Reprinted from the JAMA® Journal of the American Medical Association
July 3, 1987 Volume 258
Copyright 1987, American Medical Association

DXplain

An Evolving Diagnostic Decision-Support System

G. Octo Barnett, MD; James J. Cimino, MD; Jon A. Hupp, MD; Edward P. Hoffer, MD

DXplain is an evolving computer-based diagnostic decision-support system designed for use by the physician who has no computer expertise. DXplain accepts a list of clinical manifestations and then proposes diagnostic hypotheses. The program explains and justifies its interpretations and provides access to a knowledge base concerning the differential diagnosis of the signs and symptoms. DXplain was developed with the support and cooperation of the American Medical Association. The system is distributed to the medical community through AMA/NET—a nationwide computer communications network sponsored by the American Medical Association—and through the Massachusetts General Hospital Continuing Education Network. A key element in the distribution of DXplain is the planned collaboration with its physician-users whose comments, criticisms, and suggestions will play an important role in modifying and enhancing the knowledge base.

(*JAMA* 1987;258:67-74)

SHORTLIFFE[1] defines a computer-based medical decision-support system as a computer program designed to help health professionals make clinical decisions. Under this broad definition, there are many decision-support systems in use today. They include such applications as clinical laboratory systems, pharmacy systems, and radiology reporting systems. For the most part, these are "passive" information management tools. That is, they make information more readily available to the physician but they do not provide "active" decision support by applying medical knowledge to a specific patient's data, nor do they recommend a specific conclusion or course of action.

See also pp 61 and 86.

There has been less success in developing active systems. One strategy used in such systems applies predefined rules to the patient's data and alerts providers to conditions that might require action. The most successful system for inpatient care is the HELP system.[2] In ambulatory medicine, two examples of systems that provide patient-specific recommendations for ambulatory care are . CARE[3] and COSTAR.[4] There are also a few examples of "expert" systems that guide the physician in treating certain complex, but well-defined, conditions, eg, the ONCOCIN system.[5] These systems contain in-depth knowledge about a well-defined subset of medical knowledge and require considerable and detailed information about the patient's clinical status to provide the computer-generated consultations.

Computer-aided diagnosis in general medicine requires a much broader level of decision support. There have been a number of provocative research efforts in computer-aided diagnosis,[6-9] but these have been research prototypes whose contributions were primarily methodological. With the exception of the work by deDombal et al,[10] these programs were not made available to the practicing physician. There are a number of reasons to explain the limited impact of information technology in the diagnostic aspects of medical decision making.[1,11,12] Major problems have included difficulty in accessing and interacting with the systems, a limited area of application (eg, one of the most successful computer-aided diagnosis programs is limited to the differential diagnosis of the acute abdomen[10]), a

From Harvard Medical School (Dr Barnett); the Laboratory of Computer Science, Massachusetts General Hospital (Drs Barnett, Cimino, Hupp, and Hoffer); and Boston University School of Medicine (Dr Hoffer), Boston. Dr Hupp is now a resident in the Department of Surgery, Massachusetts General Hospital, Boston.

Reprint requests to Laboratory of Computer Science, Massachusetts General Hospital, Boston, MA 02114 (Dr Barnett).

```
DXplain Function: [ADD]

Enter SEX of Patient (M, F, ? or <return> to ignore): [M]

Enter AGE of Patient (in Years (0-120) or <return> to ignore): [66]

Is the patient's condition: ACUTE (few hours), SUBACUTE (few days),
    or CHRONIC (longer) (A, S, or C): [A]

Add Term: [GALLOP RHYTHM]

Add Term: [HEMOPTYSIS]

Add Term: [PALPITATION]

Add Term: [TACHYPNEA]

Add Term: []
DXplain Function: [SHOW]

Common Diseases:
1  -   *  EMBOLISM, PULMONARY
2  -      PAPILLARY MUSCLE, DYSFUNCTION
4  -   *  ANEMIA, POSTHEMORRHAGIC, ACUTE
5  -      HEART FAILURE, CONGESTIVE
7  -      MITRAL VALVE, INCOMPETENCY
10 -   *  PNEUMONIA, PNEUMOCOCCAL
12 -   *  PNEUMONIA, ASPIRATION
14 -      POLYCYTHEMIA, SECONDARY
15 -   *  HYPOVOLEMIC SHOCK
16 -   *  TRANSFUSION REACTION, CARDIOVASCULAR

Rare & Very Rare Diseases:
3  -   *  THROMBOSIS, PULMONARY ARTERY
6  -   *  LEGIONNAIRE DISEASE
8  -   *  LUNG, COLLAPSE, ACUTE MASSIVE
9  -   *  PNEUMONIA, KLEBSIELLA AND OTHER GRAM NEGATIVE BACILLI
11 -   *  PLAGUE, PNEUMONIC
13 -   *  PNEUMONIA, TULAREMIC
<More>  []
```

Fig 1.—Typical DXplain session. User interacts with DXplain through selection of one of series of commands from menu. For purposes of illustration, user's entries are enclosed in brackets. Empty bracket means that user pressed only "enter." User begins by using "add" command to enter clinical features. Program first asks few basic questions about patient's sex, age, and duration of condition (user is free to skip these questions). Next, user enters medical terms regarding patient's history, symptoms, physical findings, and laboratory abnormalities. In example, user enters "acute onset" of "gallop rhythm," "hemoptysis," "palpitation," and "tachypnea" in "66-" year-old "male." Once user has entered several terms, "show" command is chosen to generate and display list of diseases that might explain some or all findings. List is organized to present three different aspects of knowledge base: (1) diseases are ranked in order of relative support for each condition (rank appears in left-most column), (2) diseases are grouped into common and less common diseases (this enables user to focus on ordinary conditions first, if desired, leaving less frequent diseases for later attention), and (3) relative need for possible immediate management of each diagnosis is depicted by asterisk. Note that last line in this example asks user "More." This indicates that there are still more diseases on list, although they are ranked lower than first set. At this point, user may examine remainder of current disease list, use "add" command to include more information about patient, use other commands that will provide more information about diseases and why DXplain is (or is not) considering a specific disease, or enter new case.

limited ability of the systems to explain and justify the interpretations, and less than acceptable quality of the interpretations. Some investigators believe that there are major intellectual and technical problems that must be solved before there can be a truly reliable consulting program. In particular, it is claimed that computer programs must be able to incorporate pathophysiological knowledge and causal relationship based on pathophysiological reasoning.[12]

We accept the validity of these concerns, but take a more optimistic position concerning the potential for using currently available knowledge and technology to assist the physician in the information needs of daily practice. A recent study suggests that the information needs of physicians in office practice are not being met by printed sources such as textbooks and journal articles.[13] Of the practice-related questions that were identified, 25% involved the diagnosis of symptoms, physical findings, or syndromes. We believe that a significant portion of the practical needs of clinical medicine can be met now by providing ready access to a computer-stored knowledge base of diagnoses and their associated signs and symptoms.

We believe it is possible to develop a computer program that uses an extensive knowledge base and relatively simple computational models to provide significant diagnostic problem-solving assistance to the practicing physician. The program should allow the computer-naive user to enter a set of signs and symptoms using a typical medical vocabulary. The program would then generate a list of hypotheses that deserve consideration, comment on the diagnostic relevance of each sign and symptom, and suggest specific additional data elements that might clarify the differential diagnosis currently favored by the computer model.[9,14-16]

This article describes DXplain, an evolving computer-based diagnostic decision-support system designed for use by the physician who has no computer expertise.[14] DXplain accepts a list of clinical manifestations and then proposes diagnostic hypotheses. DXplain explains and justifies its interpretations and provides easy access to a comprehensive knowledge base concerning the differential diagnosis of the set of signs and symptoms.

DXplain does not attempt to make a single diagnosis to mimic the behavior or replace the judgment of the expert clinician. DXplain has a less ambitious, but perhaps more attainable goal: to suggest a list of diagnoses that should be considered given a particular set of signs and symptoms. Because DXplain is in a continuing state of development, we have labeled it an *evolving* diagnostic decision-support system. Our plan is that DXplain will continue to improve through collaboration with its physician-users whose comments, criticisms, and suggestions will play an important role in modifying and enhancing the knowledge base and the algorithms used in the computer-generated interpretations.

SYSTEM CAPABILITIES

We believe that the following six criteria are important for a differential diagnosis decision-support system intended for use in routine clinical practice. The system should (1) be easy to use by physicians who have little or no computer background; (2) be based on comprehensive medical content; (3) provide correct and accurate interpretations; (4) justify its interpretations; (5) be convenient to access from the physician's office, hospital, or home; and (6) evolve and improve as a result of user criticism and analysis of user sessions.

```
-----  FIGURE 2-A -----

Add Term: [ARYTHMA]
Don't understand 'ARYTHMA', how about:

1  ARRHYTHMIA

Choose KEYWORD (0 for none): [1]

-------------------------------------------------------

----- FIGURE 2-B -----

Add Term: [FEVER]
FEVER has more specific TERMS.
Would you like to see them?   [YES]
1  -- FEVER, RECURRENT
2  -- FEVER, PEL-EBSTEIN
3  -- FEVER, LOW GRADE
4  -- FEVER, HIGH GRADE

Choose TERM (<RETURN> to keep 'FEVER'):  [3]

-------------------------------------------------------

----- FIGURE 2-C -----

Add Term: [RLQ ABDOMINAL PAIN]
1  -- ABDOMINAL PAIN, RIGHT LOWER QUADRANT

Choose TERM (0 for none): [1]

-------------------------------------------------------

-----FIGURE 2-D -----

Add Term: [RENAL]
17 Terms contain 'KIDNEY'.
Would you like to see these TERMS? [YES]

-------------------------------------------------------

----- FIGURE 2-E -----

Add Term: [BLOODY NOSE]
1  -- EPISTAXIS
2  -- NOSE DISCHARGE, BLOODY

Choose TERM (0 for none): [1]

-------------------------------------------------------
```

Fig 2.—Examples of DXplain user interface features. (A) suggests correction for misspelled word; (B) provides access to more specific forms of user's term; (C) expands abbreviations; (D) identifies synonyms at word level; (E) identifies synonyms at term level.

Easy to Use

It has been the general experience that few practicing physicians will use computer programs requiring extensive training or knowledge of computer technology. DXplain demands little knowledge of computer technology and requires only the use of the vocabulary common in medical practice. On-line help is available for an explanation of the system commands.

As with any decision-support system, DXplain uses a controlled vocabulary for communicating patient information to the system. Computer technology has not yet progressed to the stage in which a computer program can recognize the free-form narrative text that a physician might use to describe clinical manifestations in a medical record or with a colleague. The predefined medical vocabulary used by DXplain consists of more than 4700 terms based on the clinical and basic laboratory data that might be collected in an ambulatory practice or in the emergency department.

DXplain has several features that assist the user in selecting the desired terms from the controlled vocabulary. For instance, there is extensive synonym and abbreviation terminology; also, the system can recognize and correct many misspellings. DXplain can recognize terms that are "close" (eg, "congestive heart failure" and "heart failure"), as well as synonyms for terms at both the single-word level (eg, "kidney" and "renal" are equivalent in many different terms) and at the full-phrase level (eg, "anisocoria" is equivalent to "pupillary inequality," "factor VIII deficiency" is considered equivalent to "hemophilia"; and "blood glucose elevated" is considered equivalent to "hyperglycemia"). In addition, many abbreviations are recognized (eg, "ESR" and "CHF") to speed data entry.

The DXplain vocabulary structure is hierarchical, which allows flexibility in specifying the precise level of detail of the clinical manifestation. Thus, DXplain groups more specific terms (such as "lower abdominal pain") under a less specific term (such as "abdominal pain"). This hierarchy is important for both the user interface and the algorithm used for interpretation. When a user enters a specific term, such as "right lower quadrant pain," all of the appropriate less-specific terms, such as "lower abdominal pain" and "abdominal pain," are assumed by the system, but related terms at the same level of specificity, such as "left lower quadrant pain," are not assumed.

A typical user interaction with DXplain is illustrated in Fig 1. Examples of how the system helps the user select the appropriate terms during input are illustrated in Fig 2.

A second important factor relating to ease of use is the rapidity with which the user can enter the clinical manifestations and extract the desired information and interpretation. The system is designed to be largely self-explanatory. There is no need to read manuals: the beginning user can take advantage of menu selection; the experienced user can make use of an abbreviated command language to speed the interaction. The requirement for rapid time response was a critical design factor that influenced decisions made about the organization of the knowledge base in the computer. It requires about two minutes to complete the dial-in sequence to log on to AMA/NET and to connect to the computer located at Massachusetts General Hospital. The entry of the seven terms illustrated in Fig 1 required about one minute. Evaluation of the clinical data by the computer program and presentation of the interpretation

```
DXplain Function: [WHAT]

Disease name:  [PULMONARY EMBOLISM]

This Disease iS # 1 on the disease list.

'EMBOLISM, PULMONARY' is a Common disease.

You have entered the following terms which support this disease:
'MIDDLE AGE', 'GALLOP RHYTHM', 'HEMOPTYSIS', 'PALPITATION',
'TACHYPNEA', 'ACUTE'

Would also expect to find the following terms:
'ANXIETY', 'CALF PAIN', 'CONGESTIVE HEART FAILURE', 'COUGH',
'CYANOSIS', 'FEVER', 'HYPOTENSION', 'JUGULAR VENOUS DISTENTION',
'RALES', 'SHOCK' .................
'HYPERVENTILATION','JUGULAR VENOUS BLOOD PRESSURE INCREASE',
'PHLEBITIS', 'PULMONARY HYPERTENSION', 'PULSE PRESSURE DECREASE',
...................
'TACHYCARDIA', 'RESPIRATORY ALKALOSIS', 'ONSET SUDDEN'

------------------------------------------------------------

Disease name:  [TESTIS, CARCINOMA, EMBRYONAL]

This disease was not initially selected for scoring because of a low
number of matching terms.

Adding this disease to your disease list.
Now # 64 on the disease list.

'TESTIS, CARCINOMA, EMBRYONAL' is a Common disease.

The following terms support this disease: 'MALE'

Would also expect to find the following terms:
'ABDOMINAL PAIN, LOWER', 'TESTICULAR SWELLING', 'TESTICULAR MASS'

The following important term(s) are not explained by this disease:
'GALLOP RHYTHM', 'HEMOPTYSIS', 'TACHYPNEA'

------------------------------------------------------------

Disease name: [COPD]

This disease was excluded from the list because of the following
Term(s):   'ACUTE'
```

Fig 3.—Examining DXplain's reasoning. This example is continuation of example in Fig 1. Using "what" command (to ask "what about . . . ?"), user examines reasons why particular disease ("pulmonary embolism") is included or excluded from consideration. Program responds with list of supporting evidence, followed by list of features that commonly appear in disease (only a few are included in this illustration). This enables user to quickly compare program's knowledge with characteristics of patient, without requiring exhaustive entry of all clinical findings. Next, user asks, "What about embryonal carcinoma of testis?" DXplain replies that there was little evidence supporting this diagnosis, but obligingly scores it, places it on list, and tells user which entered terms support diagnosis, which other terms would also support diagnosis, and which terms are not explained by diagnosis. Finally, user asks, "What about chronic obstructive pulmonary disease?" (note use of abbreviation), and is given reason that DXplain excluded this disease from consideration.

and list of diagnostic hypotheses takes from 10 to 20 s.

Accurate and Comprehensive Medical Content

The development of a comprehensive knowledge base in medicine, whether printed or computer based, is a large effort. The merit of any decision-support system depends to a significant extent on the quality of the knowledge base used in formulating its recommendations. One aspect of any knowledge base, whether printed or computer based, is the inability of the author to certify that the entire knowledge base is totally accurate and comprehensive.

There are no methods to extract automatically the relevant information from the published literature. In fact, the published literature is not as useful as desired since it often does not provide the quantitative relationships between clinical manifestations and diseases that are required for a decision-support system. For example, one frequently finds statements like "symptom X is 'occasionally found' in the disease" or "it is 'not uncommon' to have symptom Y" or "the 'great majority' of the patients with the disease will demonstrate symptom Z."

To provide the necessary knowledge base for DXplain, we begin with the computer-based version of *Current Medical Information and Terminology* (*CMIT*), which is published and supported by the American Medical Association.[17] *Current Medical Information and Terminology* contains summaries of information concerning the etiology, signs and symptoms, laboratory findings, and disease course for more than 3000 distinct diseases, disorders, and conditions.

To transform this knowledge base into one suitable for a decision-support system, it was necessary to carry out a considerable amount of manipulation and additional content acquisition. For the disease list for DXplain, we combined some diagnoses from *CMIT* to form more inclusive definitions, while others were subdivided into more specific forms of the disease (eg, separate stages of a disease, forms of the disease specific to certain demographic groups, and forms attributed to different etiologic agents). Additional diseases were added when deficiencies in *CMIT* were noted.

As the first step in specifying the controlled vocabulary of terms to represent the clinical features of each disease, terms were extracted from the text of *CMIT*, yielding a list of some 6000 words and phrases. An extensive review of this list was undertaken to lend coherence to the vocabulary. Synonyms were merged where appropriate and similar terms were linked together in a hierarchical manner to express their relationships (usually in terms of such qualifiers as severity, duration, or anatomic location). The outcome was a term directory consisting of 4000 descriptors. As disease descriptions were compiled, 700 additional concepts were added to the directory.

We used a three-step process to generate the disease-term relationships of the knowledge base. First, a list was compiled from the *CMIT* data base and from expert judgment of all the terms that had some role in either supporting or ruling out a particular disease. Next, medical texts were reviewed to determine the frequency of each clinical finding in the disease. Where necessary, in circumstances of recent discoveries or obscure conditions, literature searches were conducted to supplement standard references. Finally, an estimate of the potential for the presence or absence of

```
DXplain Function: [QUESTION]

DYSPNEA  (Y, N, U, W, Q OR ?) [W]
'DYSPNEA' is usually present in:  'EMBOLISM, PULMONARY'.

DYSPNEA  (Y, N, U, W, Q OR ?)  [N]   noted.

HEART MURMUR, SYSTOLIC  (Y, N, U, W, Q OR ?) [W]
'HEART MURMUR, SYSTOLIC' is usually present
          in: 'PAPILLARY MUSCLE,DYSFUNCTION'.

HEART MURMUR, SYSTOLIC (Y, N, U, W, Q OR ?) [Y]  added.

CONGESTIVE HEART FAILURE  (Y, N, U, W, Q OR ?)  [U]

HEART SOUND, FOURTH  (Y, N, U, W, Q OR ?) [Q]

DXplain Function: [SHOW]

Scoring diseases...

Common Diseases:
1  -   *  PAPILLARY MUSCLE, DYSFUNCTION
3  -      MITRAL VALVE, INCOMPETENCY
5  -   *  ANGINA PECTORIS
.
.
.
19 -  **  EMBOLISM, PULMONARY

DXplain Function: [WHAT]

Disease name: EMBOLISM, PULMONARY
This Disease in # 19 on the disease list

'EMBOLISM, PULMONARY' is a Common disease.

You have entered the following terms which support this disease:
'MIDDLE AGE', 'GALLOP RHYTHM', 'HEART MURMUR', 'HEMOPTYSIS',
'PALPITATION',  'TACHYPNEA', 'ACUTE'

The following term(s) make this disease less likely:
'No DYSPNEA'

.... find the following terms: ..(not included in figure)..
```

Fig 4.—Using DXplain interrogative mode. User continues example from Figs 1 and 3 by allowing program to ask questions about case (using "question" command). Program queries about specific clinical features and user responds with one of several options: W, to find out why program is asking the question; N, to indicate that finding is absent; Y, to indicate that finding is present; U, to indicate that answer is unknown; H, to get explanation of choices (not shown here); or Q, to quit interrogative mode. In this example, user has replied "no" to a question that would support pulmonary embolism and "yes" to a question that would support papillary muscle dysfunction. As a result, when diseases are rescored, there is predictable change in ordering of diseases. When asked, "What about pulmonary embolism?" program replies with usual information but now includes clinical feature that makes diagnosis less likely.

each term to evoke or refute the diagnosis was made.

This content development was an iterative process that was done by the system developers in collaboration with 13 physicians representing several medical disciplines. Because of the number of individuals involved, extensive discussions were required to reach a consensus on content and assignment of appropriate weighting factors for the relationship of terms to diseases. Once guidelines were established, authors were trained in filling out work sheets for each disease. Each work sheet included all terms that had been previously assigned, a prescribed set of demographic terms (age, sex, and duration of symptoms), and a generous number of blank lines for adding new terms. After the work sheets were completed, one of us (J.J.C.) reviewed them to ensure consistency of content, terminology, and assignment of evoking potentials.

The result of this work is DXplain's knowledge base consisting of descriptions of approximately 2000 diseases, about 4700 terms (signs, symptoms, and more), and some 65 000 relationships among them. On average, each disease description contains relationships to 35 terms. Each relationship identifies the frequency of the manifestation among patients with the disease and the strength with which a manifestation evokes a diagnosis for consideration.

From a practical viewpoint, it is impossible to verify completely all the elements in a knowledge base that is the size and complexity of DXplain. We believe that there must be a continuing effort to detect and correct incomplete or inaccurate disease descriptions. In addition, the knowledge base can never be static, but must continue to evolve as new knowledge is gained and new evidence is identified about each disease.

One of the major advantages of having the knowledge base reside on a single central computer is that the knowledge base can be updated easily and as often as necessary. The addition of a new term, disease, or synonym, the modification of the descriptor-disease relationship, or the enhancement of the program can be done quickly. Because these changes are immediately available to all users, the system can be dynamically responsive to user suggestions concerning deficiencies and improvements.

Provide Correct and Accurate Interpretations

In designing a differential diagnosis decision-support system, we assumed that the most useful assistance that can be provided to the practicing physician is to suggest the reasonable diagnoses that should be considered, given any particular set of signs and symptoms. We believe that in many complex diagnostic problems (and many not so complex), a major reason for not making the correct diagnosis is the failure of the physician to generate a differential diagnosis list that is comprehensive enough to include the correct diagnosis, yet sufficiently focused to include only the more likely diagnostic possibilities.

The goal in DXplain is to assist the physician in two phases of the process of differential diagnosis: (1) to remind the physician of diseases that should be considered as possible candidates to explain the patient's condition and (2) to provide information about specific diseases that might be eliminated from consideration. We explicitly reject the objective of determining the "correct" diagnosis. We reason that the clinician will always have a more complete picture of the patient than will be entered into the computer program and that a more realistic goal for a decision-support system is to bring to the user's attention the plausible explanations for a patient's signs and symptoms. The

user may then consider which diseases are appropriate to the case by applying common sense, clinical experience, and full knowledge about the patient and, possibly, by further interacting with the program.

DXplain is not intended to aid the specialist working in his/her specialty area. Thus, the cardiologist dealing with a patient with a complex murmur is unlikely to find DXplain very helpful; but the cardiologist dealing with a patient with abdominal pain may obtain useful assistance. The most obvious way in which DXplain can help is in suggesting obscure or rare diseases that may be rarely seen by most physicians. DXplain also may be helpful by suggesting diseases that present in atypical ways.

DXplain uses a relatively straightforward algorithm to select its list of plausible diagnoses. This algorithm has been described in a previous publication.[14] The DXplain selection rule takes advantage of conditional probabilities and a scoring system similar to a Bayesian computation. Models such as this have been frequently used in other diagnostic decision-support projects, the most influential being the ranking algorithm used in INTERNIST/QMR.[7,14] Our development of DXplain was strongly influenced by the experiences and the limitations of the original efforts of Miller and colleagues with INTERNIST, and their later development of QMR.

The knowledge base for DXplain contains more than 65 000 relationships between diseases and patient descriptors (or clinical manifestations). The format for these relationships and the data structures are similar to the format and data structures used in the INTERNIST/QMR system[7] except that in DXplain terms cannot only *support* a given diagnosis, but can also *contradict* other diagnoses. The relationships in DXplain are concerned with three different elements: (1) term importance, (2) term frequency, and (3) term-evoking power.

Term importance is used to express how significant the particular term is in indicating the presence of disease. A high term importance is given to findings that can be identified with high reliability or are rarely found in healthy persons and, therefore, should be explained by some disease within the differential diagnoses.

Term frequency is used to identify how often a particular term is expected to occur in a specific disease. The possible values for term frequency can be one of seven different states ranging from "always" to "never."

Term-evoking power is used to identify how strongly a particular term supports the possibility that a specific disease might be present. This term is related to the concept of predictive value positive (the predictive value of a positive test—this is the probability of a disease being present given the presence of a certain finding). The possible values for term-evoking power can be one of eight different states ranging from "certainly supported" to "weakly supported" and from "weakly contradicted" to "strongly contradicted."

DXplain uses the numerical values of these different relationships to derive the list of the diagnoses that should be considered (Fig 1). DXplain selects diseases to be included on this list on the basis of how well the manifestations of a particular disease match the descriptors that have been entered by the user. The diseases are presented to the user in two lists: "common diseases" and "rare diseases"; in addition, a serious disease (one that may require relatively immediate action) is so indicated on the list by an asterisk.

The user can request DXplain to explain why any specific diagnosis was included (Fig 3). DXplain will present the clinical findings entered by the user that support the selection of that disease, the clinical findings that would not be expected in that disease, and additional clinical findings that would be expected if that disease was present. In this way, DXplain assists the user in understanding the logic used by the program and facilitates pattern matching by the user in comparing DXplain's disease description with his/her knowledge about the patient. The intent is to present sufficient information so that the user can always use his/her own clinical judgment as to the appropriateness of DXplain's interpretation.

The user can ask DXplain to consider a specific diagnosis that was not included on the initial list (Fig 3). DXplain will then present the same analysis, as described in the previous paragraph, for the diagnosis under consideration and, furthermore, will include this diagnosis in any later interpretation after additional findings are entered.

The user can change DXplain into an interrogative mode wherein the system will question the user about the presence or absence of significant findings that have the potential for clarifying DXplain's current differential diagnoses (Fig 4). This mode is particularly useful in helping the user select the clinical manifestations that are important without forcing them to enter a large number of less relevant findings. At any time in this mode, the user can interrupt DXplain to ask "Why?" ie, to ask DXplain to justify why this particular clinical manifestation is important. DXplain will respond by displaying the name of the disease that is being considered at that point in the interaction and the reason the particular finding might be important in confirming the presence of that disease.

DXplain's ability to explain and justify are key elements of the system. It is critical that this system not be perceived as a magic black box that can somehow provide the "answer" to a complex diagnostic problem. We believe that physicians will not accept DXplain as a useful diagnostic assistant unless the clinical interpretations seem reasonable and unless the system can offer explanations that are understandable and persuasive.[18]

Convenient to Access

Almost a decade ago, Shortliffe wrote: "A recurring observation as one reviews the literature of computer-based medical decision making is that essentially none of the systems has been effectively utilized outside of a research environment, even when its performance has been shown to be excellent."[8] For the most part, this observation is still true today. One important barrier is that the practicing physician cannot easily access a computer-based decision-support system from his/her office, or any other location, at any time. There are a number of systems that are available in university hospitals and that are of great importance in the host institution and have considerable value as demonstration models. However, none of these systems provide support to the practicing physician on a national scale.

DXplain is unique in that the decision-support capability is easily accessible using only a simple computer terminal (a microcomputer can also be used to make the connection), a telephone modem, and a telephone call (usually a local number in most of the major cities of the United States). DXplain is also available in a similar fashion in Canada and Japan. There is no start-up cost associated with purchase of the programs or of the knowledge base. Many physicians already have the necessary technology to access DXplain since the needs are the same as those used for on-line bibliographic search services.

DXplain was developed with the support of the American Medical Association and is designed to be distributed to the medical community through AMA/NET—a nationwide computer communications network sponsored by the American Medical Association. Physicians and other professionals can access

the system through the AMA/NET. Medical schools and teaching hospitals can access the system either through AMA/NET or through the Massachusetts General Hospital Continuing Education Network. In both cases, the cost of accessing DXplain is directly dependent on the length of time one is connected to the system.

A subscriber to AMA/NET can also access information data bases (EMPIRES clinical reference citations, Medical Procedure Coding and Nomenclature, the Associated Press Medical News Service, and more), public information services (Centers for Disease Control Information Service, National Library of Medicine/National Institutes of Health Information Services, Adverse Drug Reaction Reporting Form, and more), electronic mail, and the Massachusetts General Hospital interactive medical education courses (Hoffer et al[19]). AMA/NET also provides documentation and telephone support to its subscribers. (For information on AMA/NET, call 1-800-426-2873; for information on the Massachusetts General Hospital Continuing Education Network, call 617-726-3950.)

Evolve and Improve as a Result of User Criticism

A key element in the distribution of DXplain is an integrated electronic mail capability. At any point in the interaction a user may enter a comment or question into the computer system. This electronic mail is read at frequent intervals by the system developers at Massachusetts General Hospital and responded to as appropriate.

We view the plan for the continuing improvement and enhancement of DXplain as one of the more important aspects of its development. The major potential weakness of any diagnostic decision-support system such as DXplain is the quality and completeness of the underlying knowledge base. Evaluating a clinical decision-support system is difficult, both conceptually and in practice.[1,11] Systematic clinical trials or formal outcome studies on the impact either of computer-based knowledge bases or of medical textbooks are logistically almost impossible. DXplain has been used by physicians for over 500 hours at more than 40 different test sites in the United States, Canada, and Japan. The initial user acceptance and peer review has been favorable, although the evaluation has been largely anecdotal.

The continuing refinement of DXplain will be most fruitful if the planned collaboration between the physician-users and the developers materializes. We expect and need the participation of physicians who will challenge the system with rare diseases and with uncommon manifestations of common diseases. The continuing critical review by DXplain users of the knowledge base and interpretations of the system will provide important feedback in the iterative process of knowledge base development.

SYSTEM LIMITATIONS

DXplain has a knowledge base that covers more diseases than are discussed in most textbooks of medicine, but in some areas DXplain is incomplete, eg, there is only limited coverage of dermatologic diseases, where diagnosis often depends on the visual appearance of the lesion. At present, there is only minimal coverage of diseases from psychiatry and orthopedics. DXplain presently allows the entry of only a limited set of laboratory test findings. The original design goal of DXplain was primarily focused on ambulatory medicine; as a result, the current version of DXplain does not allow the entry of many of the complex laboratory tests that are performed only in hospitals. We are continuing to add both diseases and terms, including the most common laboratory abnormalities.

DXplain can cope in only a limited fashion with the variations in the way that a disease can present based on its evolution over time, degree of severity, and the modifications introduced by therapy. In addition, DXplain does not identify complex disease patterns caused by the presence of two or more diseases in the same patient. DXplain considers each disease and its expected manifestations as unique entities. DXplain is unable to recognize how the manifestations of one disease can be modified by the presence of a second interacting disease. However, DXplain will attempt to select and present to the user all the individual diseases that might account for the more important findings so that the physician can use clinical judgment to carry out any appropriate recognition of disease patterns.

A number of authors have emphasized the importance of an explanation capability to encourage physicians to use decision-support systems.[1,18] DXplain has an explanatory capability, but it is limited to the justification of why a particular disease should be considered (or ruled out), based solely on the likelihood of occurrence of the specific clinical manifestations in that disease. DXplain has no pathophysiological or anatomic knowledge and no ability to consider pathophysiological or anatomic reasoning.

The mere existence of data in a knowledge bank is, of itself, no guarantee of completeness and accuracy. The same professional judgment and critical appraisal are required when using DXplain as required when reading a medical textbook or discussing a patient case. In fact, since the computer program does not possess the depth of medical knowledge, the wisdom of medical experience, or the ability of a colleague to reason, it would be wise to be even more critical of the computer's interpretation. One of the weaknesses that is common to every computer-based decision-support system is a lack of "common sense" and a relative inability to consider important personal, social, family, and employment factors of the particular patient. This deficiency is well illustrated by an anecdote in the conclusion of Shortliffe's article.[1]

We believe it is critical that the physician retain the ultimate responsibility for identifying the correct diagnosis or diagnoses in any given patient. Using DXplain should be considered similar to consulting a medical textbook or journal article. DXplain should be used only as an adjunct, an information base, and a well-specified medical knowledge resource; DXplain cannot be a replacement for the clinician's knowledge and experience.

CONCLUSIONS

The potential contributions of computer-based decision-support systems are based on several factors: (1) the increasing complexity and scope of the medical knowledge base, (2) the increasing fragmentation and specialization of medical practice, (3) the increasing availability and affordability of powerful computer technology, and (4) the increasing willingness among physicians to utilize computer technology in all phases of patient care activity.

DXplain is an evolving computer system that uses an extensive knowledge base and relatively simple computational models to provide significant diagnostic problem-solving assistance to the practicing physician. The program allows the computer-naive user to enter a set of patient signs and symptoms and then generates a list of hypotheses that deserve consideration. The system also comments on the diagnostic relevance of each sign and symptom and suggests specific additional data elements that might clarify the differential diagnosis currently favored by the computer model. The advantages of a dynamic, interactive, evolving reference tool, such as DXplain, over

static, passive textbooks and journal articles are exciting.

DXplain is unique in being a decision-support system that is easily and inexpensively available to a large number of physicians through nationwide medical information networks. A key element in the distribution of DXplain is the planned collaboration with its physician-users whose comments, criticisms, and suggestions will play an important role in modifying and enhancing the knowledge base and the algorithms used in the computer-generated interpretations.

The most important evaluation of DXplain will be made by the physicians themselves as they attempt to use the system in their daily practice. As DXplain is used in a variety of practices with a variety of information needs, we will learn whether this system is capable of providing clinicians with useful information in a timely manner. The measure of success is not whether the ultimate diagnosis is listed first or fourth in the DXplain list of plausible diagnoses, but whether the physician-users perceive that the use of the system has added to their understanding of the patient's problems. Through careful analysis of user interactions, comments, level of satisfaction, and system performance, we expect to learn much about DXplain's ability to provide assistance in making clinical decisions. Based on our experience during the testing phase, we are optimistic that DXplain will be a useful educational resource and an effective assistant to the physician in daily practice.

This work was supported in part by a grant from the American Medical Association, by National Library of Medicine training grant LM 07037-3, and by a grant from Hewlett-Packard.

References

1. Shortliffe EH: Computer programs to support clinical decision making. *JAMA* 1987;258:61-66.
2. Pryor RA, Garner RM, Clayton PD, et al: The HELP system. *J Med Syst* 1983;7:87-102.
3. McDonald CJ, Hui SL, Smith DM, et al: Reminders to physicians from an introspective computer medical record. *Ann Intern Med* 1984;100:130-138.
4. Barnett GO, Winickoff R, Dorsey JL, et al: Quality assurance through automated monitoring and concurrent feedback using a computer-based medical information system. *Med Care* 1978;16:962-970.
5. Shortliffe EH: Medical expert systems: Knowledge tools for physicians. *West J Med* 1986;145:830-839.
6. Pauker SG, Gorry GA, Kassirer JP, et al: Towards the simulation of clinical cognition: Taking a present illness by computer. *Am J Med* 1976;60:981-996.
7. Miller RA, Pople HE Jr, Myers JD: INTERNIST-1, an experimental computer-based diagnostic consultant for general internal medicine. *N Engl J Med* 1982;307:468-476.
8. Shortliffe EH, Buchanan BG, Feigenbaum EA: Knowledge engineering for medical decision making: A review of computer-based clinical decision aids. *Proc IEEE* 1979;67:1207-1224.
9. Blois MS, Tuttle MS, Sherertz DD: *RECONSIDER: A Program for Generating Differential Diagnoses.* New York, Institute of Electrical and Electronics Engineers, 1981, pp 263-268.
10. DeDombal FT, Leaper DJ, Horrocks JC, et al: Human and computer-aided diagnosis of abdominal pain: Further report with emphasis on performance of clinicians. *Br Med J* 1974;1:376-380.
11. Barnett GO: The computer and clinical judgment. *N Engl J Med* 1982;307:493-494.
12. Schwartz WB, Patil RS, Szolovits P: Artificial intelligence in medicine: Where do we stand? *N Engl J Med* 1987;316:685-688.
13. Covell DG, Uman GC, Manning PR: Information needs in office practice: Are they being met? *Ann Intern Med* 1985;103:596-599.
14. Hupp JA, Cimino JJ, Hoffer EP, et al: DXplain—a computer-based diagnostic knowledge base, in Salamon R, Blum B, Jorgensen M (eds): *MEDINFO 86*. Amsterdam, Elsevier Science Publishers, 1986, pp 117-121.
15. Miller RA, Masarie FE, Myers JD: Quick medical reference (QMR) for diagnostic assistance. *MD Comput* 1986;3:34-48.
16. Miller RA, McNeil MA, Challinor SM, et al: The INTERNIST-1/QUICK MEDICAL REFERENCE project: Status report. *West J Med* 1986;145:816-822.
17. Gordon B (ed): *Current Medical Information and Terminology*, ed 4. Chicago, American Medical Association, 1971.
18. Teach RL, Shortliffe EH: An analysis of physician attitudes regarding computer-based clinical consultation systems. *Comput Biomed Res* 1981;14:542-558.
19. Hoffer EP, Barnett GO, Famiglietti KT, et al: Computer-aided instruction in medicine: 16 years of MGH experience, in Salamon R, Blum B, Jorgensen M (eds): *MEDINFO 86*. Amsterdam, Eslevier Science Publishers, 1986, pp 920-922.

Printed and Published in the United States of America