Knowledge-based computer system to aid in the histopathological diagnosis of breast disease

H Heathfield, D Bose, N Kirkham

Abstract
A knowledge-based computer system, designed to assist pathologists in the histological diagnosis of breast disease, is described. This system represents knowledge in the form of "disease profiles" and uses a novel inference model based on the mathematical technique of hypergraphs. Its design overcomes many of the limitations of existing expert system technologies when applied to breast disease. In particular, the system can quickly focus on a differential problem and thus reduce the amount of data necessary to reach a conclusion. The system was tested on two sets of samples, consisting of 14 retrospective cases and five hypothetical cases of breast disease. Its recommendations were judged "correct" by the evaluating pathologist in 15 cases. This study shows the feasibility of providing "decision support" in histopathology.

Information
Technology Research
Institute, Brighton
Polytechnic, East
Sussex BN2 4GJ
H Heathfield
Department of
Mathematics,
Brighton Polytechnic
D Bose
Histopathology
Department, Royal
Sussex County
Hospital, Brighton
N Kirkham
Correspondence to:
Dr H Heathfield
Accepted for publication
10 January 1991

Invasive breast tumours of a special histological type have a recognisably lower malignant potential than those of no special type.[1] The prognostic information provided by identifying the various histological tumour types is extremely important in determining treatment and long term management. As the therapeutic options widen to include various types of chemotherapy, surgical techniques, and radiation, stratification of women on the basis of prognosis becomes extremely important.

The abundance of histological patterns, often of a complex and variable nature, which occur in breast disease, present the pathologist with several diagnostic problems. He or she is required to be fully conversant with the diversity of possible patterns and to recognise and diagnose them accurately. Furthermore, the continuum of changes that often occurs between atypical hyperplasia and in situ carcinoma may be difficult to categorise. Problems of reproducibility are compounded by the use of different classification criteria among individual pathologists. The major problem facing the pathologist, however, is differential diagnosis. There are several acknowledged differential problems for which assistance may be required, such as fibroadenoma and phyllodes tumour, radial scar and tubular carcinoma, and lobular carcinoma in situ, and atypical lobular hyperplasia.

The introduction of the Breast Screening Programme has increased the burden on pathologists and emphasised the need for training. Problems arise because mammographic screening detects a greater proportion of special histological types, with their attendant difficulties of identification, compared with clinically palpable lesions.[2]

This paper describes a computer-based support system designed to assist pathologists in the differential diagnosis of breast disease. Its design seeks to address some of the problems of existing expert system technologies for use in histopathology.

Methods
LIMITATIONS OF EXISTING TECHNOLOGIES
Many of the early efforts to apply artificial intelligence methods to medical reasoning involved the use of IF . . . THEN . . . rules. Such rules facilitate inference using deductive logic. Rule-based systems such as MYCIN[3] are based on the assumption that expert knowledge consists of a large number of independent situation-specific rules, and that computers can simulate expert reasoning by linking these together in chains of deduction.[4 5] In many well constrained medical fields rule-based systems are highly developed—for example, the evaluation of pulmonary function tests[6] and the cytological diagnosis of breast aspirates.[7] They have also been useful in a variety of commercial tasks such as configuring computer systems.[8] Many areas of medicine are so broad and complex, however, that straightforward attempts to chain together large rule sets may be faced with major difficulties.[9]

Although rules are intended to be independent fragments of knowledge, the interaction of one rule with others may not always be consistent or predictable.[10] To achieve the desired overall behaviour from a system, the system builder is required to anticipate the manner in which each rule will interact with every other rule in the knowledge base. Furthermore, as a knowledge base is expanded and new rules added, new rules may interact with the old ones in unexpected ways that are often difficult to remedy.[10] The addition of new rules may even lead to serious degradation of a system's performance.[10–12]

In addition to the practical problems of using rules, there are limitations to their representational adequacy. Pathologists naturally describe a disease in terms of the manifestations caused by that disease, not in the reverse form of a rule—that is, "if these manifestations are present then this disease caused them."

To represent even a small amount of histological knowledge in a rule format would involve the definition of large numbers of complex rules. Such a representational scheme is inelegant and, more importantly, unnatural to pathologists. It renders validation of the knowledge base by experts difficult. Furthermore, translating histological knowledge into rules would fail to exploit the natural frame-like structure which predominates in the domain.

A second major limitation of existing expert system techniques is the extremely large amount of data need by systems to reach any conclusions.[13] Such lengthy dialogue sessions are not acceptable to pathologists working under severe time constraints. Most systems do not attempt to optimise data gathering and therefore tend to ask numerous questions which contribute little or nothing to the differential process.

KNOWLEDGE REPRESENTATION

An initial familiarisation with the subject of histological diagnosis was carried out through the study of the standard breast pathology textbooks.[14–16] Journals were also used to obtain detailed information in specific areas, but most diagnostic definitions have been taken from the *Guidelines for Pathologists* document produced by the Department of Health and Royal College of Pathologists Working Party.[17]

Throughout this study numerous pathologists have provided valuable comments and advice. Further investigation was carried out through interviews with a participating pathologist. These allowed specific diagnostic points to be clarified, and potentially problematic areas of diagnosis to be identified. Such interviews have continued regularly throughout the duration of this study, providing valuable feedback.

There several important items of diagnostic knowledge that are used by all pathologists in the identification of histological types. These include information about the possible clinical symptoms of a disease, its rate of occurrence, its macroscopic and microscopic appearances, and differential diagnosis. These pieces of information can be organised into a generic template:

Class
Disease type or class—for example, invasive ductal carcinoma, or non-invasive ductal carcinoma.

Occurrence
Information about disease prevalence—for example, ductal carcinoma of no special type comprises 53% of all invasive carcinomas.[18]

Clinical features
For example, palpable lump or nipple discharge.

Macroscopic features
Features observed macroscopically such as colour or consistency.

Microscopic features
Features observed microscopically—for

example, distended ducts, tubule formation, or mitoses.

Found with
Other types of breast disease which may occur simultaneously but in separate sites—for example, epitheliosis is commonly accompanied by the formation of cysts.[19]

Associates
Diseases from the differential diagnosis list, which may be a potential source of misdiagnosis—for example, diseases to be included in the differential diagnoses of tubular carcinoma include sclerosing adenosis and complex sclerosing lesions.[14]

Such a representational scheme is rich enough to capture pathological knowledge and of sufficient structure to allow useful organisation. A total of 30 histological disease profiles have been generated. An example is shown below:

ENTITY papilloma **BEGIN**

  CLASS benign lesions;
  OCCURRENCE a rare lesion occurring primarily in middle age;
  ASSOCIATES multiple papilloma, papillary carcinoma in situ;
  FOUND WITH sclerosing adenosis, epitheliosis;

SET clinical features **BEGIN**
  tumour location = subareolar(H);
  age group = middle age (H);
  nipple discharge = blood stained (L);
END

SET microscopic features **BEGIN**
  epithelial proliferation = yes(A), no(N);
  stromal proliferation = yes(N), no(A);
  growth type = infiltrating(N), non infiltrating(A);
  lesion type = benign(A), malignant(N);
  papillary growth = yes(A), no(N);
  double cell layer = present(A);
  cytological atypia = yes(N), no(A);
  mitosis = absent(H), infrequent(H), frequent(N);
  abnormal forms = absent(M), infrequent(N), frequent(N);
  foci of papillary growth = single(A), multiple(N);
  apocrine metaplasia = absent(H);
  lesion features = necrosis(M), haemorrhage(M);
  periductal features = fibrosis(M);
END

END

The symbolic certainty factors shown in angular brackets are subjective, non-numeric estimates of how frequently an event occurs. These translate as follows: A = always; H = high likelihood; M = medium likelihood; L = low likelihood; N = never.

The complete lack of statistical data and enormous effort which would be necessary to ascertain such data make the use of precise numerical weightings prohibitive. For example, it is commonly cited that nuclear hyperchromatism is "often present" in intraductal
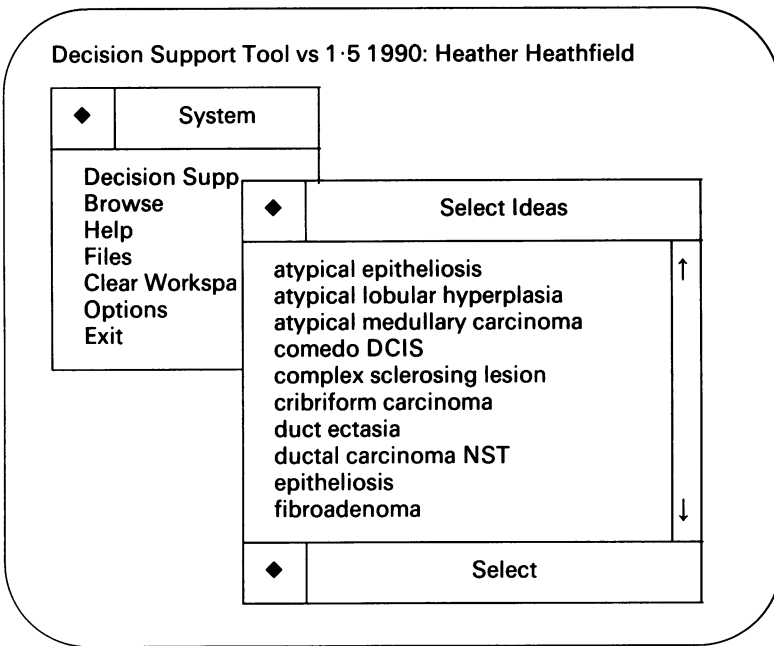
Decision Support Tool vs 1·5 1990: Heather Heathfield

◆ | System

Decision Supp
Browse
Help
Files
Clear Workspa
Options
Exit

◆ | Select Ideas

atypical epitheliosis   ↑
atypical lobular hyperplasia
atypical medullary carcinoma
comedo DCIS
complex sclerosing lesion
cribriform carcinoma
duct ectasia
ductal carcinoma NST
epitheliosis
fibroadenoma   ↓

◆ | Select

Figure 1   The "select ideas" facility

carcioma, but there is no precise quantification of this association.[17] Thus the disease description for intraductal papillary carcinoma shown above states, for example, that the proliferation is always of the type intraductal hyperplasia, and that mitotic occurrence is moderately likely to be absent or slight.

INFERENCE

The decision support system uses a novel inference technique based on the concept of hypergraphs. It is not intended to give a theoretical description of the hypergraph methodology here, however, as this is lengthy and has been described elsewhere.[20] Instead, this section will briefly present the underlying ideas embodied within this model.

The hypergraph model embodies two intuitive concepts: (1) the principle of minimum effort—that is, avoid questions that are not directly relevant to a given problem; and (2)

the idea that definitive information is preferable to doubtful information.

The principle of minimum effort states that the system should ask only those questions that are pertinent to the current differential problem. For example, in the differential problem consisting of lobular carcinoma in situ and atypical hyperplasia, cell cohesion is seen as an important differentiating factor; microcalcification is not.

The second concept states that it is preferable to base a diagnosis on those disease-feature associations that we have definite knowledge about, rather than those about which we are unsure. For example, in the differentiation of intraduct hyperplasia and intraduct carcinoma, cellular composition is a more reliable discriminator than the presence of nucleoli.

We have shown that the hypergraph model will always isolate the most clinically important differential features in any given diagnostic problem and that data gathering will be optimal.[20] The dialogue sessions produced by this model are succinct and pertinent, representing a significant reduction in the number of questions asked, when compared with other models. Furthermore, the systems line of questioning seems natural and intuitive.

THE SYSTEM

The decision support system runs on an IBM PC 386 or compatible with 640K of Ram, EGA graphics card, and Microsoft compatible mouse. It is written in C + +. The system has a user-friendly mouse driven interface.

At the start of a consultation session the user may identify the differential problem he or she wishes to explore. Pathologists are highly skilled at identifying a differential problem.[20] The "select ideas" facility shown in fig 1 supports this aspect of pathologists' problem solving skills and enables the user to concentrate on a particular diagnostic problem.

Given an initial user specified set of differential diagnoses, the system will supplement this, if necessary, with additional hypotheses that are pertinent to that particular differential
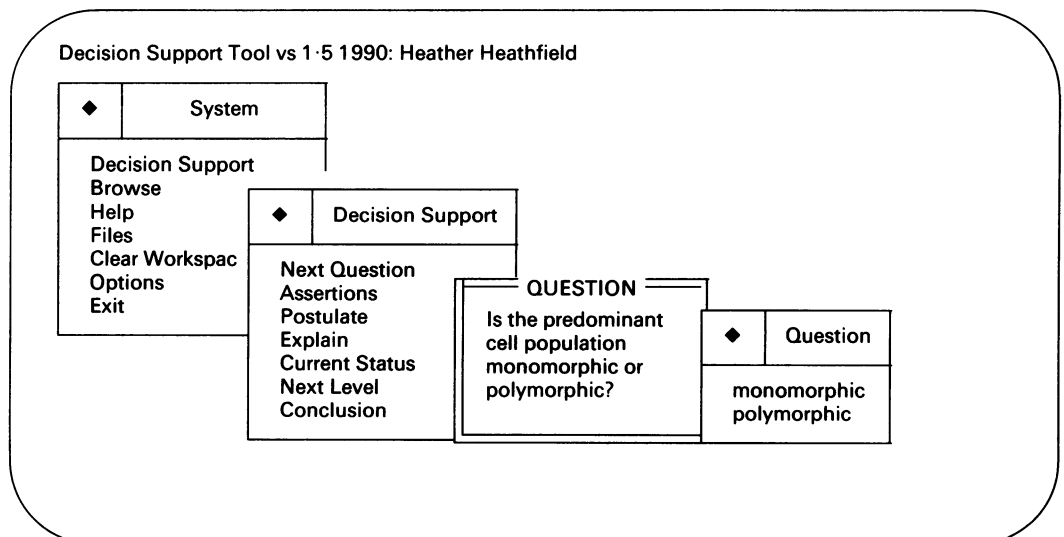
Decision Support Tool vs 1·5 1990: Heather Heathfield

◆ | System

Decision Support
Browse
Help
Files
Clear Workspac
Options
Exit

◆ | Decision Support

Next Question
Assertions
Postulate   ┌─ QUESTION ─┐
Explain   Is the predominant
Current Status   cell population   ◆ | Question
Next Level   monomorphic or
Conclusion   polymorphic?   monomorphic
                                                      polymorphic

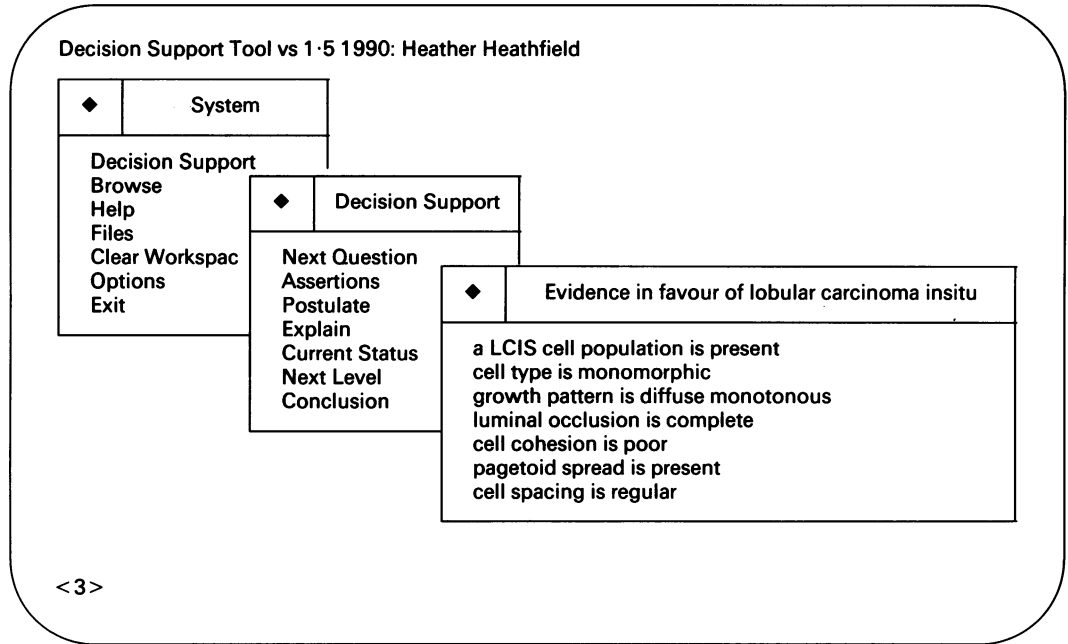*Figure 2   The format of a question*

Figure 3 Viewing evidence in favour of a disease

problem. This ensures that potential diagnostic solutions are not omitted and rare types of breast disease not overlooked. For example, if the user defines the differential problem to be between tubular carcinoma and radial scar, the system will add sclerosing adenosis and complex sclerosing lesion to the differential problem. Once the differential problem has been fully defined the system will generate a number of questions that can assist in differentiating among the alternatives. Figure 2 shows the format of a system generated question.

At any time during a consultation session, the user can view those diseases which are currently regarded as possible solutions and those which have been rejected. Evidence for these decisions can be examined. Figure 3 shows the evidence in favour of the disease lobular carcinoma in situ. A consultation session may be concluded at will by selecting the "conclude" option. This provides a summary of the current evidence in favour of a diagnosis and additional information that may assist the pathologist in deciding whether to accept a suggested diagnosis (figs 4 and 5). The system also contains a "Knowledge Base Browser".[21] This enables the user to interrogate the underlying knowledge base in various ways without actually entering into a consultation session.
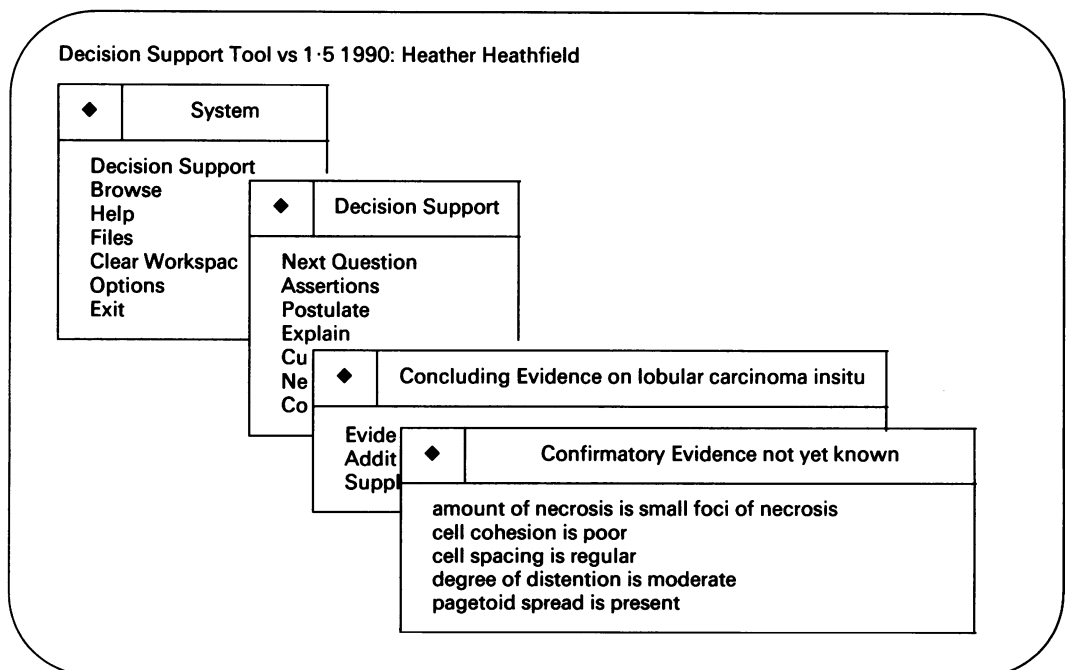

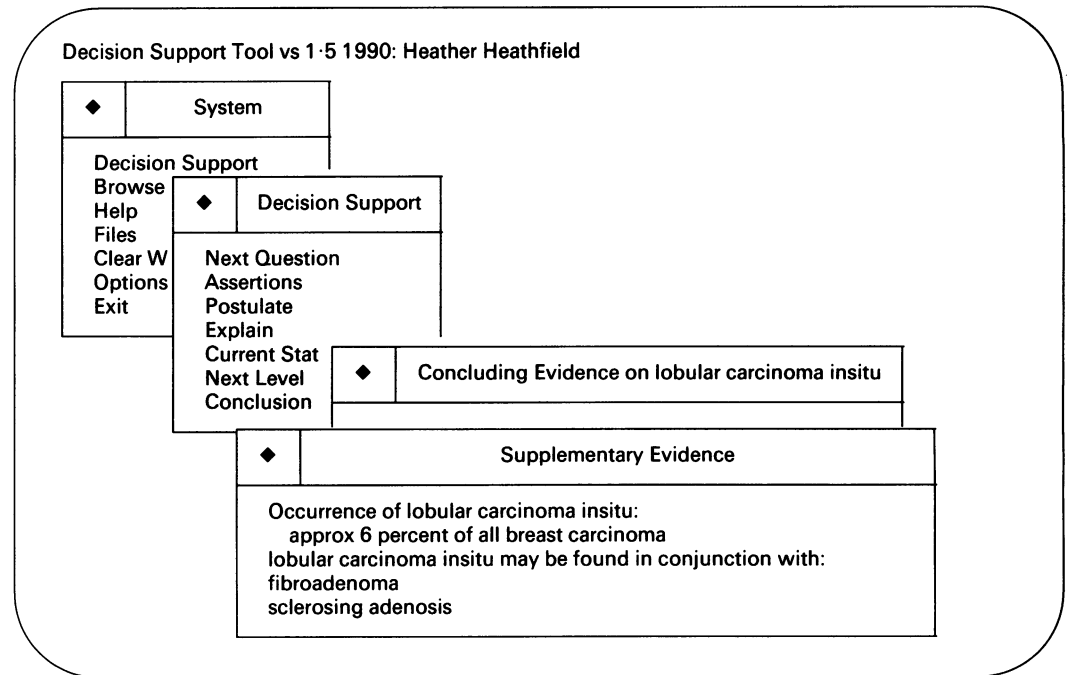*Figure 4 Viewing confirmatory evidence*

Figure 5   Viewing supplementary evidence

Results

The histology system was tested using two sets of samples: the first comprising 14 available samples of breast disease in the form of slide preparations from retrospective cases (obtained from the Royal Sussex County Hospital, Brighton in 1990); the second consisting of five hypothetical cases.

The aim of this evaluation was to determine if the histology system can assist the pathologist in making accurate and consistent diagnoses. A single pathologist evaluated the system's recommendations. The diagnoses made by the system were classified by the evaluating pathologist into one of three categories:

1 *Correct* The system's recommendation was identical with the pathologist's own diagnosis.

2 *Acceptable* The system's recommendation was different from that of the pathologist, but considered an acceptable alternative. This category also includes situations in which the system presented several alternative diagnoses (that were relevant to the given problem), one of which was the same as that given by the pathologist.

3 *Incorrect* The system's recommendation was different from that of the pathologist and not an acceptable alternative. This also includes situations in which the system gave several alternative diagnoses, one or more of which were contradictory to the differential problem as perceived by the pathologist.

Testing proceeded in the following manner. The evaluating pathologist examined the slide (or slides) of each case microscopically, identified the perceived differential problem, if possible, and then used the system to direct data gathering. When all the questions generated by the system had been answered, the pathologist compared the system's recommendations to his or her own diagnosis and classified it into one of the three categories described above.

Table 1 shows the results of using the systems in 14 cases of breast disease. The column entitled "differential problem" indicates the initial differential problem, as identified by the pathologist. The column entitled "hypothesis size" indicates the number of diseases included in the initial hypothesis after any system additions have been added to the differential problem, as defined by the pathologist.

Table 2 shows the results of five hypothetical cases of breast disease specified by the evaluating pathologist. These were included as they represent interesting and demanding diagnostic areas in which the pathologist wished to explore the system's behaviour. From a total of 19 possible correct diagnoses, the system's recommendation was judged "correct" by the evaluating pathologist in 15 cases.

The problems found in cases 3 and 6 resulted from incomplete disease specification in the knowledge base. In case 17 the system failed to differentiate between two diagnostic alternatives. All relevant data had been gathered but there was insufficient evidence to conclude in favour of one diagnosis. In such situations it is necessary to rank the diagnoses according to evidence, indicating to the pathologist which is the most likely. Possible methods of ranking final hypotheses are detailed elsewhere.[21] In case 14 the system failed to recommend the correct diagnosis because the lesion belonged to a new class of lesions, termed juvenile papillomatosis, that have not yet been included in the knowledge base.

Table 1   *System performance using 14 retrospective cases of breast disease*

Case No	Differential problem	Hypothesis size	Pathologist's diagnosis	System's recommendations	No of questions	Evaluation
1	Unknown	30	Lobular carcinoma in situ	Lobular carcinoma in situ	10	Correct
2	Unknown	30	Single papilloma	Single papilloma	8	Correct
3	Papilloma Atypical hyperplasia Sclerosing adenosis Fibrocystic disease Complex sclerosing lesion	11	Complex sclerosing lesion	Complex sclerosing lesion Sclerosing adenosis Fibrocystic disease	7	Acceptable
4	Fibroadenoma Sclerosing adenosis Complex sclerosing lesion Radial scar	6	Sclerosing adenosis	Sclerosing adenosis	6	Correct
5	Ductal carcinoma in situ Lobular carcinoma in situ	7	Ductal carcinoma in situ	Ductal carcinoma in situ	7	Correct
6	Fibroadenoma Cystic disease Sclerosing adenosis	8	Sclerosing adenosis	Sclerosing adenosis Fibroadenoma Fibrocystic disease Atypical hyperplasia	8	Acceptable
7	Duct ectasia Fibrocystic disease	3	Duct ectasia	Duct ectasia	5	Correct
8	Fibroadenoma Phyllodes tumour	2	Fibroadenoma	Fibroadenoma	3	Correct
9	Cribriform carcinoma Ductal carcinoma NST	7	Cribriform carcinoma	Cribriform carcinoma	5	Correct
10	Lobular carcinoma classic alveolar or tubulolobular form	7	Lobular carcinoma classic form	Lobular carcinoma classic form	4	Correct
11	Lobular carcinoma in situ Atypical lobular hyperplasia Ductal carcinoma in situ	12	Lobular carcinoma in situ	Lobular carcinoma in situ	5	Correct
12	Tubular carcinoma Complex sclerosing lesion	7	Ductal carcinoma NST	Ductal carcinoma NST	7	Correct
13	Lobular carcinoma Ductal carcinoma NST	7	Ductal carcinoma NST	Ductal carcinoma NST	5	Correct
14	Papilloma Papillary carcinoma in situ Papillary carcinoma Atypical ductal hyperplasia	5	Juvenle papillomatosis	Papillary carcinoma in situ Multiple papilloma	8	See text

Discussion

It is recognised that the evaluation of only one pathologist who has been associated with this study is not the ideal way in which to test the system, but these initial tests have shown the potential of the model and indicated minor areas that require attention. In particular, they have highlighted those areas of the knowledge base which need further development.

To determine if the system has fulfilled its goal—that is, to improve the accuracy and reproducibility of histological diagnosis—it is necessary to evaluate objectively its performance on a wide scale. This is a major task, requiring both laboratory and field testing,[22] that must be addressed to assess the effects of the system on routine diagnostic practice.

Much of the data requested by the system requires the user to make a subjective judgment of some histological feature, such as cell size, and thus may render system evaluation diffi-

cult. There are various ways in which this problem can be tackled, however. These include substituting quantitative morphometric data for subjective assessment where possible, and the provision of a database of example images which may be used for comparison. A natural extension of the system would be to include larger portions of textual descriptions that could be viewed by the user to clarify terms and assist in the recognition of features. It would also be useful to include reference citations to indicate the source of definitions and diagnostic criteria used in the system.

The knowledge base at present contains 30 disease profiles. These cover the main categories specified by the Breast Screening Programme. There are many more types of breast disease that have not yet been included, although it is hoped to extend the system in the future.

Table 2   *System performance using five hypothetical cases of breast disease*

Case No	Differential problem	Hypothesis size	Pathologist's diagnosis	System's recommendations	No of questions	Evaluation
15	Sclerosing adenosis Complex sclerosing lesion Radial scar	4	Sclerosing adenosis	Sclerosing adenosis	4	Correct
16	Comedo carcinoma Ductal carcinoma NST	9	Comedo carcinoma	Comedo carcinoma	5	Correct
17	Atypical lobular hyperplasia Lobular carcinoma in situ	8	Lobular carcinoma in situ	Atypical lobular hyperplasia Lobular carcinoma in situ	8	Acceptable
18	Medullary carcinoma Ductal carcinoma NST	7	Atypical medullary carcinoma	Atypical medullary carcinoma	6	Correct
19	Atypical lobular hyperplasia Lobular carcinoma in situ	7	Atypical lobular hyperplasia	Atypical lobular hyperplasia	7	Correct

The problem of pathologists using different diagnostic criteria contributes largely to the inconsistency and low reproducibility of histological diagnoses. The diagnostic guidelines produced by the Royal College of Pathologists have been circulated to all pathologists participating in the Breast Screening Programme in an attempt to mitigate this. The histology system has embodied these criteria and while its recommendations are not universally agreed by all pathologists, they do indicate which diagnoses would be made if these guidelines are adhered to. Therefore the system can help to highlight the areas where different diagnostic criteria are being used, and also educate pathologists in their use.

This study has proved the feasibility of decision support in breast histopathology. It has also successfully addressed the problems associated with existing expert system technologies. Future work will be directed towards expanding the knowledge base and testing on a wider scale.

1 Dixon JM, Page DL, Anderson TJ, *et al*. Long-term survivors after breast cancer. *Br J Surg* 1985;72:445-8.
2 Anderson TJ, Alexander F, Chetty U, *et al*. Comparative pathology of prevalent and incident cancers detected by breast screening. *Lancet* 1986;i:519-22.
3 Shortliffe EH. *Computer-based medical consultations: Mycin*. New York: Elsevier, 1976.
4 Buchanan BG, Shortliffe EH. eds. *Rule-based expert systems: the MYCIN experiments of the Stanford heuristic programming project*. Reading, MA: Addison-Wesley, 1984.
5 van Melle WJ. *System aids in constructing consultation*

programs. Ann Arbor, M: UMI Research Press, 1981.
6 Kunz JC, Fallat RJ, McClung DH, *et al*. A physiological rule-based system for interpreting pulmonary function test results. *Proc Comp Crit Pulmonary Med* 1979:375-9.
7 Heathfield HA, Kirkham N, Ellis IO, Winstanley G. Computer assisted diagnosis of fine needle aspirate of the breast. *J Clin Pathol* 1990;44:168-70.
8 McDermott J. R1: A rule-based configurer of computer systems. *Artifical Intelligence* 1982;19:39-88.
9 Szolovits P, Patil RS, Schwartz WB. Artifical intelligence in medical diagnosis. *Ann Intern Med* 1988;108:80-7.
10 Schwartz WB, Patil RS, Szolovits P. Artificial intelligence in medicine. Where do we stand? *N Engl J Med* 1987;316:385-8.
11 Szolovits P, Patil RS, Schwartz WB. Artificial intelligence in medical diagnosis. *Ann Intern Med* 1988;108:80-7.
12 Clancy WJ, Letsinger R. Neomycin: reconfiguring a rule-based expert system. *Proceedings of the 7th International Joint Conference on Artificial Intelligence, 1981*. Los Altos CA; Morgan Kaufmann; 1981:829-36.
13 Heathfield HA, Bose D, Winstanley G, Kirkham N. A computer-based decision support system for diagnostic histopathology of the breast. *Path Res Pract* (in press).
14 Page DL, Anderson TJ. *Diagnostic histopathology of the breast*. Edinburgh: Churchill Livingstone, 1987.
15 Sloane JP. *Biopsy pathology of the breast*. London: Chapman and Hall, 1985.
16 Azzopardi JG. *Problems in breast pathology*. Philadelphia: W B Saunders, 1979.
17 Department of Health and Royal College of Pathologists Working Party. *Guidelines for pathologists*. London: HMSO, 1990.
18 Fisher ER, Gregorio RM, Fisher B. The pathology of invasive breast cancer. A syllabus derived from findings of the national surgical adjuvant breast project (protocol no 4). *Cancer* 1975;36:1-85.
19 Franz VK, Pickren JW, Melcher GW, *et al*. Incidence of chronic cystic disease in so-called "normal breasts": a study based on 225 post-mortem examinations. *Cancer* 1951;4:762-83.
20 Heathfield HA. *Differential diagnosis and decision support in histopathology*. (Thesis). Information Technology Research Institute, Brighton Polytechnic, 1990.
21 Heathfield HA, Winstanley G, Kirkham N. A menu-driven knowledge base browsing tool. *J Med Inform* 1990;15:151-9.
22 Wyatt J, Spiegelhalter D. Evaluating medical expert systems: What to test and how? *J Med Inform* 1990;15:205-17.