African Journal of Mathematics and Computer Science Research Vol. 3(10), pp. 225-231, October 2010
Available online at http://www.academicjournals.org/AJMCSR
ISSN 2006-9731 ©2010 Academic Journals

Full Length Research Paper

Exploring Bayesian networks for medical decision support in breast cancer detection

Jyotirmay Gadewadikar[1]*, Ognjen Kuljaca[1], Kwabena Agyepong[2], Erol Sarigul[3], Yufeng Zheng[4] and Ping Zhang[2]

[1]The Sensors and Automation Laboratory, Systems Research Institute, Alcorn State University, 1000 ASU Drive #360, Lorman, MS, 39096 United State.
[2]Systems Research Institute, Alcorn State University, 1000 ASU Drive #360, Lorman, MS, 39096, United State.
[3]Vision Research Lab, Systems Research Institute, Alcorn State University, 1000 ASU Drive #360, Lorman, MS, 39096, United State.
[4]Pattern Recognition and Image Analysis Lab, Systems Research Institute, Alcorn State University, 1000 ASU Drive #360, Lorman, MS, 39096, United State.

Accepted 13 September, 2010

The objective of this paper is to explore the implementation of a Bayesian Belief Network for an automated breast cancer detection support tool. It is intuitive that Bayesian networks are employed as one viable option for computer-aided detection by representing the relationships between diagnoses, physical findings, laboratory test results and imaging study findings. This work brings important entities such as Radiologists, Image Processing Scientists, Data Base Specialists and Applied Mathematicians on a common platform. A brief background concerning causal networks, probability theory and Bayesian networks is given; available computational tools and platforms are described. It is explained that, by exploiting conditional independencies entailed by influence chains, it is possible to represent a large instance in a Bayesian network using little space, and it is often possible to perform probabilistic inference among the features in an acceptable amount of time. The next steps towards realizing a Bayesian Belief Network Implementation are described. Bayesian networks have an unparallel advantage of being able to exploit the explicit structure of the domain model to derive a graphical representation for learning. The encoding of independencies in the network topology admits the design of efficient procedures for performing computations over the network. For the application of computer-aided detection in mammography, the researchers intend to design an interface between the project's Bayesian network learning algorithm and the radiologists, so that the radiologists can have interaction with the system by labeling only a small number of informative images presented by the active learning algorithm.

Key words: Bayesian networks, breast cancer, computer aided decision support.

INTRODUCTION

Early diagnosis of breast cancer through mammography screening is the most effective means of reducing the death rate caused by the breast cancer. Extensive studies are done by Duffy et al. (2002), for example, in organizing service screening in 7 Swedish counties, covering approximately 33% of the population of Sweden, resulted in a 40 to 45% reduction in breast carcinoma mortality among women actually screened. In the work by Gajdos et al. (2002), the cancer characteristics are classified into groups based on mammographic presentation, that is, mass, calcification, mass and architectural distortion.

Masses defined as space occupying lesions seen in more than one different projection may be characterized by their shape and margin properties.

*Corresponding author. E-mail: jyo@alcorn.edu. Tel: 601-877-2405. Fax: 601-877-6219.

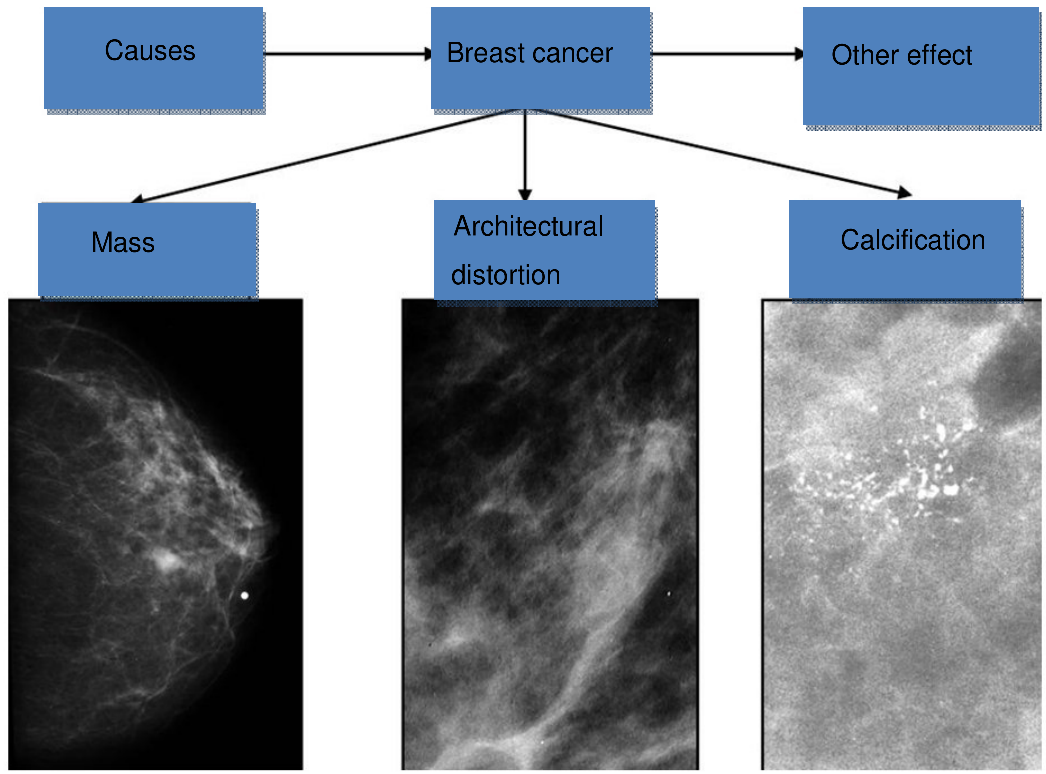

Figure 1. A Simplified cause and effect diagram.

With regard to shape, they may be round, oval, lobular, or irregular. Margins of masses may be circumscribed, obscured, microlobulated, indistinct or spiculated. Spiculated masses characterized by lines radiating from the margins (Kopans, 1998), tend to have a high likelihood of malignancy.

Calcifications are small calcium deposits on breasts. The process generating the deposits may be malignant or benign. Important characteristics include number, distribution, morphology and size. Size of calcifications ranges from 0.1 to 1 mm with an average diameter of 0.3 mm (Kopans, 1998).

Architectural Distortions (AD) are breast lesions in which the parenchyma is distorted. The distortion appears as if it is pulled into a central point, but does not have radio opaque central density (Radiology, 2003).

These features extracted by radiologists are graphically represented by cause and effect diagrams (Davis et al., 2005; Burnsidea et al., 2004; Kahn et al., 1995). A simple explanation of a cause and effect diagram is further explained. In an endeavor to correlate graphical analysis specifically with breast cancer detection, a simplified cause and effect diagram is shown in Figure 1.

In this paper, a brief background concerning causal networks, the probability theory and Bayesian networks, the available computational tools and platforms, the next steps towards a Bayesian Belief Network Implementation and finally, the conclusion are given.

BACKGROUND

Based on the various available references (Pearl, 2000; Pearl 1998; Neapolitan, 2003; Jensen, 2007); here, it is aimed to introduce Bayesian networks to radiologists, image processing scientists and data base specialists. To have an effective Bayesian network tool, multi-disciplinary inputs are required, thus, it is important for everyone involved to have a basic understanding and background.

Causality and causal networks

Consider the situation in which one feature of an entity has a direct influence on another feature of that entity. For example, the presence or absence of a disease in a human being has a direct influence on whether a test for that disease turns out positive or negative. The central aim of medical diagnosis using mammography findings is the elucidation of cause-effect relationships among variables or events. However, the appropriate methodlogy of extracting such relationships from available data and theory is a topic of current research. Contrary to the general notion that an abnormality is a cause of disease, causal relationship establishes that abnormality is the one of the effect of the disease. The abnormality may also be as a result of other causes. The two fundamental

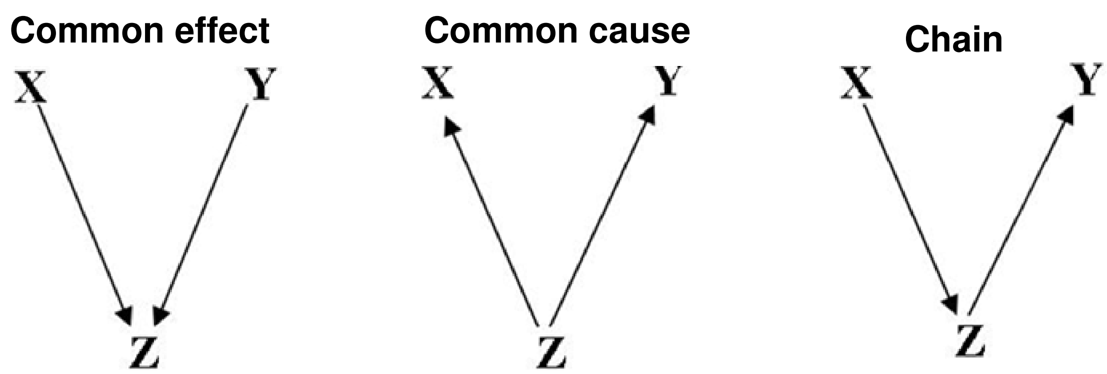

Figure 2. Causal network structures.

questions of causality are (1) what empirical evidence is required for legitimate inference of cause effect relationship? (2) given that we are still willing to accept causal information about a phenomenon, what inference can we draw from such information, and how?

Only in recent years, owing partly to advances in graphical models, causality has undergone a major transformation, that is, from a concept shrouded in mystery into a mathematical object with well defined semantics and well found logic.

The ability to infer causal relationships is crucial for scientific reasoning and, more generally, forms the basis for learning to act intelligently in the world. The knowledge of causal relationships, as opposed to mere statistical associations, gives us a sense of deep understanding of a system and a sense of potential control over the system's states, stemming from the ability to predict the consequences of actions that have not yet been performed (Pearl, 2000). There has been extensive study of how people make inferences about simple causal relationships, focusing on learning the relationship between a single cause and effect (Steyvers, 2003).

In the present research, we are interested in how networks involving multiple cause–effect relationships can be inferred for a possible breast cancer presence.

Causal networks are diagrams that indicate causal connections using arrows. As shown in Figure 2, Directed graphs provide us with an intuitive way to express causal knowledge. Nodes represent continuous or discrete state variables of a system and arrows represent direct causal relations, pointing from causes to effects. The state of each variable is modeled as some function of the states of its parents—its direct causes.

As stated in Molin (1989), physiological knowledge has a typical causal representation, which can be used in diagnostic expert systems to provide suitable explanations of the cause-effect relationships connecting findings with diagnostic conclusions. The experience gained in a diagnostic expert system combining heuristic and causal knowledge, makes it possible to use physiological information in assessing, the compatibility and consistency of diagnostic conclusions. In addition, it provides the opportunity of using the causal network to inquire directly about the domain, entering the system at different levels (findings, states and hypotheses) independently from patient data.

Bayes' theorem

In general, physicians learn medicine in a causal direction. Medical students are taught that a disease causes a certain constellation of symptoms, signs, and imaging findings, that is, probability that a finding has occurred in a given disease $P(finding|disease)$. In contrast, practicing physicians must reason in the opposite direction (Burnsidea, 2005). Given a certain constellation of symptoms, signs, and imaging findings, the patient has a probability of a certain disease, $P(disease|finding)$. Bayes' theorem enables physicians to go from the more intuitive causal conditional probability to the more clinically relevant conditional probability. In other words, Bayes' theorem calculates the probability that a disease is present in a finding, $P(disease|finding)$, using the probability of finding a given disease, $P(finding|disease)$, the probability of disease, $P(disease)$, and the probability of a finding, $P(finding)$ is given as:

$$P(disease \mid finding) = \frac{P(finding \mid disease)P(disease)}{P(finding)}$$

Physicians can calculate probabilities using Bayes' theorem without a computer if there are limited diagnostic parameters used to update the probability of a given disease. If the factors that modify the probability of disease become too numerous and interact, physicians often lack the time and computational abilities to perform these calculations.

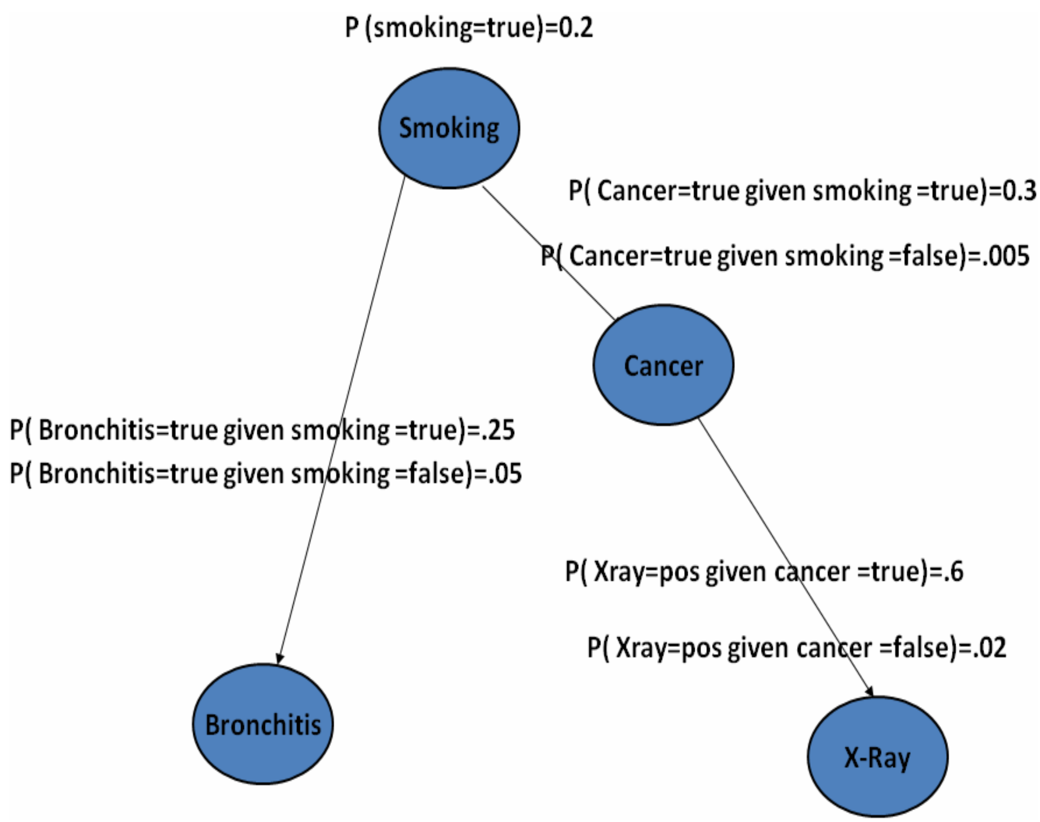

Figure 3. A Bayesian network.

We formally state the theorem:

Given two events E and F such that $P(E) \neq 0$ and $P(F) \neq 0$, we have:

$$P(E \mid F) = \frac{P(F \mid E)P(E)}{P(F)} \qquad (1)$$

Furthermore, given *n* mutually exclusive and exhaustive events $E_1, E_2, E_3, \ldots E_n$ such that $P(E_i) \neq 0$ all *i*, we have for $1 \leq i \leq n$,:

$$P(E_i \mid F) = \frac{P(F \mid E_i)P(E_i)}{P(F \mid E_1)P(E_1) + P(F \mid E_2)P(E_2) \ldots P(F \mid E_n)P(E_n)} \qquad (2)$$

Bayesian belief networks

A Bayesian Network (Pearl, 1988) is a graphical model that represents relationships of probabilistic nature among variables of interest. Such networks consist of a qualitative part (structural model), which provides a visual representation of the interactions amid variables, and a quantitative part (set of local probability distributions), which permits probabilistic inference and numerically measures the impact of a variable or sets of variables on others. Both the qualitative and quantitative parts determine a unique joint probability distribution over the variables in a specific problem. In other words, a Bayesian Network is a directed acyclic graph consisting of (Cooper, 1999): (a) Nodes (circles), which represent random variables; arcs (arrows), which represent probabilistic relationships among these variables and (b) for each node, there exists a local probability distribution attached to it, which depends on the state of its parents "(Cruz-Ramírez et al., 2007). Figure 3 taken from Neapolitan (2003) shows an example of a Bayesian Network and its correspondent conditional probabilities.

With reference to Neapolitan (2003), whether or not an individual has a history of smoking has a direct influence both on whether or not that individual has bronchitis. Also, the presence or absence of lung cancer has a direct influence on whether or not a chest X-ray is positive.

In this situation, we would want to do probabilistic inference involving features that are not related via a direct influence. We would want to deter-mine, for example, the conditional probabilities both of bronchitis and of lung cancer when it is known an individual smokes, and has a positive chest X-ray. Yet bronchitis has no direct influence (indeed no influence at all) on whether a chest X-ray is positive. Therefore, these conditional probabilities cannot be computed using a simple application of Bayes' theorem. There is a straightforward algorithm for computing them, but the probability values it requires are not ordinarily accessible; furthermore, the algorithm has exponential space and time complexity.

Bayesian networks were developed to address these difficulties. By exploiting conditional independencies entailed by influence chains, it is possible to represent a large instance in a Bayesian network using little space, and it is often possible to perform probabilistic inference among the features in an acceptable amount of time. In addition, the graphical nature of Bayesian networks gives us a much better intuitive grasp of the relationships among the features.

AVAILABLE COMPUTATIONAL TOOLS

Here, various tools available to implement inference algorithms in the context of a Bayesian Network are discussed. Jensen (2007) gives a list of such tools, some details about each tool are thus explained.

Bayesware

Bayesware discoverer

Bayesware Discoverer is an automated modeling tool that transforms a database into a network of dependencies, by searching for the most probable model responsible for the observed data. Bayesware Dis-coverer also provides sophisticated visualization tools, flexible connectivity capabilities and easy to use wizard interfaces to execute common tasks.

Bayesware classifier

Bayesware Classifier is a fast supervised classification program that is able to handle missing data. Based on a novel analytical methodology, Bayesware Classifier is able to trade-off the risk of various hypotheses about the patterns of missing data and provides both accurate analysis of your incomplete database and principled decisions about the predictions from the data.

BN PowerConstructor

A system that learns Bayesian belief network structures and parameters from data.

BN PowerPredictor

A data mining system for datamodeling/classification/ prediction. It extends BN PowerConstructor to BN based classifier learning.

DataPreprocessor

A tool used with BN PowerConstructor and BN PowerPredictor for pre-processing the training data.

BNT tool box

BNT supports many types of conditional probability distributions (nodes), decision and utility nodes, as well as chance nodes, that is, influence diagrams as well as Bayes nets. BNT supports static and dynamic Bayesian Networks (useful for modeling dynamical systems and sequence data). BNT supports many different inference algorithms, several methods for parameter learning and several methods for regularization. The source code is extensively documented, object-oriented, and free, making it an available tool for teaching, research and rapid prototyping.

Elvira

The Elvira is an available tool for the implementation of Bayesian expert systems. Elvira can operate in three modes, (1) Edit, to create and modify Bayesian networks and Influence Diagrams whose variables are only discrete. (2) Inference, to do evidence propagation and abduction. Most of the explanation capabilities of Elvira are offered only in this mode. (3) The learning mode allows the building of Bayesian networks from databases.

Elvira is implemented to be run in several languages. And, as other graphical tools, the user can open different windows simultaneously in order to perform different tasks.

Esthauge

The LIMID software system from Esthauge Decision Knowledge is a full Java software tool for construction and manipulation of graphical models in the form of Bayesian networks and decision graphs. It consists of an application programming interface for embedding with other Java software tools and a graphical user interface for visual construction, editing and maintenance of models.

Hugin

The Hugin Development Environment provides a set of tools for constructing model-based decision support systems in domains characterized by inherent uncertainty. The models supported are Bayesian networks and their extension influence diagrams. The Hugin Development Environment allows to define both discrete domain variables and to some extent continuous domain variables in your models.

The Hugin Decision Engine can be used through the Hugin Graphical User Interface. The Hugin Development Environment can also be used through one of several APIs (Application Program Interfaces), which come as libraries for C, C++, NET and Java, and as an ActiveX server.

The Hugin Development Environment can be used to construct models as components in applications for decision support, data mining and expert systems. The application communicates with the constructed component models through one of the Hugin APIs.

JavaBayes

JavaBayes is a system that handles Bayesian networks: It calculates marginal probabilities and expectations, produces explanations, performs robustness analysis, and allows the user to import, create, modify and export networks.

MSBNx

MSBNx is a component-based Windows application for creating, assessing, and evaluating Bayesian Networks, created at Microsoft Research. The application's installation module includes complete help files and sample networks. Bayesian Networks are encoded in an XML file format. The application and its components run on Windows 98, Windows 2000, and Windows XP.

Netica

Netica is a program for working with belief networks and influence diagrams. It has an intuitive and smooth user interface for drawing the networks, and the relationships between variables may be entered as individual probabilities in the form of equations, or learned from data files (which may be in ordinary tab-delimited form and have "missing data").

Samlam

Samlam is a tool for modeling and reasoning with Bayesian networks, developed in Java by the Automated Reasoning Group of Professor Adnan Darwiche at UCLA. Samiam includes two main components: A graphical user interface and a reasoning engine. The graphical interface allows users to develop Bayesian network models and to save them in a variety of formats. The reasoning engine supports many tasks including: Classical inference, parameter estimation, time-space tradeoffs, sensitivity analysis and explanation-generation.

FUTURE STEPS

The future steps towards implementing a Bayesian network based breast cancer detection aid tool are thus described.

Bayesian network model structure design

Our first task is to construct a belief network. Other Bayesian networks have been developed previously in this domain (Burnsidea et al., 2004; Kahn et al., 1995) and we will examine the differences between them and our web based requirements. The key problem is to construct the model so that we can capture all of the important aspects of system reality from the point of view of the diagnostic process (Przytula, 2000).

Data acquisition

The second task is to populate the knowledge base including conditional probability tables, it is assumed that the information needed comes from various sources such as meta data, demographical statistics and, most importantly, from experts. These sources must provide us with a simplified view of reality, which our model needs to simplify even further. Additionally, the knowledge base will be constructed from peer-reviewed medical literature, census data, and health statistics reports. When required, probability data is unavailable or the sample size is too small, an expert mammographer will provide subjective estimates of the probabilities.

Software selection and inference software implementation

A selection from the available software list as explained will be done in this stage, in this stage software requirements will be known, other image processing algorithms detection input output formats will be known as well. The web based interface will be implemented by the time of start of inference software algorithms implementation.

System integration and test

System Integration and Testing, in the context of software systems and software engineering, is the last step. A testing process will be adopted that exercises a software system's coexistence with others. Following this process, the deliverable systems will be passed onto acceptance testing.

Conclusion

Bayesian networks have an unparallel advantage of being able to exploit the explicit structure of the domain model to derive a graphical representation for learning. The encoding of independencies in the network topology admits the design of efficient procedures for performing computations over the network.

The Bayesian network is a powerful tool to describe the uncertainty and complexity of many problems in the real world. Within the idea of graphical model, a complex system consists of simple parts, which are bound together with probability theory. For the application of computer-aided detection in mammography, the researchers intend to design an interface between the project's Bayesian network learning algorithm and the radiologists, so that the radiologists can have interaction with the system by labeling only a small number of informative images presented by the active learning algorithm. In this way, the project's system only requires very little labeling effort from radiologists, while significant improvements of classifiers are achieved.

ACKNOWLEDGEMENT

Authors acknowledge the support received from the Department of Defense Award No. 05229004 “Development of a Knowledge Base to Support Detection and Diagnosis and Research in Mammography” in funding

this research.

REFERENCES

Burnside E (2005). Bayesian Networks: Computer-assisted Diagnosis Support in Radiology. Acad. Radiol., 12: 422-430.

Burnside ES, Rubinb DL, Shachter RD (2004). Using a Bayesian Network to Predict the Probability and Type of Breast Cancer Represented by Microcalcifications on Mammography. Proceedings of the 11th world congress on medical informatics, pp. 13-17.

Cooper GF (1999). An overview of the representation and discovery of causal relationships using Bayesian networks, in: C. Glymour, G.F. Cooper (Eds.), Computation, Causation and Discovery, AAAI Press/MIT Press, Cambridge, MA. pp. 3-62.

Cruz-Ramírez N, Acosta-Mesa HG, Carrillo-Calvet H, Alonso Nava-Fernández L, Barrientos-Martínez, RE (2007). Diagnosis of breast cancer using Bayesian networks: A case study. Comput. Biol. Med., 37: 1553-1564.

Davis J, Burnside B, Santos Costa ID, Page D, Shavlik J, Ramakrishnan R (2005). View Learning: An extension to SRL An application in Mammography. CS 731: Advanced methods in artificial intelligence, with biomedical applications, pages.cs.wisc.edu/~dpage/cs731/view.ppt.

Duffy SW, Tabár L, Chen H, Holmqvist M, Yen M, Abdsalah S, Epstein B, Frodis E, Ljungberg E, Hedborg-Melander C, Sundbom A, Tholin M, Wiege M, Åkerlund A, Wu H, Tung T, Chiu Y, Chiu C, Huang C, Smith RA, Rosén M, Stenbeck M, Holmberg L (2002). The impact of organized mammography service screening on breast carcinoma mortality in seven Swedish counties: A collaborative evaluation. Cancer, Am. Cancer Society, 95: 458-469.

Gajdos C, Tartter PI, Bleiweiss IJ, Hermann G, De Csepel J, Estabrook A, Rademaker AW (2002). Mammographic Appearance of Nonpalpable Breast Cancer Reflects Pathologic Characteristics. Annal. Surgery, 235(2): 246-251.

Jensen FV, Nielsen TD (2007). Bayesian Networks and Decision Graphs, (second edition). Website http://bndg.cs.aau.dk/, Springer Verlag, New York.

Kahn CE, Roberts LM, Wang K, Jenks D, Haddawy P (1995). Preliminary investigation of a Bayesian network for mammographic diagnosis of breast cancer. Proc Annu Symp Comput Appl Med Care. pp. 208-212.

Kopans DB (1998). Breast Imaging. Lippincott Williams. Philadelphia.

Molin G, Console L, Torass P (1989). Causal expert systems supporting medical decision making and medical education: explanations based on simulated situations. Engineering in Medicine and Biology Society, 1989. Images of the Twenty-First Century., Proceedings Annual Int. Conference IEEE Eng., 9: 1827-1828

Neapolitan RE (2003). Learning Bayesian Networks. Prentice Hall.

Pearl J (1988). Probabilistic reasoning in intelligent systems: networks of plausible inference, first ed. in: Representation and Reasoning, R.J.B.A.T.B. Laboratories, Morgan Kaufmann, San Mateo, California.

Pearl J (2000). Causality Modeling Reasoning and Inference. Cambridge University Press.

Przytula KW, Thompson D (2000). Construction of Bayesian networks for diagnostics. Aerospace Conference Proceedings, IEEE, 5: 193-200.

Steyvers M, Tenenbaum JB, Wagenmakers E, Blum B (2003). Inferring causal networks from observations and interventions. Cognitive Sci., 27: 453-489.