Original Investigation

Can Occult Invasive Disease in Ductal Carcinoma *In Situ* Be Predicted Using Computer-extracted Mammographic Features?

Bibo Shi, PhD, Lars J. Grimm, MD, MHS, Maciej A. Mazurowski, PhD, Jay A. Baker, MD, Jeffrey R. Marks, PhD, Lorraine M. King, PhD, Carlo C. Maley, PhD, E. Shelley Hwang, MD, MPH, Joseph Y. Lo, PhD

Rationale and Objectives: This study aimed to determine whether mammographic features assessed by radiologists and using computer algorithms are prognostic of occult invasive disease for patients showing ductal carcinoma *in situ* (DCIS) only in core biopsy.

Materials and Methods: In this retrospective study, we analyzed data from 99 subjects with DCIS (74 pure DCIS, 25 DCIS with occult invasion). We developed a computer-vision algorithm capable of extracting 113 features from magnification views in mammograms and combining these features to predict whether a DCIS case will be upstaged to invasive cancer at the time of definitive surgery. In comparison, we also built predictive models based on physician-interpreted features, which included histologic features extracted from biopsy reports and Breast Imaging Reporting and Data System-related mammographic features assessed by two radiologists. The generalization performance was assessed using leave-one-out cross validation with the receiver operating characteristic curve analysis.

Results: Using the computer-extracted mammographic features, the multivariate classifier was able to distinguish DCIS with occult invasion from pure DCIS, with an area under the curve for receiver operating characteristic equal to 0.70 (95% confidence interval: 0.59–0.81). The physician-interpreted features including histologic features and Breast Imaging Reporting and Data System-related mammographic features assessed by two radiologists showed mixed results, and only one radiologist's subjective assessment was predictive, with an area under the curve for receiver operating characteristic equal to 0.68 (95% confidence interval: 0.57–0.81).

Conclusions: Predicting upstaging for DCIS based upon mammograms is challenging, and there exists significant interobserver variability among radiologists. However, the proposed computer-extracted mammographic features are promising for the prediction of occult invasion in DCIS.

Key Words: Breast cancer; ductal carcinoma *in situ*; digital mammogram; microcalcification; CAD.

INTRODUCTION

Ductal carcinoma *in situ* (DCIS) is a preinvasive tumor confined within the ducts of the mammary glands (1) and lies along the breast cancer continuum between atypical ductal hyperplasia and invasive ductal carcinoma. The incidence of DCIS has increased substantially since the introduction of mammographic screening, with over 60,000 women in the United States diagnosed with DCIS every year, representing approximately 20% of all new breast neoplasm diagnoses (2). However, despite the increased incidence of DCIS, there has not been a concomitant decrease in invasive breast cancer (3). Since the risk of progression from DCIS to invasive cancer is unclear, with estimates ranging from 14% to 53% (4), there is a growing debate about overdiagnosis and consequent overtreatment of DCIS. Furthermore, among DCIS-only cases diagnosed at core biopsy, approximately 26% will be shown to contain invasive ductal carcinoma at surgical excision (5). This upstaging, specifically from DCIS diagnosed at core biopsy to invasive ductal carcinoma at excision, has important consequences for patient management.

Many studies have sought to predict the occult invasion in DCIS. Different factors or markers, including immunohistochemical biomarkers, histologic features, and mammographic or sonographic findings, have been described and associated with

Acad Radiol 2017; ■:■■–■■

From the Carl E. Ravin Advanced Imaging Laboratories, Department of Radiology, Duke University School of Medicine, 2424 Erwin Rd, Suite 302, Durham, NC 27705 (B.S., L.J.G., M.A.M., J.Y.L.); Department of Radiology (J.A.B.); Department of Surgery, Duke University School of Medicine, Durham, North Carolina (J.R.M., L.M.K., E.S.H.); Biodesign Center for Personalized Diagnostics and School of Life Sciences, Arizona State University, Tempe, Arizona (C.C.M.); Centre for Evolution and Cancer, Institute of Cancer Research, London, United Kingdom (C.C.M.). Received December 2, 2016; revised March 15, 2017; accepted March 16, 2017. This work was supported in part by the National Institutes of Health/National Cancer Institute (R01-CQA185138) and the Department of Defense Breast Cancer Research Program (W81XWH-14-1-0473). **Address correspondence to:** B.S. e-mail: bibo.shi@duke.edu

http://dx.doi.org/10.1016/j.acra.2017.03.013

outcomes in DCIS (5–11). However, none of these factors have been accepted as a definitive predictor of this upstaging or are sufficiently reliable for clinical use. Overall, it still remains a difficult task and unmet need to accurately predict occult invasive disease in DCIS.

Breast microcalcifications (MCs) appear in 30%–50% of mammographically detected cancers (12), and over 90% of women with DCIS have suspicious MCs on mammography (13). There has been much work using computer-aided detection (CAD) and computer-aided diagnosis (CADx) for mammography including MC clusters (14–20). Other CAD or CADx studies have focused on DCIS (21–24), but those studies have not utilized the diagnostic magnification views routinely available during the workup of suspicious calcifications, which offer additional details not appreciable on routine full-field screening mammographic views. In this work, we hypothesize that computer-vision techniques as well as various mammographic features developed for screening detection or diagnosis can be used to help predict the presence of occult invasive disease associated with DCIS. We have, therefore, developed a computer-vision algorithm-based approach to extract mammographic features for patients with DCIS, and built a classification model relying on these features to distinguish between pure DCIS and DCIS with occult invasive disease.

MATERIALS AND METHODS

The processing pipeline to analyze the digital mammography magnification views is shown in Figure 1. A core component of our prediction model is a computer-vision algorithm-based approach to extract mammographic features (denoted as computer-vision features in the remaining sections). In addition, we also extracted and evaluated some histologic features and Breast Imaging Reporting and Data System (BI-RADS)-related mammographic features that have been used in the literature (1,5,6,25–29) (denoted as physician-interpreted features in the remaining sections), and compared their predictive power for occult invasive disease in DCIS with the computer-vision features.

Subject Selection

The present study was approved by the institutional review board of Duke University. All procedures were compliant with the Health Insurance Portability and Accountability Act. Our subjects included women aged 40 and older who underwent stereotactic core needle biopsy and were diagnosed with pure DCIS that presented as calcifications only, and for whom at least one digital magnification view was available. Exclusion criteria included the presence of any masses, asymmetries, or architectural distortion on a mammogram; a history of breast cancer or prior surgery; and the presence of microinvasion at the time of initial biopsy. The excluded subjects were deemed to be already at high risk of invasion and therefore were not appropriate for the present study. Overall, 99 subjects from our institution with biopsy-proven DCIS only were retrieved from 2009 to 2014. Of those, 25 were upstaged to invasive cancer at the time of definitive surgery. All magnification views were

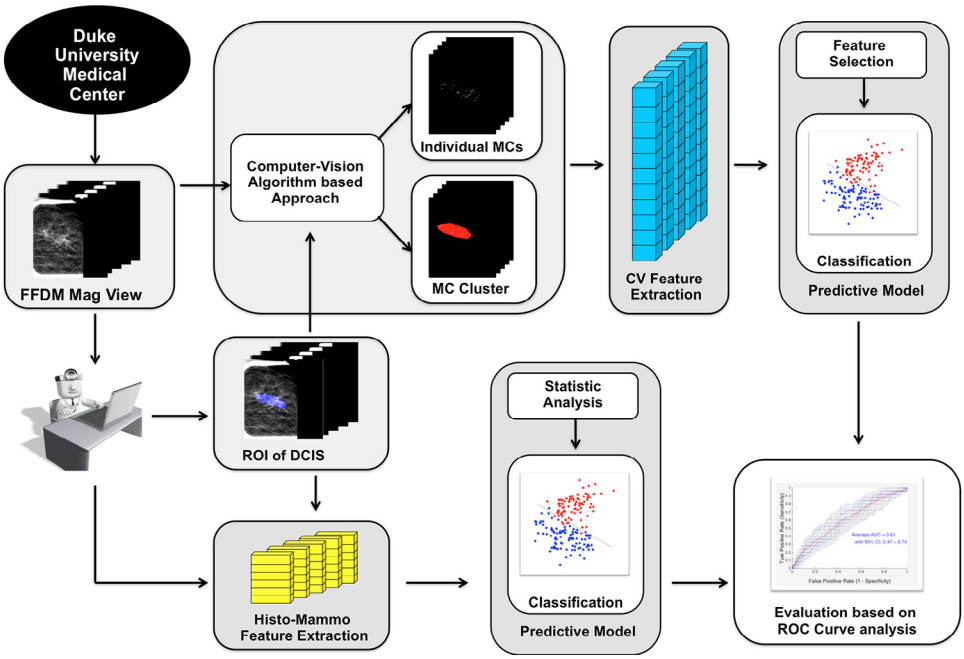

Figure 1. Flowchart of the proposed methodology. CV, computer-vision; DCIS, ductal carcinoma *in situ*; FFDM, full-field digital mammography; MC, microcalcification; ROC, receiver operating characteristic; ROI, region of interest.