Computational Techniques to Improve the Early Diagnosis of Breast Cancer

Elizabeth S. Burnside, Ryan Woods, Turgay Ayer, Mehmet Ayvaci, Oguzhan Alagoz, Louis Oliphant, Jie Liu, Houssam Nassif, C. David Page, Jude Shavlik, David H. Gustafson

University of Wisconsin
SCHOOL OF MEDICINE AND PUBLIC HEALTH
Department of Radiology

INTRODUCTION

More than 20 million mammograms are performed annually in the US. Mammograms are the only proven way to detect breast cancer early and improve survival, but variability of practice can diminish efficacy. Radiologists estimate breast cancer risk based on imaging observations and make the decision to pursue additional imaging exams, recommend further imaging, or biopsy. We are developing Computer Assisted Diagnosis (CADx) tools to help physicians with interpretation, risk prediction, and decision–making.

We have developed two breast cancer risk prediction models—the **Mammography Bayesian Network (MBN)** and the **Mammography Artificial Neural Network (MANN)**—based on mammographic descriptors and demographic factors. We have also created a model to differentiate invasive and in situ breast cancer.

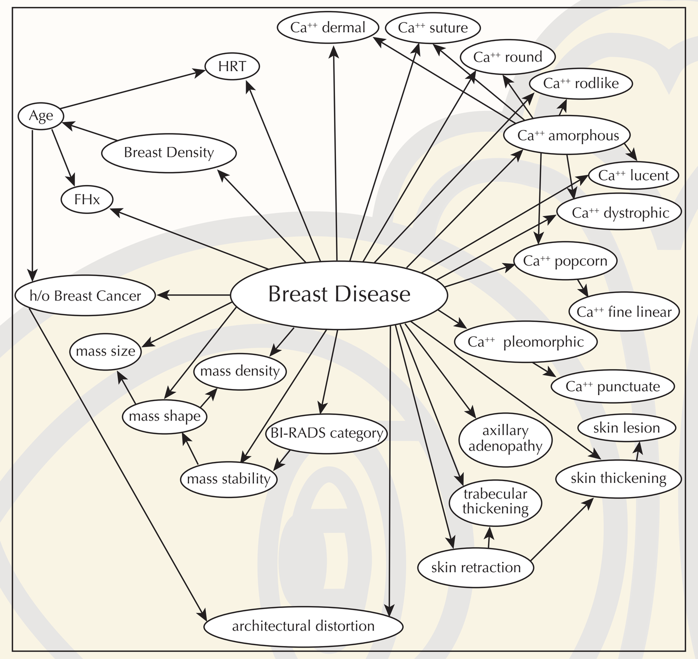


Figure 1

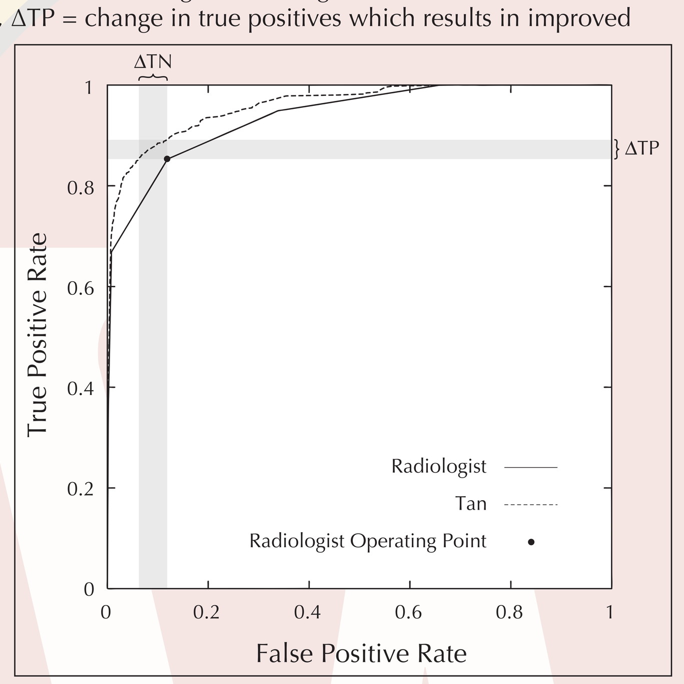
Mammography Bayesian Network

Graph shows the structure of the trained BN. Labeled circles represent nodes and arrows (arcs) represent dependence relationships.
Note: Ca ++ = calcifications
HRT = Hormone replacement therapy
FHx = Family history of breast cancer
h/o = history of

Fig 2

Mammography Bayesian Network Performance

ROC curves constructed from 1) BI–RADS categories of the radiologists, 2) the predicted probabilities of the BN. Note– ROC= Receiver Operator Characteristic, BN= Bayesian Network, ΔTN = change in true negatives which results in improved specificity, ΔTP = change in true positives which results in improved sensitivity. The radiologists operating point is considered the BI–RADS 3 point corresponding to a threshold above which biopsy would be recommended.



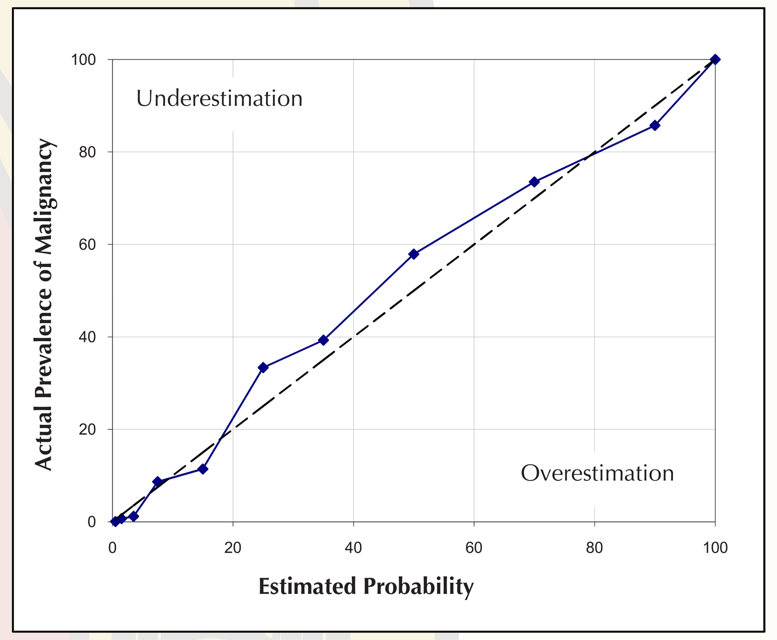
MAMMOGRAPHY BAYESIAN NETWORK

A Bayesian network is a probabilistic graphical model that uses nodes (ovals) to represent predictive variables, and arcs (arrows) to represent dependence relationships among these variables. In our application, we are interested in predicting the probability of breast cancer from demographic risk factors and mammographic findings. The predictions generated by the MBN have the potential to both inform radiologists to aid in mammographic interpretation, as well as aid patients and referring physicians in clinical decision–making. A recent study (in press) demonstrates that the MBN can exceed the performance of radiologists in predicting the risk of breast cancer in a large, population based dataset. We collected structured reports from 48,744 consecutive pooled screening and diagnostic mammography examinations on 18,270 patients from 4/5/1999 to 2/9/2004. We matched 62,219 mammographic findings with our state cancer registry which served as our reference standard. Using an algorithm called Tree Augmented Naïve Bayes (TAN) and 10-fold cross–validation, we built a Mammography Bayesian network (Figure 1) from this data and evaluate the performance of radiologists compared with the BN using area under the ROC curve (AUC), sensitivity, and specificity. The BN significantly exceeded the performance of interpreting radiologists in terms of AUC (0.960 vs. 0.939, *P* < 0.002, Figure 2), sensitivity (90.0% vs. 85.3%, *P* < 0.001) and specificity (93.9% vs. 88.1%, *P* < 0.001).

Figure 4

Mammography Artificial Neural Network Performance

Calibration curve of the MANN (solid line) and the curve representing perfect calibration (dashed line). This is a plot of the actual prevalence of malignancy versus estimated risk of malignancy for each quartile of the probability scale.



MAMMOGRAPHY ARTIFICIAL NEURAL NETWORK

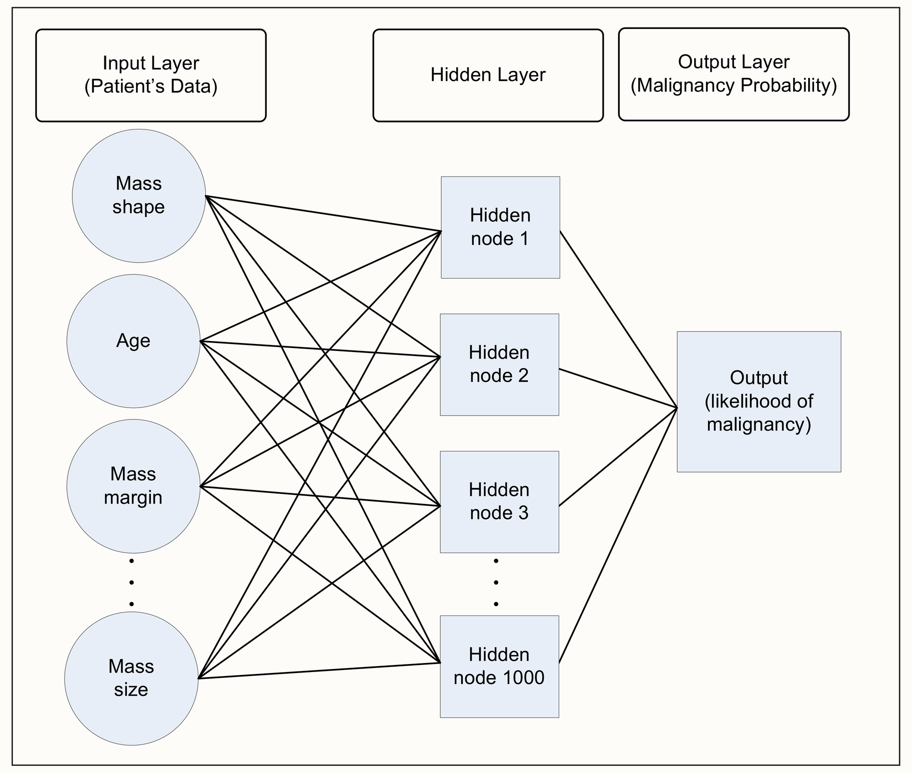
Artificial Neural Networks are machine learning tools that we use to aid in mammographic interpretation, as well as aid patients and referring physicians in clinical decision–making.

We built a three–layer feed–forward neural network to estimate the likelihood of malignancy using the same dataset detailed above. The layers of Mammography Artificial Neural Network (MANN) included an input layer of mammography findings, a hidden layer of nodes, and an output layer with a single node generating the probability of malignancy (Figure 3).

Figure 3

Mammography Artificial Neural Network

The structure of the MANN. Inputs to the MANN include lesion descriptors and demographic risk factors of the patient. The single output shows the likelihood of malignancy.



Discrimination (the ability to separate benign abnormalities from malignant ones) and calibration (the ability to stratify individuals into accurate risk categories) are the two main components of accuracy in a risk prediction model.

We measured both discriminative ability (AUC), and calibration capacity (using the Hosmer–Lemeshow (H–L) test and calibration curves) to assess the accuracy of risk prediction for our MANN.

The AUC of the MANN model, 0.965, was significantly better than that of the radiologists alone, 0.939 (*P*<0.01).

The H–L statistic for the MANN was 12.46 (*P*=0.13, df=8) [*P*>.05 indicated good calibration]. The precision of the predictive probabilities was shown by a calibration curve (Figure 4), which presented the ability of the model to enable prediction of probabilities across all ranges of risk. Our MANN model is well calibrated as demonstrated by low values of the H–L statistic and the corresponding high *P* value, as well as the calibration curve.

DISTINGUISHING INVASIVE VERSUS IN SITU DISEASE: AGE MATTERS

We developed a Bayesian Network to predict the risk of ductal carcinoma in situ (DCIS) versus invasive breast cancer using mammography features and patient demographic factors. The Bayesian Network helped quantify the probability of invasive cancer given demographic factors and mammographic findings. Out of 2211 malignant cases, there were 1790 correct classification with 1305 of them being invasive cancer, 239 false negatives and 182 false positives when 50% is deemed as the level above which the prediction will be invasive (positive). Closely examining the misclassifications, we found out that the total error rates and the False Negative Rate (Predicting in situ cancer as being invasive) tend to go down for older ages (Table 1).

Table 1

Error Rates in Age Groups

Age Range	Total Cancers	# Invasive	Total Error Rate
LT 45	231	156	29.44%
45-49	244	157	20.49%
50-54	300	204	17.33%
55-59	306	196	15.03%
60-64	265	195	19.25%
65-69	226	159	19.91%
70-74	234	167	17.95%
75-79	218	155	15.60%
GTE 80	187	155	17.65%
Total	2211	1544	19.04%

We then demonstrate that the Bayesian Network performed better in older women compared to younger women. First, we show that the ROC curve for the older women dominates the ROC curve for the younger women Figure 5.

Figure 5

ROC curve of TAN Trained Model with Age Stratification

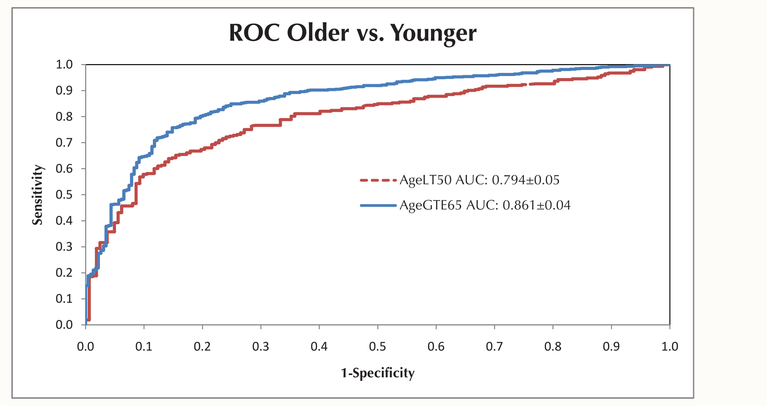
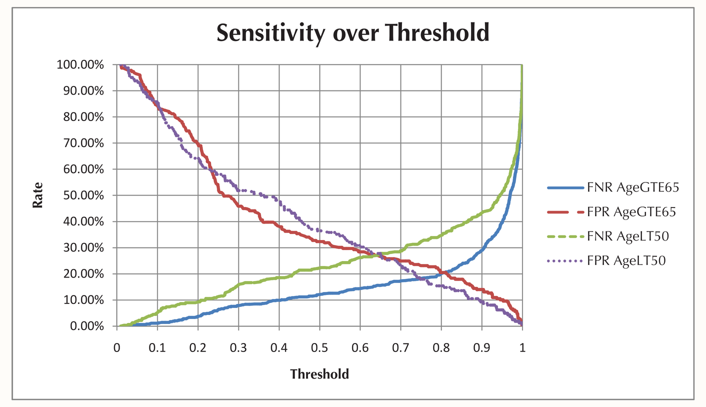


Figure 6

Stratified Error Rate Trade–Off



Secondly, we showed that the improved performance of the Bayesian Network in older women is attributable to a decrease in False Negatives which is important clinically. Figure 6 shows both the improved performance in FNR via age stratification and also the trade–off between FNR and FPR over different threshold values.

Understanding the risk of invasive disease can aid in the clinical management decisions such as the need for increased sampling at biopsy and the appropriate selection of surgical interventions. The ability to accurately predict the probability of DCIS versus invasive disease would enable older women and their referring physicians to make more informed decisions about managing their breast health in the context of their expected life span and co–morbidities.

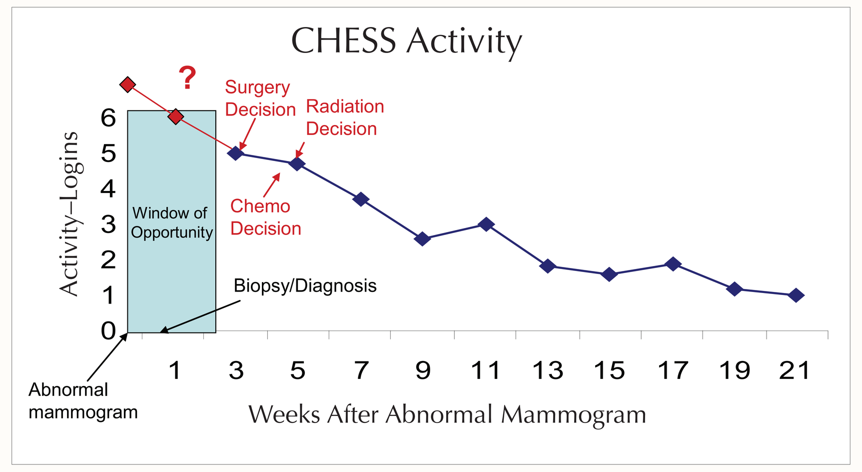
Window of Opportunity

Developing decision models coupled with CHESS may:

- Improve patient QOL during diagnosis
- Decrease anxiety, increase information competence, and increase autonomy
- Facilitate shared decision–making
- Improve quality of care

Figure 7

Opportunity for Decision–making



DISSEMINATION

When facing the breast cancer diagnostic process, women and their families need information and support to help sort through key decisions they must make. An Interactive Cancer Communication System (ICCS) and the Comprehensive Health Enhancement Support System (CHESS), have each been accepted, used, and endorsed by patients, and have successfully improved the quality of life for patients diagnosed with breast cancer. The current model of CHESS helps women *after* a diagnosis of breast cancer. We believe that, with the help of the decision–making models we are creating, CHESS could help women *during* the diagnostic process. Interactive education during the diagnostic process provides an opportunity for patients to explore data before they are overwhelmed by the information given at the time of diagnosis.

Corresponding Author:
Elizabeth Burnside
University of Wisconsin Department of Radiology
F6/186 CSC
Madison, WI 53792–3252
Ph: 608.263.5585
Fax: 608.263.0876