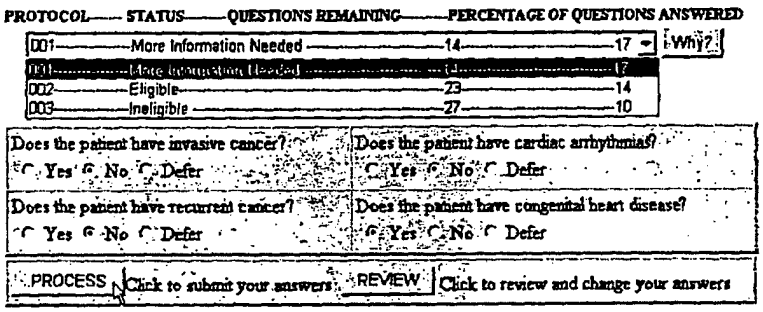


Fig. 8. Entering medical data.

(a) Eligibility criteria

1. The patient is female.
2. She is at most forty-five years old.
3. Either
   - her cancer is not invasive, or
   - her cancer is not recurrent.
4. Either
   - at most three lymph nodes have tumor cells, or
   - all tumors are smaller than 2.5 centimeters.
5. Either
   - the patient has no cardiac arrhythmias, or
   - the patient has no congenital heart disease.

(b) Acceptance expression

sex = FEMALE and
age $\leq$ 45 and
(*invasive* = NO or *recurrent* = NO) and
(*lymph-nodes* $\leq$ 3 or *tumor-size* $\leq$ 2.5) and
(*arrhythmias* = NO or *congenital* = NO)

(c) Reduced expression

sex = FEMALE and
age $\leq$ 45 and
invasive-and-recurrent = NO and
(*lymph-nodes* $\leq$ 3 or *tumor-size* $\leq$ 2.5) and
arrhythmias-and-congenital = NO

Fig. 9. Reducing the number of disjunctions. The conversion of the eligibility criteria (a) into a logical expression (b) leads to an explosion in the size of the corresponding disjunctive normal form. We can prevent the explosion by replacing some disjunctions with single questions (c).

The linear scalability is an important advantage over Bayesian systems, which do not scale to a large number of clinical trials [7, 21, 23]. The authors of these systems have reported that the sizes of the underlying networks are superlinear in the number of trials [22, 37], and the training time is superlinear in the network size [24, 34].

If the agent uses the cost-reduction heuristics, it converts the acceptance and rejection expressions into disjunctive normal form, which can potentially lead to an explosion in their size. For example, if eligibility conditions are as shown in Figure 9(a), the agent initially generates the expression in Figure 9(b). If the agent converts it to disjunctive normal form, the resulting expression consists of eight clauses.

Although the conversion may result in impractically large expressions, experiments have shown that this problem does not arise in practice because the number of nested disjunctions is usually small. Furthermore, we can eliminate some disjunctions by combining their elements into longer questions. For instance, we can represent Condition 3 in Figure 9(a) by a single question: "Does the patient have both invasive and recurrent cancer?" If we apply this modification to Conditions 3 and 5, then we obtain the expression in Figure 9(c), and its conversion to disjunctive normal form results in an expression with two clauses.

IX. Concluding Remarks

We have developed an agent that automatically assigns patients to clinical trials. We have described the representation of selection criteria, heuristics for ordering of tests, and a web-based interface for entering patients' data, which will enable physicians across the country to access a central repository of clinical trials.

Experiments have confirmed that the agent has the potential to find more participants for clinical trials. They have also shown that the ordering of medical tests affects their overall cost, and the implemented heuristics can reduce the cost of finding trial participants. The heuristics do not account for the probabilities of possible test results, and we plan to add probabilistic reasoning as part of the future work.

Acknowledgments: This work has been partially supported by the Breast Cancer Research Program of the U.S. Army Medical Research and Materiel Command under contract DAMD17-00-1-0244, and by the H. Lee Moffitt Cancer Center.

References

[1] D. Bareford and A. Hayling. Inappropriate use of laboratory services: Long term combined approach to modify request patterns. *British Medical Journal*, 301(6764):1305–1307, 1990.

[2] Sanjukta Bhanja, Lynn M. Fletcher, Lawrence O. Hall, Dmitry B. Goldgof, and Jeffrey P. Krischer. A qualitative expert system for clinical trial assignment. In *Proceedings of the Eleventh International Florida Artificial Intelligence Research Society Conference*, pages 84–88, 1998.

[3] Jacques Bouaud, Briggite Séroussi, Éric-Charles Antoine, Mary Gozy, David Khayat, and Jean-François Boisvieux. Hypertextual navigation operationalizing generic clinical practice guidelines for patient-specific therapeutic decisions. *Journal of the American Medical Informatics Association*, 5(suppl.):488–492, 1998.

[4] Jacques Bouaud, Briggite Séroussi, Éric-Charles Antoine, Laurent Zelek, and Marc Spielmann. Reusing ONCODOC, a guideline-based decision support system, across institutions: A successful experiment in sharing medical knowledge. In *Proceedings of the American Medical Informatics Association Annual Symposium*, volume 7, 2000.

[5] Bruce G. Buchanan and Edward H. Shortliffe. *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project.* Addison-Wesley, Reading, MS, 1984.

[6] Robert W. Carlson, Samson W. Tu, Nancy M. Lane, Tze L. Lai, Carol A. Kemper, Mark A. Musen, and Edward H. Shortliffe. Computer-based screening of patients with HIV/AIDS for clinical trial eligibility. *Online Journal of Current Clinical Trials*, 4(179), 1995.

[7] Francisco J. Diez, José Mira, E. Iturralde, and S. Zubillaga. DIAVAL, a Bayesian expert system for echocardiography. *Artificial Intelligence in Medicine*, 10(1):59–73, 1997.

[8] Lesley Fallowfield, D. Ratcliffe, and Robert Souhami. Clinicians' attitudes to clinical trials of cancer therapy. *European Journal of Cancer*, 33(13):2221–2229, 1997.

[9] John H. Gennari and Madhu Reddy. Participatory design and an eligibility screening tool. In *Proceedings of the American Medical Informatics Association Annual Fall Symposium*, pages 290–294, 2000.

[10] Carolyn Cook Gotay. Accrual to cancer clinical trials:

Directions from the research literature. *Social Science and Medicine*, 33(5):569–577, 1991.

[11] Peter Hammond and Marek J. Sergot. Computer support for protocol-based treatment of cancer. *Journal of Logic Programming*, 26(2):93–111, 1996.

[12] M. Korver and A. R. Janssens. Development and validation of HEPAR, an expert system for the diagnosis of disorders of the liver and biliary tract. *Medical Informatics*, 16(3):259–270, 1993.

[13] M. Korver and Peter J. F. Lucas. Converting a rule-based expert system into a belief network. *Medical Informatics*, 18(3):219–241, 1993.

[14] Cyrus Kotwall, Leo J. Mahoney, Robert E. Myers, and Linda Decoste. Reasons for non-entry in randomized clinical trials for breast cancer: A single institutional study. *Journal of Surgical Oncology*, 50:125–129, 1992.

[15] Peter J. F. Lucas. Refinement of the HEPAR expert system: Tools and techniques. *Journal of Artificial Intelligence in Medicine*, 6(2):175–188, 1994.

[16] Peter J. F. Lucas, R. W. Segaar, and A. R. Janssens. HEPAR: An expert system for the diagnosis of disorders of the liver and the biliary tract. *Liver*, 9:266–275, 1989.

[17] Michael D. McNeely and Beverly J. Smith. An interactive expert system for the ordering and interpretation of laboratory tests to enhance diagnosis and control utilization. *Canadian Medical Informatics*, 2(3):16–19, 1995.

[18] Ian R. Morrison, B. A. Schaefer, and Beverly J. Smith. Knowledge acquisition: The ACQUIRE approach. In *Proceedings of the First Semi-Annual Conference in Policy Making and Knowledge Systems*, 1991.

[19] Mark A. Musen. *Automated Generation of Model-Based Knowledge Acquisition Tools*. Morgan Kaufmann, San Mateo, CA, 1989.

[20] Mark A. Musen, Samson W. Tu, Amar K. Das, and Yuval Shahar. EON: A component-based approach to automation of protocol-directed therapy. *Journal of the American Medical Informatics Association*, 3(6):367–388, 1996.

[21] Lucila Ohno-Machado, Eduardo Parra, Suzanne B. Henry, Samson W. Tu, and Mark A. Musen. AIDS^2: A decision-support tool for decreasing physicians' uncertainty regarding patient eligibility for HIV treatment protocols. In *Proceedings of the Seventeenth Annual Symposium on Computer Applications in Medical Care*, pages 429–433, 1993.

[22] Agnieszka Oniśko, Marek J. Druzdzel, and Hanna Wasyluk. Learning Bayesian network parameters from small data sets: Application of noisy-OR gates. In *Proceedings of the Workshop on Bayesian and Causal Networks: From Inference to Data Mining*, 2000.

[23] Agnieszka Onisko, Mark J. Druzdzel, and Hanna Wasyluk. Application of Bayesian belief networks to diagnosis of liver disorders. In *Proceedings of the Third Conference on Neural Networks and Their Applications*, pages 730–736, 1997.

[24] Constantinos Papaconstantinou, Georgios Theocharous, and Sridhar Mahadevan. An expert system for assigning patients into clinical trials based on Bayesian networks. *Journal of Medical Systems*, 22(3):189–202, 1998.

[25] Franco Perraro, Paolo Rossi, Carlo Liva, Adolfo Bulfoni, G. Ganzini, and Adriano Giustinelli. Inappropriate emergency test ordering in a general hospital: Preliminary reports. *Quality Assurance Health Care*, 4:77–81, 1992.

[26] Briggite Séroussi, Jacques Bouaud, and Éric-Charles Antoine. Enhancing clinical practice guideline compliance by involving physicians in the decision process. In Werner Horn, Yuval Shahar, Greger Lindberg, Steen Andreassen, and Jeremy C. Wyatt, editors, *Artificial Intelligence in Medicine*, pages 76–85. Springer-Verlag, Berlin, Germany, 1999.

[27] Briggite Séroussi, Jacques Bouaud, and Éric-Charles Antoine. Users' evaluation of ONCODOC, a breast cancer therapeutic guideline delivered at the point of care. *Journal of the American Medical Informatics Association*, 6(5):384–389, 1999.

[28] Briggite Séroussi, Jacques Bouaud, and Éric-Charles Antoine. ONCODOC: A successful experiment of computer-supported guideline development and implementation in the treatment of breast cancer. *Artificial Intelligence in Medicine*, 22(1):43–64, 2001.

[29] Briggite Séroussi, Jacques Bouaud, Éric-Charles Antoine, Laurent Zelek, and Marc Spielmann. Using ONCODOC as a computer-based eligibility screening system to improve accrual onto breast cancer clinical trials. In Silvana Quaglini, Pedro Barahona, and Steen Andreassen, editors, *Artificial Intelligence in Medicine*, pages 421–430. Springer-Verlag, Berlin, Germany, 2001.

[30] Edward H. Shortliffe. MYCIN: *A Rule-Based Computer Program for Advising Physicians Regarding Antimicrobial Therapy Selection*. PhD thesis, Computer Science Department, Stanford University, 1974.

[31] Edward H. Shortliffe, Randall Davis, Stanton G. Axline, Bruce G. Buchanan, Cordell C. Green, and Stanley Cohen. Computer-based consultations in clinical therapeutics: Explanation and rule acquisition capabilities of the MYCIN system. *Computers and Biomedical Research*, 8:303–320, 1975.

[32] Edward H. Shortliffe, A. Carlisle Scott, Miriam B. Bischoff, William van Melle, and Charlotte D. Jacobs. ONCOCIN: An expert system for oncology protocol management. In *Proceedings of the Seventh International Joint Conference in Artificial Intelligence*, pages 876–881, 1981.

[33] Beverly J. Smith and Michael D. McNeely. The influence of an expert system for test ordering and interpretation on laboratory investigations. *Clinical Chemistry*, 45(8):1168–1175, 1999.

[34] Georgios Theocharous. An expert system for assigning patients into clinical trials based on Bayesian networks. Master's thesis, Computer Science and Engineering Department, University of South Florida, 1996.

[35] Samson W. Tu, Carol A. Kemper, Nancy M. Lane, Robert W. Carlson, and Mark A. Musen. A methodology for determining patients' eligibility for clinical trials. *Journal of Methods of Information in Medicine*, 32(4):317–325, 1993.

[36] Carl Van Walraven and C. David Naylor. Do we know what inappropriate laboratory utilization is? A systematic review of laboratory clinical audits. *Journal of the American Medical Association*, 280(6):550–558, 1998.

[37] Haiqin Wang and Marek J. Druzdzel. User interface tools for navigation in conditional probability tables and elicitation of probabilities in bayesian networks. In *Proceedings of the Sixteenth Conference on Uncertainty in Artificial Intelligence*, pages 617–625, 2000.

[38] Salim Yusuf, Peter Held, K. K. Teo, and Elizabeth R. Toretsky. Selection of patients for randomized controlled trials: Implications of wide or narrow eligibility criteria. *Statistics in Medicine*, 9:73–86, 1990.

Knowledge Acquisition for Clinical-Trial Selection

Savvas Nikiforou, Eugene Fink, Lawrence O. Hall, Dmitry B. Goldgof, and Jeffrey P. Krischer

nikiforo@csee.usf.edu, eugene@csee.usf.edu, hall@csee.usf.edu,
goldgof@csee.usf.edu, jpkrischer@moffitt.usf.edu

Computer Science and Engineering, University of South Florida, Tampa, Florida 33620

Abstract— **When medical researchers test a new treatment procedure, they recruit patients with appropriate medical histories. An experiment with a new procedure is called a clinical trial. The selection of patients for clinical trials has traditionally been a labor-intensive task, which involves the matching of medical records with a list of eligibility criteria, and studies have shown that clinicians can miss up to 60% of the eligible patients. A recent project at the University of South Florida has been aimed at the automation of this task. We have developed an intelligent agent that selects trials for eligible patients. We report the work on the representation and entry of the related knowledge about clinical trials. We describe the structure of the agent's knowledge base and the interface for adding new trials.**

Keywords— **Knowledge representation, medical expert systems, user interfaces.**

I. Introduction

Cancer causes 550,000 deaths in the United States every year, and the treatment of cancer is an active research area. Medical experts explore new treatment methods, such as drugs, surgery techniques, and radiation therapies. An experiment with a new treatment procedure is called a *clinical trial*. When researchers conduct a trial, they recruit patients with an appropriate cancer type and medical history. The selection of patients has traditionally been a manual procedure, and studies have shown that clinicians can miss up to 60% of the eligible patients [12, 22, 30].

A recent project at the University of South Florida has been aimed at automatic selection of patients for clinical trials. We have developed an intelligent agent that prompts a clinician for a patient's data and identifies all matching trials [1, 11]. It includes a knowledge base with information about available clinical trials, criteria for selecting patients, and related medical tests.

We report the work on a web-based interface that enables a clinician to enter new trials without the help of a programmer. We have used the interface to build a knowledge base for clinical trials at the Moffitt Cancer Center, located at the University of South Florida. We review the previous work on medical expert systems (Section II), explain the knowledge representation in the developed agent (Section III), and describe the interface for adding new knowledge (Section IV).

II. Previous Work

Researchers began to work on medical applications of artificial intelligence in the early seventies. Shortliffe and his colleagues developed the MYCIN system, which diagnosed bacterial diseases [5, 25, 26]. Experiments showed the effectiveness of MYCIN, which led to the development of other medical systems [5, 14], such as NEOMYCIN, PUFF, CENTAUR, and VM.

Musen *et al.* built a rule-based system, called EON, that selected AIDS patents for clinical trials [17]. Ohno-Machado *et al.* developed the AIDS^{2}sy stem, which also assigned AIDS patients to clinical trials [19]. Bouaud *et al.* created a cancer expert system, called ONCODOC, that suggested alternative trials for each patient and allowed a physician to choose among them [3, 4]. Séroussi used ONCODOC to select participants for clinical trials at two hospitals, which helped to increase the number of selected patients by a factor of three [23, 24].

Early expert systems did not have knowledge-acquisition tools, and programmers hand-coded the related rules. To simplify knowledge entry, researchers implemented specialized tools for some systems [13, 15].

Eriksson pointed out the need for tools that would allow efficient knowledge acquisition, and described a system for building such tools [6]. Tallis *et al.* developed a library of scripts for modifying knowledge bases, which helped to enforce the consistency of the modified knowledge [7, 27, 28, 29]. Kim and Gil considered the use of scripts for building new knowledge-acquisition tools, and created a system for evaluating these tools [9, 10]. Blythe *et al.* designed a general knowledge-acquisition interface based on previous techniques [2].

Musen developed the PROTÉGÉ environment for creating knowledge-acquisition tools [14, 16], which proved effective for the development of knowledge systems, including the AIDS expert systems [20], asthma treatment selection [8], and elevator-design rules [21].

III. Knowledge Base

Physicians at the Moffitt Cancer Center have about 150 clinical trials available for cancer patients. They have identified criteria that determine a patient's eligibility for each trial, and they use these criteria to select trials for eligible patients. Traditionally, physicians have selected trials by a manual analysis of patients' data. The review of resulting selections has shown that they usually do not check all clinical trials and occasionally miss an appropriate trial.

To address this problem, we have built an intelligent agent that helps to select trials for each patient. It prompts a clinician to enter the results of medical tests, and uses them to identify appropriate trials.

In Figure 1(a), we give a simplified example of eligibility criteria for a clinical trial. This trial is for young and

(a) Eligibility criteria

1. The patient is female.
2. She is at most forty-five years old.
3. Her cancer stage is II or III.
4. Her cancer is not invasive.
5. At most three lymph nodes have tumor cells.
6. Either
   - the patient has no cardiac arrhythmias, or
   - all tumors are smaller than 2.5 centimeters.

(b) Tests and questions

General information
What is the patient's sex?
What is the patient's age?

Mammogram, Cost is $150
What is the cancer stage?
Does the patient have invasive cancer?

Biopsy, Cost is $300
What is the cancer stage?
How many lymph nodes have tumor cells?
What is the greatest tumor diameter?

Electrocardiogram, Cost is $200
Does the patient have cardiac arrhythmias?

(c) Eligibility expression.

$sex =$ FEMALE and
$age \leq 45$ and
$cancer\text{-}stage \in \{II, III\}$ and
$invasive\text{-}cancer =$ NO and
$lymph\text{-}nodes \leq 3$ and
$(arrhythmias =$ NO or
$tumor\text{-}diameter \leq 2.5)$

Fig. 1. Example of eligibility criteria, tests, and questions.

middle-aged women with a noninvasive cancer at stage II or III. When testing a patient's eligibility, a clinician has to order three medical tests (Figure 1b). The agent first prompts the clinician to enter the patient's sex and age. If the patient satisfies the corresponding conditions, the agent asks for the mammogram results and verifies Conditions 3 and 4; then, it requests the biopsy and electrocardiogram data.

The agent's knowledge base includes questions, tests, and logical expressions that represent eligibility for each trial. We give an example of tests and questions in Figure 1(b), and a logical expression in Figure 1(c).

The agent supports three types of questions; the first type takes a yes/no response, the second is multiple choice, and the third requires a numeric answer. For example, the cancer stage is a multiple-choice question, and the tumor diameter is a numeric question. The description of a medical test includes the test name, dollar cost, and list of questions that can be answered based on the test results. For instance, the mammogram in Figure 1 has a cost of $150, and it allows the answering of two questions. Different tests may answer the same question; for example, both mammogram and biopsy show the cancer stage.

We encode the eligibility for a clinical trial by a logical expression, which may include variables that represent the available medical data, as well as equalities, inequalities, "set-element" relations, conjunctions, and disjunctions. For example, we encode the criteria in Figure 1(a) by the expression in Figure 1(c).

The agent collects data until it can determine whether the eligibility expression is TRUE or FALSE. For instance, if a patient's sex is MALE, then the expression in Figure 1(c) is FALSE, and the agent immediately rejects this trial. If the sex is FEMALE, the agent has to ask more questions. If the knowledge base includes many clinical trials, the agent checks a patient's eligibility for each of them. It first asks for the tests related to multiple trials, and then requests additional tests for specific trials.

IV. Entering Eligibility Criteria

We have designed a web-based interface for adding new clinical trials [18], which consists of two main parts; the first part is for adding information about medical tests (Figure 2), and the second is for eligibility criteria (Figure 3). The interface includes ten screens; two of them are "start screens," which can be reached from any other screen. We give an example of entering eligibility criteria, describe the two parts of the interface, and present experiments on its effectiveness.

Example: Suppose that a user needs to enter the criteria shown in Figure 1. First, she utilizes the *"Adding tests"* screen to enter the three tests (Figure 4). Then, she adds the related questions; to enter questions for a specific test, she selects the test and clicks *"Modify"* (Figure 4), and the agent displays the *"Modifying a test"* screen (Figure 5). To add a question, she clicks the appropriate button at the bottom (Figure 5) and then types the question (Figure 6).

After adding the questions for all tests, the user goes to the *"Adding clinical trials"* screen and initializes a new trial (Figure 7). She gets the *"Selecting tests"* screen and chooses the tests related to the current trial (Figure 8). Then, she marks relevant questions and the answers that make a patient eligible (Figure 9). If the eligibility criteria include disjunctions, she has to use the screen for composing logical expressions (Figure 10).

Tests and questions: The interface for adding tests and questions includes six screens (Figure 2). The start screen is for viewing the available tests and defining new ones, whereas the other screens are for modifying tests and adding questions.

We show the start screen in Figure 4; its left-hand side allows viewing questions and going to a modification screen. If the user selects a test and clicks *"View,"* the agent shows the questions related to this test. If the user clicks *"Modify,"* it displays the *"Modifying a test"* screen (Figure 5). The right-hand side of the start screen allows adding a new test by specifying its name and cost.

The *"Modifying a test"* screen shows the information about a specific test, which includes the test name, cost, and related questions. The user can change the test name and cost; the four bottom buttons allow moving to the screens for adding and deleting questions.

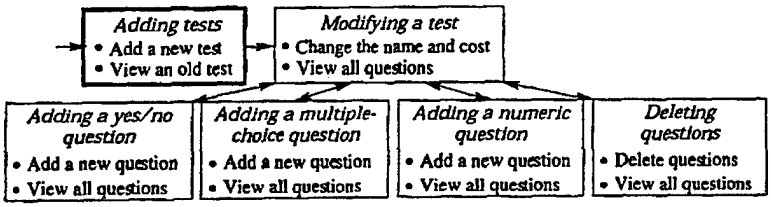


Fig. 2. Entering tests and questions. We show the screens by rectangles and the transitions between them by arrows. The bold rectangle is the start screen.

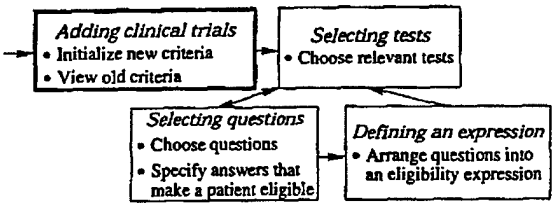


Fig. 3. Entering eligibility criteria.

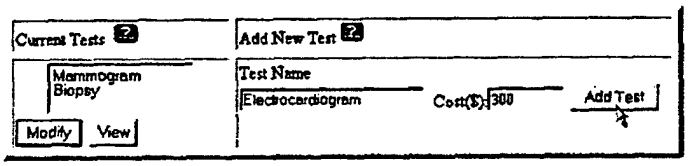


Fig. 4. Adding a new test.

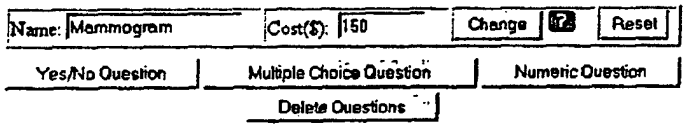
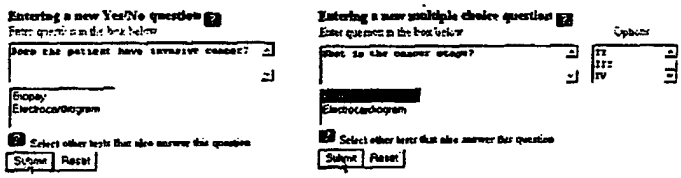


Fig. 5. Modifying a test; the bottom buttons are for moving to question-entry screens.



(a) Yes/no question. (b) Multiple-choice question.

Fig. 6. Adding new questions; the user enters a question and answer options.

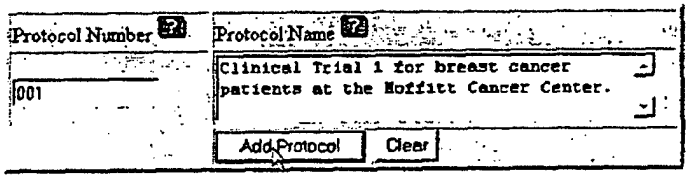


Fig. 7. Adding a new clinical trial.

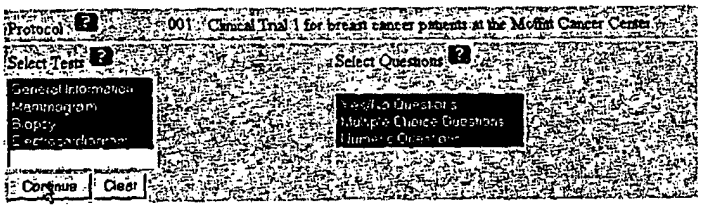


Fig. 8. Choosing tests and question types.

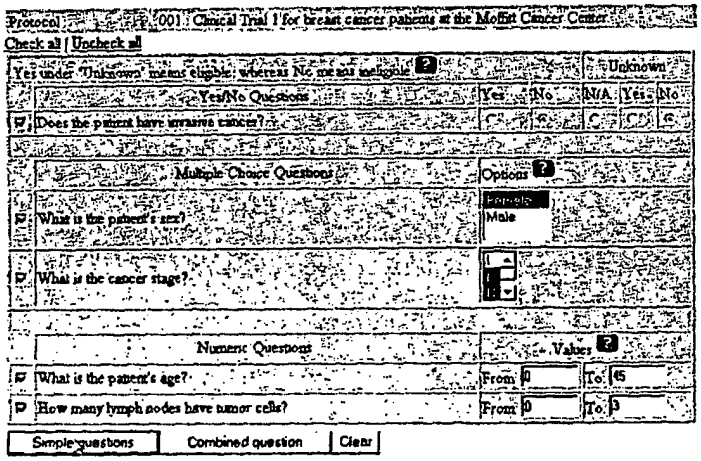


Fig. 9. Selecting questions and answers. The user checks the questions for the current clinical trial and marks the answers that satisfy the eligibility criteria.

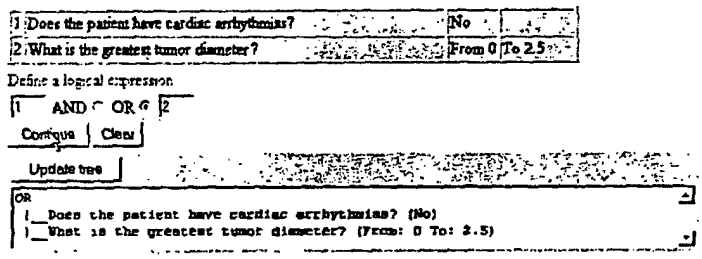


Fig. 10. Combining questions into a logical expression.

We show the screens for adding yes/no and multiple-choice questions in Figure 6; the screen for numeric questions is similar. The user can enter a new question for the current test, along with a set of allowed answers. If the question is also related to other tests, the user has to mark them in the lower box. The *"Deleting questions"* screen is for removing old questions.

Eligibility conditions: The mechanism for entering eligibility criteria consists of four screens (Figure 3). The start screen allows the user to initialize a new clinical trial and view the criteria for old trials. If the user needs to modify a clinical trial, the agent first displays the test-selection screen (Figure 8). The user then chooses related tests and question types, and clicks *"Continue"* to get the question list.

The next screen (Figure 9) allows the user to select specific questions and mark the answers that make a patient eligible. For a multiple-choice question, the user may specify several eligibility options; for example, a patient may be eligible if her cancer stage is II or III. For a numeric question, the user has to specify a range of values; for instance, a patient may be eligible if her age is between 0 and 45 years. If the user clicks *"Simple questions,"* the agent generates a conjunction of the

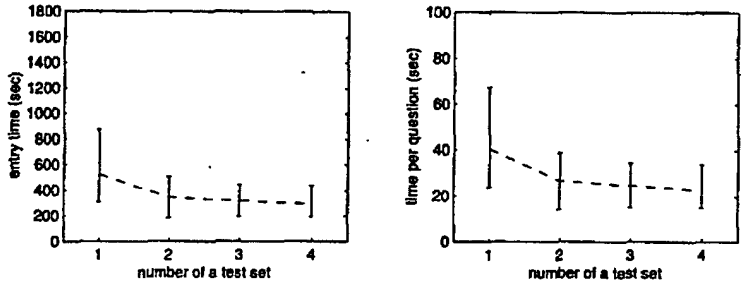


Fig. 11. Entry time for test sets (left) and the mean time per question for each set (right). We plot the average time (dashed lines) and the time of the fastest and slowest users (vertical bars).

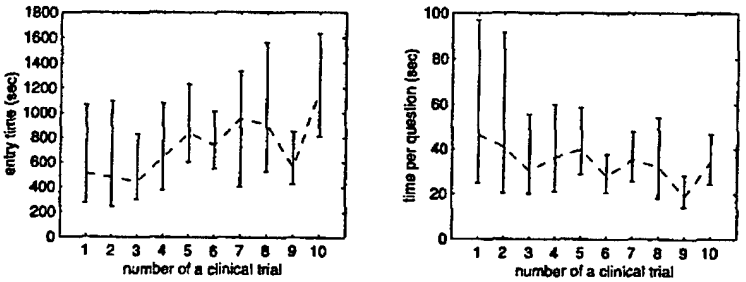


Fig. 12. Entry time for eligibility criteria. We show the average time for each clinical trial and the time per question (dashed lines), along with the performance of the fastest and slowest users (vertical bars).

selected criteria. If the eligibility conditions involve a more complex expression, the user has to click *"Combined question"* and then use the screen for composing logical expressions (Figure 10).

Entry time: We have run experiments with sixteen novice users, who had no prior experience with the interface. First, every user has entered four sets of medical tests; each set has included three tests and ten questions. Then, each user has added eligibility expressions for ten clinical trials used at the Moffitt Cancer Center; the number of questions in an eligibility expression has varied from ten to thirty-five.

We have measured the entry time for each test set and each eligibility expression. In Figure 11, we show the mean time for every test set and the time per question for the same sets. All users have entered the test sets in the same order, from 1 to 4; since they had no prior experience, their performance has improved during the experiment. In Figure 12, we give similar graphs for the entry of eligibility expressions.

The experiments have shown that novices can efficiently use the interface; they quickly learn its full functionality, and their learning curve flattens after about an hour. The average time per question is 31 seconds for the entry of medical tests and 37 seconds for eligibility criteria, which means that a user can enter all 150 cancer trials used at Moffitt in about two weeks.

V. Concluding Remarks

We have developed knowledge-acquisition tools for an agent that automatically assigns cancer patients to clinical trials. We have described the representation of eligibility criteria and a web-based interface for adding new trials. The experiments have shown that a user can enter a new trial in fifteen to thirty minutes. Novices can use the interface without prior instructions, and they reach their full speed after about an hour. Although cancer research at Moffitt has provided the motivation for this work, the agent is not limited to cancer, and we can use it for trials related to other diseases.

Acknowledgments: This work has been partially supported by the Breast Cancer Research Program of the U.S. Army Medical Research and Materiel Command under contract DAMD17-00-1-0244, and by the Moffitt Cancer Center.

References

[1] Sanjucta Bhanja, Lynn M. Fletcher, Lawrence O. Hall, Dmitry B. Goldgof, and Jeffrey P. Krischer. A qualitative expert system for clinical trial assignment. In *Proceedings of the Eleventh International Florida Artificial Intelligence Research Society Conference*, pages 84–88, 1998.

[2] Jim Blythe, Jihie Kim, Surya Ramachandran, and Yolanda Gil. An integrated environment for knowledge acquisition. In *Proceedings of the International Conference on Intelligent User Interfaces*, pages 13–20, 2001.

[3] Jacques Bouaud, Briggite Séroussi, Éric-Charles Antoine, Mary Gozy, David Khayat, and Jean-François Boisvieux. Hypertextual navigation operationalizing generic clinical practice guidelines for patient-specific therapeutic decisions. *Journal of the American Medical Informatics Association*, 5(suppl.):488–492, 1998.

[4] Jacques Bouaud, Briggite Séroussi, Éric-Charles Antoine, Laurent Zelek, and Marc Spielmann. Reusing ONCODOC, a guideline-based decision support system, across institutions: A successful experiment in sharing medical knowledge. In *Proceedings of the American Medical Informatics Association Annual Symposium*, volume 7, 2000.

[5] Bruce G. Buchanan and Edward H. Shortliffe. *Rule-Based Expert Systems: The* MYCIN *Experiments of the Stanford Heuristic Programming Project.* Addison-Wesley, Reading, MS, 1984.

[6] Henrik Eriksson. Specification and generation of custom-tailored knowledge-acquisition tools. In *Proceedings of the Thirteenth International Joint Conference on Artificial Intelligence*, volume 1, pages 510–518, 1993.

[7] Yolanda Gil and Marcelo Tallis. A script-based approach to modifying knowledge bases. In *Proceedings of the Fourteenth National Conference on Artificial Intelligence*, pages 377–383, 1997.

[8] Peter D. Johnson and Mark A. Musen. Development of a guideline authoring tool with PROTÉGÉ-II based on the DILEMMA generic protocol and guideline model. Technical Report SMI-96-0620, Stanford Medical Informatics Group, Stanford School of Medicine, 1996.

[9] Jihie Kim and Yolanda Gil. Acquiring problem-solving knowledge from end users: Putting interdependency models to the test. In *Proceedings of the Seventeenth National Conference on Artificial Intelligence*, pages 223–229, 2000.

[10] Jihie Kim and Yolanda Gil. User studies of an interdependency-based interface for acquiring problem-solving knowledge. In *Proceedings of the International Conference on Intelligent User Interfaces*, pages 165–168, 2000.

[11] Princeton K. Kokku, Lawrence O. Hall, Dmitry B. Goldgof, Eugene Fink, and Jeffrey P. Krischer. A cost-effective agent for clinical trial assignment. In *Proceedings of the* IEEE *International Conference on Systems, Man, and Cybernetics*, 2002.

[12] Cyrus Kotwall, Leo J. Mahoney, Robert E. Myers, and Linda Decoste. Reasons for non-entry in randomized clinical trials for breast cancer: A single institutional study. *Journal of Surgical Oncology*, 50:125–129, 1992.

[13] Sandra Marcus and John P. McDermott. SALT: A knowledge acquisition language for propose and revise systems. *Artificial Intelligence*, 39(1):1–37, 1989.

[14] Mark A. Musen. *Automated Generation of Model-Based Knowledge Acquisition Tools.* Morgan Kaufmann, San Mateo, CA, 1989.

[15] Mark A. Musen, David M. Combs, Edward H. Shortliffe, and Lawrence M. Fagan. OPAL: Toward the computer-aided design of oncology advice systems. In Perry L. Miller, editor, *Selected Topics in Medical Artificial Intelligence*, pages 166–180. Springer-Verlag, New York, NY, 1988.

[16] Mark A. Musen, Ray W. Fergerson, William E. Grosso, Natalya F. Noy, Monica Crubezy, and John H. Gennari. Component-based support for building knowledge-acquisition systems. In *Proceedings of the Conference on Intelligent Information Processing of the International Federation for Information Processing World Computer Congress*, 2000.

[17] Mark A. Musen, Samson W. Tu, Amar K. Das, and Yuval Shahar. EON: A component-based approach to automation of protocol-directed therapy. *Journal of the American Medical Informatics Association*, 3(6):367–388, 1996.

[18] Savvas Nikiforou. Selection of clinical trials: Knowledge representation and acquisition. Master's thesis, Department of Computer Science and Engineering, University of South Florida, 2002.

[19] Lucila Ohno-Machado, Eduardo Parra, Suzanne B. Henry, Samson W. Tu, and Mark A. Musen. AIDS^2: A decision-support tool for decreasing physicians' uncertainty regarding patient eligibility for HIV treatment protocols. In *Proceedings of the Seventeenth Annual Symposium on Computer Applications in Medical Care*, pages 429–433, 1993.

[20] Angel R. Puerta, John W. Egar, Samson W. Tu, and Mark A. Musen. A multiple-method knowledge-acquisition shell for the automatic generation of knowledge-acquisition tools. *Knowledge Acquisition*, 4(2):171–196, 1992.

[21] Thomas E. Rothenfluh, John H. Gennari, Henrik Eriksson, Angel R. Puerta, Samson W. Tu, and Mark A. Musen. Reusable ontologies, knowledge-acquisition tools, and performance systems: PROTÉGÉ-II solutions to Sisyphus-2. *International Journal of Human-Computer Studies*, 44(3–4):303–332, 1996.

[22] Briggite Séroussi, Jacques Bouaud, and Éric-Charles Antoine. Enhancing clinical practice guideline compliance by involving physicians in the decision process. In Werner Horn, Yuval Shahar, Greger Lindberg, Steen Andreassen, and Jeremy C. Wyatt, editors, *Artificial Intelligence in Medicine*, pages 76–85. Springer-Verlag, Berlin, Germany, 1999.

[23] Briggite Séroussi, Jacques Bouaud, and Éric-Charles Antoine. Users' evaluation of ONCODOC, a breast cancer therapeutic guideline delivered at the point of care. *Journal of the American Medical Informatics Association*, 6(5):384–389, 1999.

[24] Briggite Séroussi, Jacques Bouaud, and Éric-Charles Antoine. ONCODOC: A successful experiment of computer-supported guideline development and implementation in the treatment of breast cancer. *Artificial Intelligence in Medicine*, 22(1):43–64, 2001.

[25] Edward H. Shortliffe. MYCIN: *A Rule-Based Computer Program for Advising Physicians Regarding Antimicrobial Therapy Selection.* PhD thesis, Computer Science Department, Stanford University, 1974.

[26] Edward H. Shortliffe, Randall Davis, Stanton G. Axline, Bruce G. Buchanan, Cordell C. Green, and Stanley Cohen. Computer-based consultations in clinical therapeutics: Explanation and rule acquisition capabilities of the MYCIN system. *Computers and Biomedical Research*, 8:303–320, 1975.

[27] Marcelo Tallis. A script-based approach to modifying knowledge-based systems. In *Proceedings of the Tenth Innovative Applications of Artificial Intelligence Conference*, pages 1183–1195, 1998.

[28] Marcelo Tallis and Yolanda Gil. Designing scripts to guide users in modifying knowledge-based systems. In *Proceedings of the Sixteenth National Conference on Artificial Intelligence*, pages 242–249, 1999.

[29] Marcelo Tallis, Jihie Kim, and Yolanda Gil. User studies of knowledge acquisition tools: Methodology and lessons learned. In *Proceedings of the Twelfth Workshop on Knowledge Acquisition, Modeling and Management*, 1999.

[30] Salim Yusuf, Peter Held, K. K. Teo, and Elizabeth R. Toretsky. Selection of patients for randomized controlled trials: Implications of wide or narrow eligibility criteria. *Statistics in Medicine*, 9:73–86, 1990.

Using Probabilistic Methods to Optimize Data Entry in Accrual of Patients to Clinical Trials

Bhavesh D. Goswami, Lawrence O. Hall, Dmitry B. Goldgof,
Eugene Fink, and Jeffrey P. Krischer

bgoswami@csee.usf.edu, hall@csee.usf.edu, goldgof@csee.usf.edu,
e.fink@cs.cmu.edu, jpkrischer@moffitt.usf.edu

Computer Science and Engineering, University of South Florida, Tampa, FL 33620

Abstract

A clinical trial is a study conducted on a group of patients to evaluate a new treatment procedure. Usually, clinicians manually select patients for a clinical trial; the choice of eligible patients is a labor-intensive process, and clinicians are often unable to identify sufficient number of patients, which delays the evaluation of new treatments. We have developed a web-based system that helps clinicians to determine the eligibility of patients for multiple clinical trials. It uses probabilistic techniques that minimize the amount of manual data entry, by ordering the related data-entry steps. We describe the developed system and give the results of applying it to retrospective data of breast cancer patients at the Moffitt Cancer Center.

1. Introduction

A clinical trial is an experimental evaluation of a new medical procedure. When medical researchers conduct a trial, they specify a list of criteria that determines a patient's eligibility for this trial, and use these criteria to select potential participants among available patients. The selection of patients has traditionally been a manual procedure, and studies have shown that clinicians can miss up to 60% of eligible patients, which often delays the evaluation of new treatments [4, 11].

Computer scientists have developed several artificial-intelligence systems to address this problem. In particular, Musen *et al.* built a rule-based system, called EON, which selected AIDS patients for clinical trials [5]. Ohno-Machado *et al.* developed the AIDS^2 system, which also assigned AIDS patients to clinical trials [7]. Bouaud *et al.* created a decision-tree system, called ONCODOC, which helped to select patients for cancer trials [9, 10]. Papaconstantinou *et al.* built a Bayesian system for assigning patients to breast-cancer trials [8].

The developed probabilistic systems used Bayesian networks, and they inherited the usual drawbacks of Bayesian systems, including complex structure, difficulty of adding new trials, and significant running time. On the other hand, the decision-tree system could check a patient's eligibility for only one trial, and it did not scale to the use of multiple trials. To address this problem, we developed an analytical rule-based system, which efficiently processed multiple clinical trials [1–3]; however, it was unable to estimate the probably of a patient's eligibility for available trials in the absence of complete information. We have then combined the rule-based system with probabilistic techniques, described in this paper.

The developed system consists of knowledge-entry tools [6] and a patient-selection mechanism [3], as shown in Figure 1. The user accesses the system through the web-based

interface, which allows retrieving old patient data and adding new patients. For each new patient, the system presents a list of related questions, and then uses the answers to select matching clinical trials. After each answer, it estimates the probability of the patient's eligibility for each trial, and re-orders the remaining questions to minimize the expected amount of data entry. If the system decides that a patient is ineligible for a trial, it shows the conditions that make the patient ineligible. Furthermore, the system uses the new answers to augment its probabilistic knowledge and to revise the probability estimates.

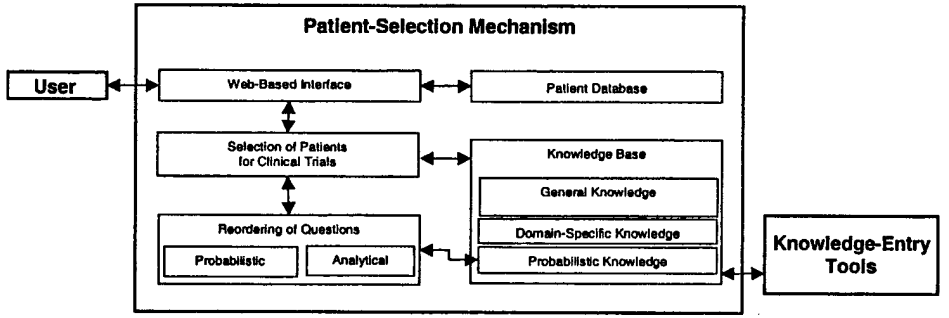


Figure 1: Architecture

2. Ordering of questions

We first explain the use of probabilities for reordering questions to reduce the amount of data entry. The underlying idea is to determine a patient's ineligibility as soon as possible. If a patient is ineligible, the information that is most likely to show her ineligibility should be obtained first. For each question, the system keeps track of how many times it was asked in the past, and how many times a patient was determined ineligible based on this question. It first asks the question that has the highest probability of showing that the patient is ineligible.

These probabilities are also used to estimate the eligibility probability of a patient for a given trial. We assume that all questions have independent probabilities. Although this assumption is not guaranteed to be true, in practice most questions are either completely dependent on each other or completely independent. For example, if a patient's cancer stage is 0 or 1, then the patient has no lymph nodes with cancer cells. Thus, the questions "Does the patient have positive lymph nodes?" and "What is the cancer stage?" are dependent. We can account for such situations by including implication rules into the knowledge base; for example, the system includes the rule "If cancer stage 0 or 1, then the patient has no positive lymph nodes."

We use the Bayes rule to compute eligibility probabilities. We can think of eligibility decisions as a classification problem with two classes, "eligible" and "ineligible." The attributes are the questions, and their values are "favorable" and "unfavorable" for eligibility. For each question and each clinical trial, we can determine the probability that the answer to this question is favorable for the trial. Thus, we have probabilities for the occurrence of each attribute value. To use the Bayes rule, we also need the probabilities of the occurrence of the classification types "Eligible" and "Ineligible." To obtain these probabilities, the system records how many patients were tested for each clinical trial and how many of them were eligible.

For example, suppose that we have a trial T with questions Q_1, Q_2 and Q_3, and that we have tested 100 patients, and found 40 of them eligible. Question Q_1 has been asked 90 times and disqualified patients 10 times, Q_2 has been asked 80 times and disqualified 5 patients, and Q_3

has been asked 70 times and disqualified 15 patients. Then, the probability that a patient is eligible for trial T is $P(T_E) = 40/100 = 0.4$. The probability that question Q_1 is answered favorably is $P(Q_1) = 80/90 = 0.89$; similarly $P(Q_2) = 75/80 = 0.94$, and $P(Q_3) = 55/70 = 0.79$.

Now suppose that we have answers to questions Q_1 and Q_2 for some patient, and both answers are favorable. According to the Bayes rule, the eligibility probability is

$$P(T_E \mid Q_1, Q_2) = \frac{P(T_E)\,P(Q_1, Q_2 \mid T_E)}{P(Q_1, Q_2)},$$

where $P(Q_1, Q_2 \mid T_E)$ is the probability that answers to Q_1, Q_2 are favorable given that the patient is eligible. If a patient is eligible, then all the questions are answered favorably, which means that $P(Q_1, Q_2 \mid T_E) = 1$. Furthermore, we have assumed that all questions are independent, which implies that $P(Q_1, Q_2) = P(Q_1)\,P(Q_2)$, and therefore

$$P(T_E \mid Q_1, Q_2) = \frac{P(T_E)}{P(Q_1)\,P(Q_2)} = \frac{0.4}{0.89 \cdot 0.94} = 0.48.$$

If the system collects more answers that satisfy the eligibility criteria, the eligibility probability becomes larger. On the other hand, if some answer does not satisfy the eligibility criteria, the system immediately concludes that the patient is ineligible. In general, if the system has collected n favorable answers and no unfavorable answers, the eligibility probability is

$$P(T_E \mid Q_1, Q_2, \ldots, Q_n) = \frac{P(T_E)}{P(Q_1)\,P(Q_2) \ldots P(Q_n)}.$$

3. Experiments

We have tested the developed technique on the retrospective data of six clinical trials and ninety patients at the Moffitt Cancer Center. We have used ten-fold cross validation; that is, we trained the system on the data of eighty-one patients, and then tested it on the other nine patients. We have repeated this test ten times, using different sets of nine test patients. In Table 1(a), we show the number of questions asked by the developed system in each test ("with probs."), and compare it with the number of questions asked by an earlier version of the system, which did not use probabilities ("without probs."). The use of probabilities has reduced the number of questions by 13%, and the t-test has shown that this difference is statistically significant with 99.99% confidence.

Table 1(a): Selection of all matching trials for each patient

Test number	Mean number of questions: With probs.	Without probs.	Difference	Percentage difference
1	16.7	20.8	4.1	20%
2	15.2	17.0	1.8	11%
3	15.8	17.6	1.8	10%
4	15.8	18.3	2.5	14%
5	13.8	16.7	2.9	17%
6	15.6	17.8	2.2	12%
7	15.8	18.3	2.5	13%
8	15.5	16.8	1.3	8%
9	16.5	18.5	2.0	11%
10	15.8	19.2	3.4	17%
Mean	15.7	18.2	2.4	13%

Table 1(b): Selection of one matching trial for each patient

Test number	Mean number of questions: With probs.	Without probs.	Difference	Percentage difference
1	20.7	28.7	8.0	28%
2	29.0	34.3	5.3	16%
3	31.7	24.3	-7.4	-30%
4	26.3	33.0	6.7	20%
5	22.3	25.0	2.7	11%
6	18.7	31.7	13.0	41%
7	25.7	33.0	7.3	22%
8	22.7	36.7	14.0	38%
9	19.3	22.7	3.4	15%
10	17.3	24.3	7.0	29%
Mean	23.4	29.4	6.0	20%

Sometimes, clinicians may need to find only one matching trial for a patient, rather than checking the patient's eligibility for all available trials. Thus, we have also experimented with using the system to select one matching trial for each patient. In this experiment, the system uses the probabilistic data to identify the most likely matching trial, and chooses questions relevant to this trial. It continues asking questions until it finds one matching trial or determines that the patient is ineligible for all trials. In Table 1(b), we give the results of this experiment ("with probs."), and compare them with a similar experiment without probabilistic reasoning ("without probs."). The use of probabilities has led to 20% reduction in the number of questions, and the t-test has shown that this difference is statistically significant with 95% confidence. The probabilistic system has given better results than the system without probabilities in nine out of ten cases. It has given worse results for test 3, because one of the patients in this test turned out ineligible for clinical trials that initially had a high eligibility probability.

4. Concluding remarks

We have developed a system that helps clinicians to select patients for clinical trials; it reduces the related manual work and helps to avoid human errors, thus increasing the number of selected patients. The system includes a probabilistic mechanism for ordering of the related questions, which helps to minimize the amount of data entry. The web-based interface allows a remote access to the system, and it can potentially enable physicians across the country to access a central repository of clinical trials.

5. References

[1] Eugene Fink, Lawrence O. Hall, Dmitry B. Goldgof, Bhavesh D. Goswami, Matthew Boonstra, and Jeffrey P. Krischer. *Experiments on the automated selection of patients for clinical trials.* In Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics, pages 4541–4545, 2003.

[2] Eugene Fink, Princeton K. Kokku, Savvas Nikiforou, Lawrence O. Hall, Dmitry B. Goldgof, and Jeffrey P. Krischer. Selection of patients for clinical trials: An interactive web-based system. Artificial Intelligence in Medicine, to appear.

[3] Princeton K. Kokku, Lawrence O. Hall, Dmitry B. Goldgof, Eugene Fink, and Jeffrey P. Krischer. *A cost-effective agent for clinical trial assignment.* In Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics, 2002

[4] Cyrus Kotwall, Leo J. Mahoney, Robert E. Myers, and Linda Decoste. *Reasons for non-entry in randomized clinical trials for breast cancer: A single institutional study.* Journal of Surgical Oncology, 50:125–129, 1992.

[5] Mark A. Musen, Samson W. Tu, Amar K. Das, and Yuval Shahar. *EON: A component based approach to automation of protocol-directed therapy.* Journal of the American Medical Informatics Association, 3(6):367–388, 1996.

[6] Savvas Nikiforou. *Selection of clinical trials: Knowledge representation and acquisition.* Master's thesis, Department of Computer Science and Engineering, University of South Florida, 2002.

[7] Lucila Ohno-Machado, Eduardo Parra, Suzanne B. Henry, Samson W. Tu, and Mark A. Musen. *AIDS^2: A decision-support tool for decreasing physicians' uncertainty regarding patient eligibility for HIV treatment protocols.* In Proceedings of the Seventeenth Annual Symposium on Computer Applications in Medical Care, pages 429–433, 1993.

[8] Constantinos Papaconstantinou, Georgios Theocharous, and Sridhar Mahadevan. *An expert system for assigning patients into clinical trials based on Bayesian networks.* Journal of Medical Systems, 22(3):189–202, 1998.

[9] Brigitte Séroussi, Jacques Bouaud, and Eric-Charles Antoine. *Users' evaluation of ONCODOC, a breast cancer therapeutic guideline delivered at the point of care.* Journal of the American Medical Informatics Association, 6(5):384–389, 1999.

[10] Brigitte Séroussi, Jacques Bouaud, and Eric-Charles Antoine. *ONCODOC: A successful experiment of computer-supported guideline development and implementation in the treatment of breast cancer.* Artificial Intelligence in Medicine, 22(1):43–64, 2001.

[11] Samson W. Tu, Carol A. Kemper, Nancy M. Lane, Robert W. Carlson, and Mark A. Musen. A methodology for determining patients' eligibility for clinical trials. *Journal of Methods of Information in Medicine*, 32(4):317–325, 1993.

Session FA3: Telemedicine