Pergamon

Comput. Biol. Med. Vol. 27, No. 1, pp. 19–29, 1997

Printed in Great Britain
0010-4825/97 $17.00+0.00

PII: S0010-4825(96)00039-X

CONSTRUCTION OF A BAYESIAN NETWORK FOR MAMMOGRAPHIC DIAGNOSIS OF BREAST CANCER

CHARLES E. KAHN JR[a]*, LINDA M. ROBERTS[b], KATHERINE A. SHAFFER[a] and PETER HADDAWY[a]

[a] Department of Radiology, Medical College of Wisconsin, Milwaukee, WI 53226, U.S.A.

[b] Department of Electrical Engineering and Computer Science, University of Wisconsin-Milwaukee, Milwaukee, WI 53201, U.S.A.

(*Received* 13 *May* 1996; *revised* 30 *July* 1996)

Abstract—Bayesian networks use the techniques of probability theory to reason under uncertainty, and have become an important formalism for medical decision support systems. We describe the development and validation of a Bayesian network (MammoNet) to assist in mammographic diagnosis of breast cancer. MammoNet integrates five patient-history features, two physical findings, and 15 mammographic features extracted by experienced radiologists to determine the probability of malignancy. We outline the methods and issues in the system's design, implementation, and evaluation. Bayesian networks provide a potentially useful tool for mammographic decision support. © 1997 Elsevier Science Ltd

Bayesian networks Artificial intelligence Breast cancer Mammography
Computer-aided diagnosis Expert systems

INTRODUCTION

In 1995, an estimated 183,400 women in the United States were newly diagnosed with breast cancer, and 46,240 died of the disease [1]. Screening mammography effectively detects early breast cancers and can increase the likelihood of cure and long-term survival [2]. Differentiating between benign and malignant mammographic findings, however, is difficult. Only 15–30% of biopsies performed on nonpalpable but mammographically suspicious lesions prove malignant [3]. Automated classification of mammographic findings using discriminant analysis and artificial neural networks has indicated the potential usefulness of computer-aided diagnosis [4,5]. We explored the use of Bayesian networks as a diagnostic decision aid in mammography.

Bayesian networks

A Bayesian network—also called a belief network or causal probabilistic network—is a graphical representation of probabilistic information: it is a directed, acyclic graph in which nodes represent random (stochastic) variables, and links between nodes represent direct probabilistic influences between the variables [6]. In this formalism, propositions are given numerical probability values signifying the degree of belief accorded them, and the values are combined and manipulated according to the rules of probability theory. The graph is "directed" in that the connections between nodes have directionality, that is, they are "one way". The graph is "acyclic" in that it cannot contain cycles or "feedback" loops. Each node represents a variable and has two or more possible states. For example, the variable "Breast Cancer" has two states: "present" and "absent". Each state is associated with a probability value; for each node, these probability values sum to 1.

Typically, the direction of a connection between nodes indicates a causal influence or class–property relationship. For example, a link may indicate that age influences the presence of breast cancer. Indirect influences are represented by paths through the network.

* Author to whom correspondence should be addressed.

The lack of certain types of paths between two nodes indicates probabilistic independence. The precise relationship between probabilistic independence and connectedness of nodes is defined in terms of a property called *d*-separation [6].

For example, consider the network shown in Fig. 1, in which the patient's current age and her age at menarche are two possible factors contributing to the risk of breast cancer. According to the criterion of *d*-separation, the topology of the network indicates that age and age at menarche are probabilistically independent. But if we know that the patient has breast cancer, then they are no longer independent and knowing that the patient is over 50 decreases the chance that she had early menarche. This is due to the fact that each cause is sufficient to increase the probability of breast cancer. In this way, the directionality of the arrows in a belief network allows encoding of intercausal interactions. The strengths of influences are represented with conditional-probability matrices associated with the links. The conditional-probability matrix specifies the probabilities of the possible values that a node can assume given all possible combinations of values of its parents.

The primary operation performed with Bayesian networks is the computation of posterior probabilities. Efficient inference procedures exist for performing computations over the network [7–9]. They permit one to specify values of observed variables and compute the posterior probabilities of the remaining variables. For example, one can use a Bayesian network to perform a differential diagnosis by specifying the observed symptoms and computing the posterior probabilities of the various diagnoses. Because links represent conditional probabilities, propagation across links is bi-directional. One can easily see why this is the case since the conditional probability associated with a link can be reversed by simply applying Bayes' rule. The algorithms gain efficiency by exploiting the independence information encoded in the network topology.

Bayesian networks can express the relationships between diagnoses, physical findings, laboratory test results, and imaging study findings. Physicians can determine the *a priori* ("pre-test") probability of a disease, and then incorporate laboratory and imaging results to calculate the *a posteriori* ("post-test") probability [10]. Bayesian networks can be used to

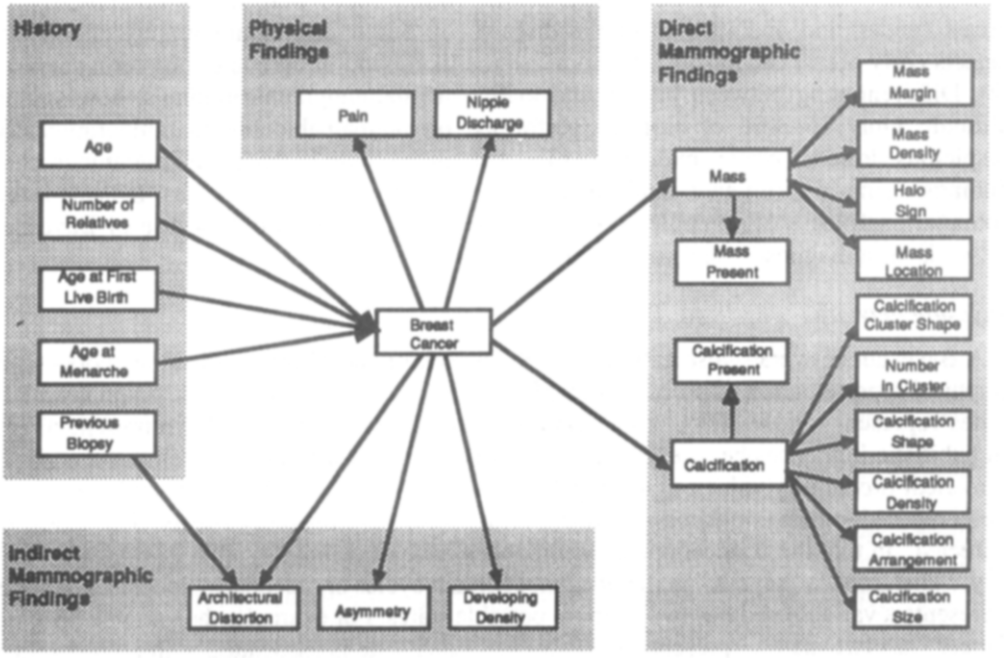

Fig. 1. Topology of MammoNet's Bayesian network model for mammographic diagnosis. Nodes, such as Breast Cancer, have two or more possible states (e.g. "present" and "absent"). Arcs between nodes indicate probabilistic influences; arrows indicate the direction of influence. After setting nodes for which information is available, the network can infer the probabilitiy values of the remaining nodes. In particular, the Breast Cancer node is evaluated to determine the probability that breast cancer is present given the available evidence.

plan diagnostic tests and therapeutic interventions [11,12]. Bayesian networks have been used to aid in interpretation of electromyographic findings [13], cytological diagnosis of fine-needle aspirates of the breast [14], and histological diagnosis of lymph-node pathology [15–17]. In radiology, Bayesian networks have been applied to the diagnosis of liver lesions on MR images [18] and to the selection of imaging procedures for patients with suspected gallbladder disease [12]. Efforts are underway to reformulate large, general medical decision support systems such as Iliad [19] and QMR [20,21] into Bayesian networks.

Problem overview

Mammography is an important tool in early detection of breast cancer. Unfortunately, many mammographic findings cannot be classified easily as malignant or benign. Successful diagnosis depends on the ability of a physician to detect mammographic abnormalities and to integrate clinical information such as risk factors and physical findings to determine the likelihood of breast cancer [22].

Epidemiological investigations have reported several risk factors that may increase a woman's chance of developing breast cancer [23]. The incidence of breast cancer increases with age, and is higher after menopause. Early menarche, late childbearing (first live birth after age 30 yr, or nulliparity), and first-degree relatives with breast cancer increase the probability of malignancy. Breast pain, nipple discharge, and skin thickening are reported by women with breast cancer, but few early-stage cancers are detected by these indicators [24]. The risk factors and physical findings alone are not sufficiently sensitive for malignancy; thus, mammography is an important screening tool.

Mammographic evidence of breast cancer can be grouped into direct and indirect findings. Direct mammographic findings include masses with or without calcifications and calcifications alone. Indirect mammographic signs include architectural distortion, asymmetry, ductal dilatation, and developing density.

Benign and malignant masses are differentiated using shape, margin, density, and the presence of the halo sign [25,26]. Round, low-density masses with smooth, sharply defined margins are more likely to be benign. Malignant masses are more likely to be high-density, spiculated, knobby, or with poorly defined margins [25]. Frequently, though, masses are classified as indeterminate, not clearly benign or malignant. Instead of spiculations, many malignant masses display poorly defined or irregular margins [25].

Calcifications can occur with or without an associated mass. The attributes of size, shape, density, distribution pattern, and number are examined to differentiate benign and malignant calcifications [27,28]. Benign calcifications are typically larger (1–4 mm in diameter), coarse, round or oval, and monomorphic (uniform in size and shape). Their distribution pattern is typically scattered or diffuse. If the calcifications are clustered, they number less than 5 per cluster. Malignant calcifications are typically smaller (<0.5 mm in diameter), linear, branching, rod-shaped, punctate, stellate, or pleomorphic (varying in size and shape). In general, the greater the number of calcifications in a cluster, the greater the likelihood of malignancy. Both benign and malignant calcifications, however, can appear tiny and clustered [28].

Almost 20% of nonpalpable cancers can present with neither mass nor calcifications, but with subtle or "indirect" signs of malignancy [29]. Architectural distortion or a developing density (an enlarging area of glandular-tissue density) may indicate cancer. Surgical interventions can confound the diagnosis: a breast biopsy can produce architectural distortion.

IMPLEMENTATION

We created a Bayesian network model of breast cancer diagnosis, called MammoNet, that incorporates five patient–history features, two physical findings, and 15 mammographic findings. The model assumes that all of the evidence pertains to one particular site identified by mammography. MammoNet infers the posterior probability of breast cancer at that site based on the available evidence.

Model structure

The nodes and their states are enumerated in Table 1. MammoNet uses the standardized terminology of the American College of Radiology's Breast Imaging Reporting and Data System (BI-RADS) lexicon [30,31]. Because some findings, such as Mass Margin, are completely dependent on the presence of a mass, these nodes attain a state "not applicable" (NA) if the preceding node has the value "none". Calcification Size has several states to indicate the size ranges of the observed calcifications; several of these ranges overlap. In practice, the mammographer selects the one range that best describes the sizes of calcifications seen.

Four of the patient-history features influence the presence of breast cancer, which in turn influences the presence of the physical findings and mammographic findings (Fig. 1). The mammographic findings are divided into direct manifestations of malignancy, such as mass or calcification, and indirect signs, such as architectural distortion. One of the patient-history factors, that of a prior biopsy, serves as a competing cause of the mammographic finding of architectural distortion.

Mammographically detectable mass and calcification are modeled as conditionally independent manifestations of malignancy. The Mass and Calcification nodes have three states: "malignant", "benign", and "none". If no mass is evident, for example, the Mass Present node is set to "no", which forces the Mass node to the state "none" and nodes such as Mass Margin to the state "not applicable" (NA). The Mass Present node allows one to express uncertainty regarding the presence of mass independently of the descriptive features. The mammographic features of a mass (Mass Margin, Mass Density, etc.)—although conditionally independent of Breast Cancer given Mass—affect the diagnosis by

Table 1. Definitions of MammoNet's nodes and their states

Category	Node	States
Diagnosis	Breast cancer	Present, absent
History	Age (yr)	20–24, 25–29, ..., 75-79
	Age at menarche (yr)	<12, 12–13, ≥14
	Age at first live birth (yr)	<20, 20–24, 25–29, ≥30
	Number of first-degree relatives with breast cancer	0, 1, 2
	Previous biopsy	Yes, no
Physical findings	Pain	Present, absent
	Nipple discharge	Present, absent
Indirect mammographic findings	Architectural distortion	Present, absent
	Asymmetry	Present, absent
	Developing density	Present, absent
Direct mammographic findings	Mass	Malignant, benign, none
	Mass present	Yes, no
	Mass margin	Spiculated, irregular, relatively well defined, NA
	Mass density	High, low, NA
	Halo sign	Present, absent, NA
	Tumor location	Upper outer, upper inner, lower outer, lower inner, retroareolar, NA
	Calcification	Malignant, benign, none
	Calcification present	Yes, no
	Calcification cluster shape	Punctate, round, linear, variable, NA
	Number of calcifications in cluster	≤5, 6–10, 11–15, 16–25, 26–50, >50, NA
	Calcification shape	Linear branching, irregular, indeterminate, round, NA
	Calcification density	1–2, 1–3, 2–3, 3–4, NA
	Calcification arrangement	Scattered, clustered, scattered & clustered, single, NA
	Calcification size (mm)	0.05–0.1, 0.05–0.2, 0.01–1, 0.01–2, 1–3, NA

their influence on the Mass node's "malignant" and "benign" states. The model treats Calcification and its related nodes in similar fashion.

Data acquisition

MammoNet's knowledge base was constructed from peer-reviewed medical literature, census data, and health statistics reports. When required probability data were unavailable or the sample size was too small, an expert mammographer provided subjective estimates of the probabilities. Conditional probabilities for architectural distortion, previous biopsy at the same site, and the halo sign were estimated as well. Values for demographic variables were derived from published epidemiological data [23,32,33]. Statistical studies published in peer-reviewed radiology journals provided most of the data for MammoNet's knowledge base, such as values of sensitivity and specificity of the physical findings and mammographic findings for breast cancer [24,25,28,29,34–36].

Inference software

We used the Bayesian Network Generation (BNG) system [37,38] to generate a problem-specific Bayesian network for each case from a knowledge base of rules representing parent–child relationships in a Bayesian network. Given a knowledge base in the form of probabilistic rules, a set of evidence E, and a query Q, the BNG system generates the structurally minimal network to compute the conditional probability of Q given the evidence E, that is, $P(Q|E)$. A network is structurally minimal if it contains only nodes relevant to the query, where relevance is determined by the network's topology. The generated network is passed to the public-domain IDEAL system [9] to perform the probabilistic computation. Inference (i.e. calculation of posterior probabilities) was performed on a SPARCstation 20 (Sun Microsystems Computer Co., Mountain View, CA). Because BNG only generates those nodes necessary for an inference problem, computation time for large models can be reduced significantly [39]. We used BNG because we expect this model to eventually become quite large.

A rule in a BNG knowledge base has the general form:

Antecedent:	$A_1,\ldots,A_n$
Consequent:	A_0
Matrix:	(conditional probabilities)

where the A_i represent nodes in the Bayesian network. For example, the rule representing the direct influence of a benign or malignant calcification on the calcification's density is shown in Fig. 2. A knowledge base of such rules can be provided with a formal semantics defining the probability distribution they represent [40,41].

Given a query Q and a set of evidence nodes, E, BNG constructs a network to compute $P(Q|E)$, such that the probability computed with the network is equal to that defined by the knowledge base semantics [38]. The key idea behind the algorithm is that since the rules in the knowledge base are structurally similar to Horn-clauses (Prolog-type rules), we can use backward-chaining to search through the rules for paths between the query and related random variables. The generated network is just the resulting search tree. The BNG system is capable of generating temporal and nontemporal Bayesian networks and of reasoning with context constraints [39,41].

An example

A 53-yr-old woman presented for mammography. She had no breast pain or nipple discharge, and no first-degree relatives with a history of breast cancer. Her mammogram showed a high-density, spiculated mass in the upper outer quadrant of the right breast. No microcalcifications were present. MammoNet generated a Bayesian network from this evidence that included only those nodes relevant to the query about breast cancer (Fig. 3), and calculated a posterior probability of breast cancer of 94.7%. Malignancy was confirmed on subsequent biopsy.

```
;P(CalcificationDensity | Calcification)
(CalcificationDensity ?x)
((Calcification ?x))
 (; p(1-2  1-3  2-3  3-4   NA | Calcification)
      0.18 0.76 0.04 0.02  0  ; malignant
      0.12 0.48 0.26 0.14  0  ; benign
      0    0    0    0     1) ; none
     (1-2  1-3  2-3  3-4   NA)
)
```

Fig. 2. Representation of conditional probabilities using BNG. The variable x represents the patient. Comments are preceded by a semicolon (;). The data indicate, for example, that the conditional probability is 0.18 that Calcification Density=1–2 given that the calcification is malignant.

PRELIMINARY EVALUATION

To test MammoNet, we encoded cases from a mammography atlas [42] and a clinical teaching file. Each case included clinical data, mammographic findings, the expert mammographer's diagnosis, and the histological diagnosis based on clinical follow-up and/or biopsy results. The atlas provided a set of 67 relatively straightforward cases, of which 24 were malignancies. The clinical teaching file provided 10 cases considered diagnostically challenging, of which only one was positive for breast cancer. By selecting cases in a wide range of difficulty, we hoped to maximize the statistical power of our test, as has been suggested by Metz [43]. Overall, 25 (32%) were positive for breast cancer; other observer-performance studies of mammographic interpretation have included a high proportion of malignancies in their test panels (e.g. [44,45]). MammoNet computed the posterior probability of breast cancer given each case's constellation of demographic, clinical, and mammographic features. Analysis using the LABROC1 program [46] yielded an estimated area under the receiver operating characteristic (ROC) curve of 0.881 ± 0.045 (Fig. 4).

At a probability threshold for breast cancer of 15% (which approximates the positive

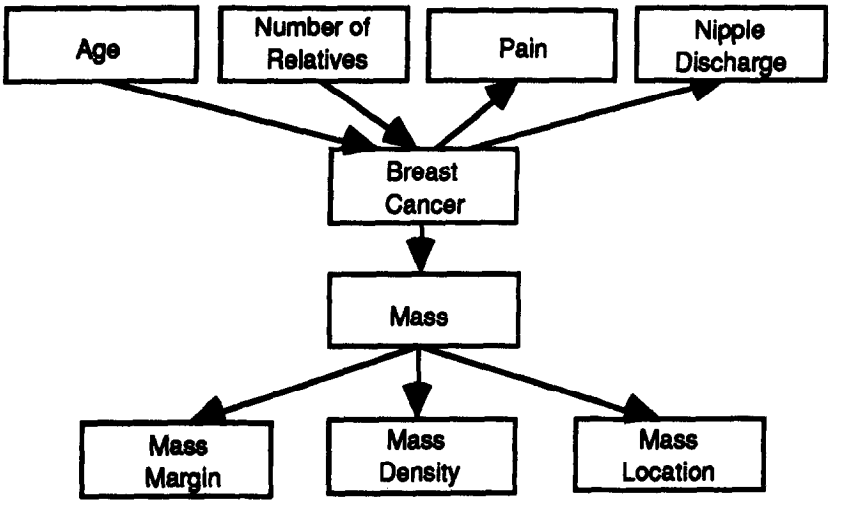

Fig. 3. Bayesian network generated by BNG for the query *BreastCancer* (Present). Only nodes relevant to query (and for which evidence is available) are included in the network.

predictive value of mammographic suspicion), MammoNet correctly identified 23 of the 25 actually positive cases (sensitivity, 92.0% ; 95% confidence interval [CI], 73.9–99.0). MammoNet's specificity at this threshold was 88.5% (95% CI, 76.5–95.7). Three benign lesions that MammoNet falsely identified as positive were considered suspicious by the mammographers and were referred for biopsy.

DISCUSSION

Automated classification using discriminant analysis and artificial neural networks has indicated the potential usefulness of computer-aided diagnosis in mammography [4,5]. A computer-assisted decision aid for mammographic interpretation included a checklist of 12 features determined to have particular diagnostic value [47]. Given a quantitative assessment of the 12 features, the decision aid estimated the probability of malignancy using weighting factors obtained from discriminant analysis. Receiver operating characteristic (ROC) analysis of this model showed a gain of about 0.05 in sensitivity or specificity when the other value remained constant at 0.85[47].

Most of the mammographic decision aids developed to date have been artificial neural networks (ANNs). An ANN is a collection of interconnected elements that can learn to recognize patterns [5,48–51]. An ANN with 14 input features achieved an area under the ROC curve (A_z) of 0.89 (vs 0.84 for attending radiologists) [5]; this ANN was trained and tested on cases from the same mammography atlas [42] as used in this study. The chief advantage of ANNs their ability to learn patterns directly from observations.

ANNs have several disadvantages. The training process, which involves incrementally adjusting the numerical weights associated with the connections between nodes, is a "hill-climbing" process; although the network may reach a locally optimal solution based on the randomized starting values of the weights, there is no guarantee that one has found the most accurate model overall. There is strong evidence that the choice of clinical cases used to

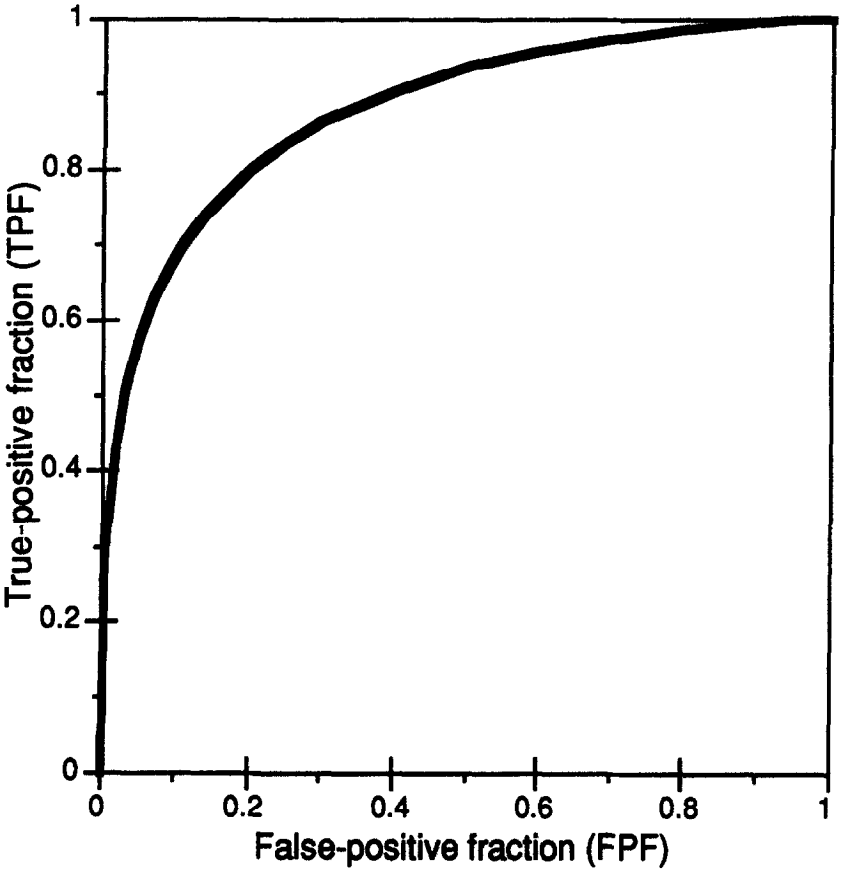

Fig. 4. Receiver operating characteristic (ROC) curve for MammoNet.

train and test computer-aided diagnostic systems can affect their performance significantly [52]. Because its tangle of numerical connection values, particularly in a multi-layered network, makes an ANN's knowledge inaccessible to human understanding, one cannot be certain what the network has "learned". ANNs cannot meaningfully explain their reasoning [53]. Although the lack of explanatory power may be acceptable for perceptual tasks, such as finding microcalcifications on a radiograph, physicians will not accept and act on a computer system's advice without knowing the basis for the system's decision [54].

Bayesian networks represent a promising technique for clinical decision support and provide a number of powerful capabilities for representing uncertain knowledge. They provide a flexible representation that allows one to specify dependence and independence of variables in a natural way through the network topology. Because dependencies are expressed qualitatively as links between nodes, one can structure the domain knowledge qualitatively before any numeric probabilities need be assigned. The graphical representation also makes explicit the structure of the domain model: a link indicates a causal relation or known association. The encoding of independencies in the network topology admits the design of efficient procedures for performing computations over the network. A further advantage of the graphical representation is the perspicuity of the resulting domain model. Because Bayesian networks represent uncertainty using standard probability, one can collect the necessary data for the domain model by drawing directly on published statistical studies.

Bayesian networks can explain their reasoning [55–57]. One system for Bayesian-network explanation, called BANTER, generates explanations in two steps: it identifies the evidence that has the most influence on the given hypothesis and then identifies the strongest and most comprehensible paths linking the influential evidence with the hypothesis [57]. To explain the current belief in the presence of breast cancer, BANTER first identifies those nodes among the specified history, physical, and mammographic findings that were most influential in producing the reported posterior probability of breast cancer, and then finds the paths along which that influence flows. Using information from MammoNet's knowledge base and supplemental information that defines the semantic relationships among the terms, BANTER can generate English-language explanations, as shown in Fig. 5.

FUTURE RESEARCH

MammoNet is undergoing preclinical testing to compare its performance to that of radiologists with varying levels of mammographic expertise. We are considering the addition of demographic features such as race and geographic location, and patient-history

```
Before presenting any evidence, the probability of Breast
   Cancer is 0.060.

With clinical data only, the probability of Breast Cancer
   is 0.042.


The following pieces of evidence are considered important:

Mass Margin (Spiculated) raises the probability of Breast
   Cancer to 0.987.  Mass Margin (Spiculated) is a feature
   of Mass (Malignant), which is caused by Breast Cancer.

Mass Density (High-Density) raises the probability of
   Breast Cancer to 0.300.  Mass Density (High-Density) is
   a feature of Mass (Malignant), which is caused by Breast
   Cancer.
```

Fig. 5. Explanation generated by the BANTER system using MammoNet's Bayesian-network model for the example case.

features such as diet, body habitus, history of hormone therapy, and previous cancers. The granularity of the model's variables could be increased by partitioning the Breast Cancer node into more than the current two states to represent the numerous types of cancer and benign conditions.

We plan to integrate MammoNet with a mammography reporting system and to evaluate its performance in our screening and problem-solving mammography populations. We are developing links between MammoNet and a database to allow collection and analysis of a large set of clinical cases. We have developed a system that generates explanations from the knowledge contained in a Bayesian network [57,58], and plan to apply this system to MammoNet. Our goal is to create a decision support tool to improve the diagnostic accuracy and cost-effectiveness of screening mammography. Such a decision aid must be reliable, integrated with a clinical database and reporting system, and able to generate explanations to the physicians who use it.

SUMMARY

Bayesian networks use the techniques of probability theory to reason under uncertainty, and have become an important formalism for medical decision support systems. We developed a Bayesian-network decision support system, called MammoNet, to assist in mammographic diagnosis of breast cancer. MammoNet integrates 15 mammographic features identified by radiologists, five patient-history features, and two physical findings to determine the probability of malignancy. MammoNet uses the standardized terminology of the American College of Radiology's Breast Imaging Reporting and Data System (BI-RADS) lexicon. The Bayesian Network Generation (BNG) system was used to generate problem-specific Bayesian networks from a knowledge base of first-order probability logic sentences. MammoNet was tested on a panel of 77 cases with known outcomes; 25 cases were positive for breast cancer. MammoNet yielded an estimated area under the receiver operating characteristic (ROC) curve of 0.881 ± 0.045. At an arbitrary probability threshold for breast cancer of 15% (which approximates the positive predictive value of mammographic suspicion), MammoNet correctly identified 23 of the 25 actually positive cases (sensitivity, 92.0%; specificity, 88.5%). Further work is underway to compare MammoNet's performance to that of radiologists. Bayesian networks provide a potentially useful tool for mammographic decision support.

Acknowledgements—This paper was presented in part at the 19th Annual Symposium on Computer Applications in Medical Care, New Orleans, November 1995. Supported by grants from the Will Ross Memorial Foundation, the Medical College of Wisconsin Cancer Center, the National Science Foundation (IRI-9509165), Sun Microsystems (Academic Equipment Grant), and the National Library of Medicine (USPHS G08 LM05705). We thank Robert A. Krieger, MS, for many helpful discussions and Charles E. Metz, PhD, for ROC analysis software.

REFERENCES

1. P.A. Wingo, T. Tong and S. Bolden, Cancer statistics, 1995, *CA*, **45**, 8 (1995).
2. S.A. Feig, Decreased cancer mortality through mammographic screening: results of clinical trials, *Radiol.*, **167**, 659 (1988).
3. F.M. Hall, J.M. Storella, D.Z. Silverstone and G. Wyshak, Nonpalpable breast lesions: recommendations for biopsy based on suspicion of carcinoma at mammography, *Radiol.*, **167**, 353 (1988).
4. D.J. Getty, R.M. Pickett, C.J. D'Orsi and J.A. Swets, Enhanced interpretation of diagnostic images, *Invest. Radiol.*, **23**, 240 (1988).
5. Y. Wu *et al.*, Artificial neural networks in mammography: application to decision making in the diagnosis of breast cancer, *Radiol.*, **187**, 81 (1993).
6. J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann, San Mateo, CA (1988).
7. S.L. Lauritzen and D.J. Spiegelhalter, Local computations with probabilities on graphical structures and their applications to expert systems, *J. R. Statist. Soc.*, **50**, 157 (1988).
8. S.K. Andersen, K.G. Olesen, F.V. Jensen and F. Jensen, HUGIN—a shell for building Bayesian belief universes for expert systems, Proceedings of the Eleventh International Joint Conference on Artificial Intelligence (AAAI Press, Menlo Park, CA, 1989), pp. 1080 (1989).
9. S. Srinivas and J. Breese, IDEAL: a software package for analysis of influence diagrams, Proceedings of the Sixth Conference on Uncertainty in Artificial Intelligence (Morgan Kaufmann, San Mateo, CA, 1990), pp. 212 (1990).
10. S. Andreassen, F.V. Jensen and K.G. Olesen, Medical expert systems based on causal probabilistic networks, *Int. J. Biomed. Comput.*, **28**, 1 (1991).

11. S. Andreassen, Planning of therapy and tests in causal probabilistic networks, *Art. Intell. Med.*, **4**, 227 (1992).
12. P. Haddawy, C.E. Kahn Jr and M. Butarbutar, A Bayesian network model for radiological diagnosis and procedure selection: work-up of suspected gallbladder disease, *Med. Phys.*, **21**, 1185 (1994).
13. S. Andreassen, M. Woldbye, B. Falcka and S.K. Andersen, MUNIN—a causal probabilistic network for interpretation of electromyographic findings, Proceedings of the International Joint Conference on Artificial Intelligence (AAAI Press, Menlo Park, CA, 1987), pp. 366 (1987).
14. P.W. Hamilton, N. Anderson, P.H. Bartels and D. Thompson, Expert system support using Bayesian belief networks in the diagnosis of fine needle aspiration biopsy specimens of the breast, *J. Clin. Path.*, **47**, 329 (1994).
15. D.E. Heckerman, E.J. Horvitz and B.N. Nathwani, Toward normative expert systems: Part I. The Pathfinder project, *Meth. Inform. Med.*, **31**, 90 (1992).
16. D.E. Heckerman and B.N. Nathwani, Toward normative expert systems: Part II. Probability-based representations for efficient knowledge acquisition and inference, *Meth. Inform. Med.*, **31**, 106 (1992).
17. D.E. Heckerman and B.N. Nathwani, An evaluation of the diagnostic accuracy of Pathfinder, *Comput. Biomed. Res.*, **25**, 56 (1992).
18. R. Tombropoulos, S. Shiffman and C. Davidson, A decision aid for diagnosis of liver lesions on MRI, C. Safran (Ed.), Proceedings of the 17th Annual Symposium of Computer Applications in Medical Care (McGraw-Hill, New York, 1993), pp. 439 (1993).
19. Y.C. Li, P.J. Haug and H.R. Warner, Automated transformation of probabilistic knowledge for a medical diagnostic system, J.P. Ozbolt (Ed.), Proceedings of the 18th Annual Symposium on Computer Applications in Medical Care (Hanley and Belfus, Philadelphia, 1994), pp. 765 (1994).
20. M.A. Shwe *et al.*, Probabilistic diagnosis using a reformulation of the INTERNIST-1/QMR knowledge base. I. The probabilistic model and inference algorithms, *Meth. Inform. Med.*, **30**, 241 (1991).
21. B. Middleton *et al.*, Probabilistic diagnosis using a reformulation of the INTERNIST-1/QMR knowledge base. II. Evaluation of diagnostic performance, *Meth. Inform. Med.*, **30**, 256 (1991).
22. S.A. Feig, Breast masses: mammographic and sonographic evaluation, *Radiol. Clin. N. Amer.*, **30**, 67 (1992).
23. G.A. Colditz *et al.*, Family history, age, and risk of breast cancer. Prospective data from the Nurses' Health Study, *J. Am. Med. Ass.*, **270**, 338 (1993).
24. R.E. Symmonds Jr and J.W. Roberts, Management of nonpalpable breast abnormalities, *Ann. Surg.*, **205**, 520 (1987).
25. E.A. Sickles, Breast masses: mammographic evaluation, *Radiol.*, **173**, 297 (1989).
26. C.A. Swann *et al.*, The halo sign and malignant breast lesions, *Am. J. Roentgen.*, **149**, 1145 (1987).
27. L.W. Bassett, Mammographic analysis of calcifications, *Radiol. Clin. N. Amer.*, **30**, 93 (1992).
28. E.A. Sickles, Breast calcifications: mammographic evaluation, *Radiol.*, **160**, 289 (1986).
29. E.A. Sickles, Mammographic features of 300 consecutive nonpalpable breast cancers, *Am. J. Roentgen.*, **146**, 661 (1986).
30. C.J. D'Orsi and D.B. Kopans, Mammographic feature analysis, *Seminars in Roentgenology*, **28**, 204 (1993).
31. *Breast Imaging Reporting and Data System (BI-RADS).* American College of Radiology, Reston, VA (1993).
32. J.L. Kelsey and M.D. Gammon, *The Epidemiology of Breast Cancer.* American Cancer Society, New York (1991).
33. M.H. Gail *et al.*, Projecting individualized probabilities of developing breast cancer for white females who are being examined annually, *J. Natl Canc. Inst.*, **81**, 1879 (1989).
34. E.A. Sickles, Nonpalpable, circumscribed, noncalcified solid breast masses: likelihood of malignancy based on lesion size and age of patient, *Radiol.*, **192**, 439 (1994).
35. R.L. Egan, M.B. McSweeney and C.W. Sewell, Intramammary calcifications without an associated mass in benign and malignant diseases, *Radiol.*, **137**, 1 (1980).
36. V.P. Jackson, K.A. Dines, L.W. Bassett, R.H. Gold and H.E. Reynolds, Diagnostic importance of the radiographic density of noncalcified breast masses: analysis of 91 lesions, *Amer. J. Roentgen.*, **157**, 25 (1991).
37. P. Haddawy, Generating Bayesian networks from probability logic knowledge bases, Uncertainty in Artificial Intelligence: Proceedings of the Tenth Conference (Morgan Kaufmann, San Mateo, CA, 1994), pp. 262 (1994).
38. P. Haddawy, L. Ngo, J. W. Helwig and R. Krieger, Clinical simulation using context-sensitive temporal probability models, R.M. Gardner (Ed.), Proceedings of the 19th Annual Symposium on Computer Applications in Medical Care (Hanley and Belfus, Philadelphia, 1995), pp. 203 (1995).
39. L. Ngo, P. Haddawy, J. Helwig and R.A. Krieger, Efficient temporal probabilistic reasoning via context-sensitive model construction, *Comput. Biol. Med.* (in press).
40. L. Ngo and P. Haddawy, Probabilistic logic programming and Bayesian networks, Algorithms, Concurrency and Knowledge (Proceedings ACSC95) (Springer-Verlag, 1995), pp. 286 (1995).
41. L. Ngo and P. Haddawy, Answering queries from context-sensitive probabilistic knowledge bases, *Theoret. Comput. Sci.*, **171**, 147 (1997).
42. L. Tabár and P.B. Dean, *Teaching Atlas of Mammography.* Thieme-Stratton, New York (1985).
43. C.E. Metz, Some practical issues of experimental design and data analysis in radiological ROC studies, *Invest. Radiol.*, **24**, 234 (1989).
44. Y. Wu, K. Doi, M.L. Giger and R.M. Nishikawa, Computerized detection of clustered microcalcifications in digital mammograms: applications of artificial neural networks, *Med. Phys.*, **19**, 555 (1992).
45. C.A. Beam, P.M. Layde and D.C. Sullivan, Variability in the interpretations of screening mammograms by US radiologists: findings from a national sample, *Arch. Int. Med.*, **156**, 209 (1995).
46. C.E. Metz, ROC methodology in radiologic imaging, *Invest. Radiol.*, **21**, 720 (1986).
47. C.J. D'Orsi *et al.*, Reading and decision aids for improved accuracy and standardization of mammographic

diagnosis, *Radiol.*, **184**, 619 (1992).
48. C.E. Floyd Jr, J.Y. Lo, A.J. Yun, D.C. Sullivan and P.J. Kornguth, Prediction of breast cancer malignancy using an artificial neural network, *Cancer*, **74**, 2944 (1994).
49. J.A. Baker, P.J. Kornguth, J.Y. Lo, M.E. Williford and C.E. Floyd Jr, Breast cancer: prediction with artificial neural network based on BI-RADS standardized lexicon, *Radiol.*, **196**, 817 (1995).
50. J.Y. Lo, J.A. Baker, P.J. Kornguth and C.E. Floyd Jr, Computer-aided diagnosis of breast cancer: artificial neural network approach for optimized merging of mammographic features, *Acad. Radiol.*, **2**, 841 (1995).
51. J.A. Baker, P.J. Kornguth, J.Y. Lo and C.E. Floyd Jr, Artificial neural network: improving the quality of breast biopsy recommendations, *Radiol.*, **198**, 131 (1996).
52. R.M. Nishikawa etal.,, Effect of case selection on the performance of computer-aided detection schemes, *Med. Phys.*, **21**, 265 (1994).
53. J. Diederich, Explanation and artificial neural networks, *Int. J. Man-Machine Stud.*, **37**, 335 (1992).
54. R.L. Teach and E.H. Shortliffe, An analysis of physician attitudes regarding computer-based clinical consultation systems, *Comput. Biomed. Res.*, **14**, 542 (1981).
55. M. Henrion and M.J. Druzdzel, Qualitative propagation and scenario-based approaches to explanation of probabilistic reasoning. In P.P. Bonissone, M. Henrion, L.N. Kanal and J.F. Lemmer (Eds), *Uncertainty in Artificial Intelligence 6* (Elsevier Science Publishers B.V., New York, 1991), pp. 17 (1991).
56. H.J. Suermondt and G.F. Cooper, An evaluation of explanations of probabilistic inference, *Comput. Biomed. Res.*, **26**, 242 (1993).
57. P. Haddawy, J. Jacobson and C.E. Kahn Jr, Generating explanations and tutorial problems from Bayesian networks, J.P. Ozbolt (Ed.), Proceedings of the 18th Annual Symposium on Computer Applications in Medical Care (Hanley and Belfus, Philadelphia, 1994), pp. 770 (1994).
58. P. Haddawy, J. Jacobson and C.E. Kahn Jr, BANTER: a Bayesian network tutoring shell, *Artif. Intell. Med.* (in press).

About the Author—CHARLES E. KAHN JR received a BA in Mathematics from the University of Wisconsin-Madison and an MD from the University of Illinois. He is Associate Professor and Director of the Section of Information and Decision Sciences in the Department of Radiology of the Medical College of Wisconsin, Milwaukee, Wisconsin.

About the Author—LINDA M. ROBERTS received a BS in Mathematics from the University of Wisconsin-Madison, an MS in Computer Science from the University of Wisconsin-Milwaukee. She is enrolled in the PhD program at the University of Wisconsin-Milwaukee.

About the Author—KATHERINE A. SHAFFER received an AB from DePauw University and an MD from the University of Michigan. She is Professor of Radiology and Director of Breast Imaging at the Department of Radiology of the Medical College of Wisconsin.

About the Author—PETER HADDAWY received a BA in Mathematics from Pomona College, and MS and PhD degrees in Computer Science from the University of Illinois at Urbana-Champaign. He is Associate Professor and Director of the Decision Systems and Artificial Intelligence Laboratory in the Department of Electrical Engineering and Computer Science of the University of Wisconsin-Milwaukee.