Original Investigation

Association between Breast Parenchymal Complexity and False-Positive Recall From Digital Mammography Versus Breast Tomosynthesis: Preliminary Investigation in the ACRIN PA 4006 Trial

Shonket Ray, PhD, Lin Chen, PhD, Brad M. Keller, PhD, Jinbo Chen, PhD, Emily F. Conant, MD, Despina Kontos, PhD

Abbreviations and Acronyms

FP
false-positive

DM
digital mammography

PPV
positive predictive value

DBT
digital breast tomosynthesis

ACR BI-RADS
American College of Radiology Breast Imaging-Reporting and Data System

IRB
institutional review board

ACRIN
American College of Radiology Imaging Network

TP
true-positive

MLO
mediolateral oblique

Rationale and Objectives: We investigate associations between measures of mammographic parenchymal complexity and false-positive (FP) recall from screening with digital mammography (DM) versus digital breast tomosynthesis (DBT).

Materials and Methods: We retrospectively analyzed data from 541 women recruited by the American College of Radiology Imaging Network 4006 trial, designed to evaluate callback and detection rates from screening with DM versus combined DM and DBT. Of these, 68 and 56 were FPs based on DM alone versus the combined DM/DBT readings, respectively. Mammographic complexity was quantified with computerized texture analysis and percent density. Logistic regression was performed to evaluate associations between extracted features and FP recall, after adjusting for age and number of previous benign biopsies. Odds ratios and area under the curve (AUC) of the receiver operating characteristic were used to assess association strength.

Results: For DM, age, previous benign biopsies and texture features of correlation, inverse difference moment, sum average, and sum variance were deemed as significant predictors ($P < .05$) of FP recall, with an AUC = 0.77. For DBT, age was the only significant predictor of FP recall with AUC = 0.64. Using exploratory receiver operating characteristic thresholds for which no true-positives would be missed, a potential FP reduction of 23.5% and 8.9% was demonstrated, respectively, for DM alone versus DM/DBT.

Conclusion: Measures of breast complexity measured on 2D digital mammograms are indicative of the likelihood for FP recall from screening with DM, and could help identify women who could benefit from supplemental screening, including DBT.

Key Words: Digital mammography; digital breast tomosynthesis; breast cancer risk assessment; texture analysis; false-positive recall.

Acad Radiol 2016; ■:■■–■■

From the Department of Radiology, University of Pennsylvania, 1 Silverstein Bldg., 3400 Spruce St, Philadelphia PA 19104 (S.R., L.C., B.M.K., E.F.C., D.K.); Department of Biostatistics and Epidemiology, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania (J.C.). Received February 9, 2016; accepted February 29, 2016. **Sources of Support:** This work was supported in part by the National Institutes of Health (NIH) 5R01CA161749-02 and U54CA163313-01 grants, and a Breast Cancer Alliance Young Investigator Grant. We also thank the American College of Radiology Imaging Network for the access to the 4006 trial data. **Address correspondence to:** D.K. e-mail: Despina.Kontos@uphs.upenn.edu

http://dx.doi.org/10.1016/j.acra.2016.02.019

DICOM
Digital Imaging and Communications in Medicine

PD
percent density

LIBRA
Laboratory for Individualized Breast Radiodensity Assessment

FCM
fuzzy c-means

ROI
region of interest

GLCM
Gray Level Co-occurrence Matrix

VIF
variance inflation factors

OR
odds ratio

AUC
area under the curve

ROC
receiver operating characteristic

CI
confidence interval

TN
true negative

BCI
Breast Complexity Index

INTRODUCTION

False-positive (FP) recall from breast cancer screening is defined as being called back from screening for further assessment without breast cancer being subsequently detected. Currently, the majority of recalls from screening with digital mammography (DM) are FP, contributing to an overall low positive predictive value (PPV) (1) and concerns for over-imaging (1). This issue of FP recalls has been cited as a major contributing factor in many recent studies calling into question the true benefits and effectiveness of breast cancer screening programs (2–4). Reducing unnecessary FP events, without compromising the sensitivity of cancer detection, could reduce the overall cost related to any breast cancer screening program (5) due to less diagnostic workup and fewer biopsy procedures, while also reducing unnecessary physical pain, psychological trauma and, potentially, scarring from unnecessary biopsy procedures (6). Digital breast tomosynthesis (DBT) is emerging as an alternative modality to complement or even replace DM, which has spurred the investigation of its potential benefits within the screening paradigm, especially in reducing FP recalls (7,8). In DBT, tomographic breast images are reconstructed from multiple X-ray projections to produce a "quasi-3D" rendition of the breast, allowing superior breast tissue visualization and reducing the effect of tissue superposition inherent to 2D imaging (9–11).

In an ideal personalized paradigm, a prior test could be used to identify women at high-risk for such FP events, allowing for better implementations of individualized breast cancer screening protocols where supplemental screening tests could be offered to women for whom conventional mammography may have low specificity. Presently, mammographic breast density is a known risk factor for FP events (12). Breast density, which is commonly used to characterize the complexity of the breast tissue, is also known to confound the sensitivity of screening mammography (13). However, the commonly used mammographic density measures (ie, percent density [PD] and/or the ACR BI-RADS categories) are global imaging descriptors that may not adequately capture the spatially localized conspicuities and complexity of the breast parenchymal tissue, which may be driving FP recalls (14). We hypothesize that mammographic features of parenchymal texture, computed in a spatially adaptive manner which characterizes local regions of confounding tissue texture, can provide complementary information to the more global density measures in describing the complexity of the breast tissue and therefore augment the