ORIGINAL ARTICLE

Effect of Arrival of Prior Mammograms on Recall Negation for Screening Mammograms Performed With Digital Breast Tomosynthesis in a Clinical Setting

SA-CME

Lara A. Hardesty, MD[a], *Kimberly E. Lind, PhD*[a], *Eric J. Gutierrez, BS*[b]

Credits awarded for this enduring activity are designated "SA-CME" by the American Board of Radiology (ABR) and qualify toward fulfilling requirements for Maintenance of Certification (MOC) Part II: Lifelong Learning and Self-assessment. To access the SA-CME activity visit https://cortex.acr.org/Presenters/CaseScript/CaseView?CDId=alITT4LNIck%3d.

Abstract

Purpose: This retrospective study evaluates the effect of comparison with prior mammograms on recall negation for screening mammography performed with digital breast tomosynthesis (DBT) in a clinical setting and compares it with that performed without DBT.

Methods: This is an Institutional Review Board–approved, HIPAA-compliant retrospective review of the electronic medical record for all nonbaseline screening mammograms performed in clinical practice over 13 months. For each mammogram, we recorded if DBT were used, the BI-RADS assigned at initial interpretation, and whether prior mammograms were available at initial interpretation. If prior mammograms arrived later for comparison, the final BI-RADS assigned after comparison was recorded. A mammogram assigned a BI-RADS 0 at initial interpretation and assigned a BI-RADS 1 or BI-RADS 2 after prior mammograms arrived for comparison was labeled as a recall that was negated by the arrival of prior mammograms. The number of recalls negated for mammograms that used DBT was compared with that for mammograms that did not use DBT.

Results: Arrival of prior mammograms for comparison negated the need for recall for mammograms performed with DBT by 67.67% and negated the need for recall for mammograms performed without DBT by 55.80%. After adjusting for age, density, and time between mammograms, the percentage of recalls negated by comparison with prior mammograms was not significantly different for mammograms performed with DBT than it was for those performed without DBT.

Conclusion: Comparison with prior mammograms remains important for the minimization of recall rates during the use of DBT for screening mammography in the clinical setting.

Key Words: Digital breast tomosynthesis, screening mammography, recall rates, comparison with prior mammograms

J Am Coll Radiol 2018;■:■-■. Copyright © 2018 American College of Radiology

INTRODUCTION

Screening mammography reduces mortality from breast cancer by 20% to 40% for women 40 years old and older [1]. However, the majority of women recalled from screening mammography for additional evaluation do not have breast cancer. These false-positive recalls contribute to patient anxiety and to the cost of screening mammography programs. Minimizing false-positive recalls is essential for a high-quality screening mammography program. Multiple studies of film-screen

[a]Department of Radiology, University of Colorado Anschutz Medical Campus School of Medicine, Aurora, Colorado.

[b]University of Colorado Anschutz Medical Campus School of Public Health, Aurora, Colorado.

Corresponding author and reprints: Lara A. Hardesty MD, Associate Professor of Radiology, University of Colorado Anschutz Medical Campus School of Medicine, Dept of Radiology, 12401 East 17th Avenue, Mailstop L954, Aurora, CO 80045; e-mail: Lara.Hardesty@ucdenver.edu.

The authors have no conflicts of interest related to the material discussed in this article.

1546-1440/18/$36.00 ■ https://doi.org/10.1016/j.jacr.2018.05.003

mammography and full-field digital mammography (FFDM) have shown that comparing the current screening mammogram to prior mammograms significantly decreases the number of false-positive recalls [2-11]. The new mammographic technology of digital breast tomosynthesis (DBT) increases sensitivity and specificity of screening mammography relative to FFDM [12-14]. With the improvement in interpretation provided by DBT, the benefit of comparison with prior mammograms in a clinical setting is unknown. Reader performance studies in a laboratory setting show that comparison with prior mammograms decreases recall rates for test sets of screening mammograms performed with DBT [15,16]. However, radiologists' interpretations in reader performance studies can vary from their interpretations in the clinical environment [17]. The purpose of this study was to evaluate the impact of comparison with prior mammograms on the number of women recalled from screening mammography using DBT in a clinical setting. We compared the impact of the availability of prior mammograms on the interpretation of mammograms performed with DBT with the impact on the interpretation of mammograms performed without DBT.

METHODS

With Institutional Review Board approval, we assembled a retrospective cohort of all women receiving nonbaseline mammograms between December 1, 2012, and December 31, 2013, at our institution. During that time, we offered DBT screening mammography using combined FFDM and DBT and we offered FFDM screening mammography using only FFDM without DBT. The use of DBT versus FFDM was based upon the location where each patient self-selected to undergo her mammogram, with the mobile unit offering only FFDM and the hospital outpatient office and the satellite outpatient office both offering only DBT.

During the period of this study, one of four full-time breast-imaging radiologists, each of whom was fellowship-trained in breast imaging, interpreted each mammogram performed at our institution. The rotating clinical schedule assigned to each radiologist the responsibility for the interpretation of screening mammograms performed on the mobile mammography unit using FFDM or the responsibility for the interpretation of screening mammograms performed at the outpatient offices using DBT. The number of days interpreting each type of study and location was divided equally among the four breast imagers.

Our institution began performing DBT mammography in May 2012. Each of the four breast imagers who interpreted mammograms during the study period of December 2012 to December 2013 had similar experience (6 months of clinical experience at the start of the study) with the interpretation of DBT mammograms. Each of the four breast imagers had more than 6 months of clinical experience interpreting FFDM mammograms before the start of the study.

During clinical interpretation, if prior mammograms were available and the radiologist saw a finding that needed additional evaluation, the mammogram was interpreted; a BI-RADS Assessment Category 0, Need Additional Imaging [18], was assigned; and the patient was recalled to complete the additional imaging. During initial clinical interpretation, if prior mammograms were not available and the radiologist saw a finding that needed additional evaluation, the mammogram was interpreted and a BI-RADS Assessment Category 0, Need Additional Imaging and/or Prior Mammograms for Comparison, was assigned. This prompted the file room staff to request the patient's prior mammograms from the outside institution for comparison. If the prior mammograms arrived from the outside institution, they were provided to the radiologist who performed the original interpretation and the original report was added to provide the final interpretation. If the prior mammograms were not received from the outside institution within 2 weeks, the patient was recalled to complete the additional evaluation that had been recommended in the initial report.

For this study, a fellowship-trained breast-imaging radiologist with 20 years of experience abstracted data from the electronic medical record for each nonbaseline mammogram. Data recorded for each mammogram included how it was performed (DBT or FFDM), patient's age and breast density, whether prior mammography was available for comparison at the time of initial interpretation, and the BI-RADS Assessment Category that was assigned at initial interpretation. If prior mammograms had not been available at the time of initial interpretation, the radiologist recorded if prior mammography arrived later for comparison and the BI-RADS Assessment Category that was assigned at the time of subsequent comparison and final interpretation. For the purposes of this study, if the initial interpretation when prior mammograms were not available was a BI-RADS Assessment Category 0 and the arrival of prior mammograms for comparison led to a final BI-RADS Assessment Category of 1 Negative or 2 Benign, then the arrival of the prior mammograms led to a "negation" of the original recommendation for recall.