Computational Techniques to Improve the Early Diagnosis of Breast Cancer

Elizabeth S. Burnside, Jagreet Chhatwal, Karin E. Witte, Jie Liu, Turgay Ayer, Oguzhan Alagoz, C. David Page, David H. Gustafson

INTRODUCTION

More than 20 million mammograms are performed annually in the U.S. Many mammograms are the only pathway to detect breast cancer early and are the best available tool for breast cancer detection with the greatest efficacy. mammographic interpretation involves two components—

- **mammographic interpretation:** Radiologist task for observation.
- **multimodal intermodual interpretation:** Radiologist task for observation.

In the mammographic interpretation, radiologists look for abnormalities on digital mammograms and make a diagnosis based on their clinical judgment. They use computer-aided detection (CAD) tools to help radiologists identify abnormalities, but no such tools are available to help radiologists interpret mammographic images. The MBN and the MLRM models identify abnormalities, but no such tools are available to help radiologists interpret mammographic images.

We developed two breast cancer risk prediction models—the Mammography Bayesian Network (MBN) and the Mammography Logistic Regression Model (MLRM). The MBN and the MLRM models can help radiologists improve their diagnostic efficacy. Mammographic interpretation involves two components—

- **early and improve survival:** While breast cancer intervals can decrease the probability of breast cancer in normal cases.
- **variability of practice:** Decreases the probability of breast cancer in normal cases.

The purpose of this study is to demonstrate that computational techniques can improve the early diagnosis of breast cancer.

MAMMOGRAPHY BAYESIAN NETWORK

A Bayesian network (BN) is a probabilistic graphical model that assumes local tendency to represent predicted variables, and arcs among arrows indicate dependencies among these variables. The BN has the ability to incorporate available data as well as breast image knowledge. In our application, we are interested in predicting the probability of breast cancer or cancer node from mammographic risk factors and mammographic findings. The predictors generated by the MBN have the potential to help radiologists and mammographic interpretation, as well as with pathologists and referring physicians in clinical decision-making.

We constructed the MBN using all variables observed by radiologists in their daily practice of interpreting mammograms. Results show that the MBN significantly improves the performance of the radiologists in their daily practice of interpreting mammograms. The performance of each model was included all the variables used in Model–1, but added the radiologists' coefficient of the independent variable (dose-risk score) that is estimated using the available data set.

The probability of cancer was the outcome in both models. Model–2 included all the variables used in Model–1, but added the radiologists' coefficient of the independent variable (dose-risk score) that is estimated using the available data set.

In our application, we are interested in predicting the probability of breast cancer or cancer node from mammographic risk factors and mammographic findings. The predictors generated by the MBN have the potential to help radiologists and mammographic interpretation, as well as with pathologists and referring physicians in clinical decision-making.

We created two logistic regression models based on the mammographic and demographic data for breast cancer patients included in the Breast Imaging–Reporting and Data System (BI–RADS) lexicon, and some recorded in the MDIS format between 1992 and 2017 at the University of Wisconsin–Milwaukee. State cancer registry outcomes matched with our data set in the aforementioned categories.

MAMMOGRAPHY LOGISTIC REGRESSION MODEL

A logistic regression is a statistical model that is used for predicting the probability of breast cancer or cancer node, when the dependent variable (cancer node) is a binary variable. In the logistic regression model, the dependent variable (cancer node) is a binary variable.

The logistic regression model can be written as:

\[ \log \left( \frac{p}{1-p} \right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k \]

where \( p \) is the probability of breast cancer, \( x_1, x_2, \ldots, x_k \) are the independent variables, and \( \beta_0, \beta_1, \beta_2, \ldots, \beta_k \) are the coefficients of the independent variables.

In our application, we are interested in predicting the probability of breast cancer or cancer node from mammographic risk factors and mammographic findings. The predictors generated by the MBN have the potential to help radiologists and mammographic interpretation, as well as with pathologists and referring physicians in clinical decision-making.

OPTIMAL BIOPSY DECISION MODEL

We designed decision trees (DTs), a mathematical framework for sequential decision making—determine the optimal time of biopsy. Biopsy is often performed because a patient requests a biopsy (e.g., biopsy for cancer), and the radiologist designs an action (e.g., recommending biopsy or suggesting further investigation). In the next few lines and in 12 months, the patient moves to other possible states probabilistically and the next transition results in an event (e.g., expected biopsy). If the biopsy is performed, the patient can move to either a benign or malignant biopsy result. If the radiologist diagnoses the patient as having breast cancer, the patient can move to the malignant biopsy result. If the result of the biopsy is benign, the patient can move to the benign biopsy result. By solving a set of mathematical equations, the DT framework provides the optimal biopsy decisions by estimating the probability of cancer (risk score) based on mammographic findings and demographic factors.

Optimal Policies of the ODEMA

Figure 1 shows the optimal probability, threshold for biopsy decisions in different age groups. The optimal policy can be interpreted as the door to the decision tree that helps the radiologist to determine the optimal age of biopsy.

**Figure 1:** Optimal Age–dependent policy to perform biopsy.

When facing a breast cancer diagnosis, women and their families need information and support to help them make decisions they must make. An Intersecting Cancer Communication System (ICCS) and the Complementary Health Education Support Program (CHESS) have 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 ODEMA helps women after a diagnosis of breast cancer. We believe that, while helping the decision-making process in other health areas, interaction 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.

**Figure 6:** Opportunity for Decision–making

**References**


