Computerized Classification Can Reduce Unnecessary Biopsies in BI-RADS Category 4A Lesions

Leichter Isaac[1,2], Lederman Richard[3], Buchbinder Shalom[4], Srour Yossi[1], Bamberger Philippe[1], and Sperber Fanny[5]

[1] Siemens Computer Aided Diagnosis, PO Box 45202, Jerusalem 91450, Israel
{isaac.leichter, yossi.srour, philippe.bamberger}@siemens.com
[2] Department of Medical Engineering, Jerusalem College of Technology, JCT PO Box 16031, Jerusalem 91160, Israel
[3] Department of Radiology, Hadassah University Hospital, P.O. Box 54202, Jerusalem 91540, Israel
rlederman@hadassah.org.il
[4] Department of Radiology, Staten Island University Hospital Staten Island, 475 Seaview, Staten Island, New York 10305-3498
sbuchbinder@siuh.edu
[5] Department of Radiology, Tel Aviv Sourasky Medical Center, 6 Weizman St., Tel Aviv 64239, Israel
sperber@tasmc.health.gov.il

Abstract. The objective of the study was to assess the potential of a CAD device with computer aided classification capabilities to reduce interventional procedures for BI-RADS category 4A lesions. 113 such lesions (17 masses, 96 clusters), forwarded for biopsy (103 benign) were analyzed retrospectively by a CAD device that generated descriptors. The device extracted quantitative features characterizing the lesions by shape, margins, size and distribution. Descriptors taken from the BI-RADS lexicon for the appearance of the lesion were generated based on the values of the quantitative features. A paradigm based on the computer generated descriptors was developed to assist in assigning a level of suspicion. The paradigm deemed malignant, all 10 malignant cases of the study (100% sensitivity) and correctly classified 38 of the 103 benign lesions. The CAD-generated descriptors, thus, eliminated 36.9% of unnecessary biopsies without decreasing the sensitivity.

1 Introduction

Computer Aided Diagnosis (CAD) in mammography [1] has received FDA approval and entered the mainstream of clinical practice. Its exact role, however, has yet to be defined [2,3], and its widespread implementation is hindered by the relatively large number of false marks [4]. Also, the current generation of CAD systems serves only as a second reader, designed to avoid missed lesions [5], without offering the radiologist a second opinion regarding the nature of the finding. As with all screening tests, mammography is subject to a lack of specificity, which leads to further evaluation of suspicious findings [6]. The need for breast biopsy, frequently with benign results

Susan M. Astley et al. (Eds.): IWDM 2006, LNCS 4046, pp. 76 – 83, 2006.

[7,8], has both a financial and psychological cost, which can be cut by increasing the specificity of diagnosis in mammography.

The addition of classification capabilities could potentially improve the efficacy of such systems by calculating the level of suspicion of any finding either detected by the first tier of the system, or considered suspicious by the radiologist. Several machine learning methods based on neural networks and Support Vector Machines have been applied for the classification of mammographic lesions [9,10]. It was found that for microcalcifications, a classifier based on kernel-based methods, such as Support Vector Machines and Kernel Fisher Discriminant, yielded a significantly better performance than neural network [11].

In this study a classification scheme, based on Kernel Fisher Discriminant, is described, and its use is tested in a subdivision of BI-RADS category 4 cases with both benign and malignant pathologies. BI-RADS category 4 includes findings that do not have the classic appearance of malignancy but have a wide range of probability of malignancy. It is the most problematic and subjective category resulting in a high percentage of benign biopsies. Category 4A is a subdivision, which includes findings with the lowest level of suspicion, for which interventional procedures are nevertheless still recommended. It has been shown that the BI-RADS descriptor categories stratify suspicious micro-calcifications appropriately into intermediate and higher probability of malignancy groups [12]. In this study, an attempt was made to further refine which lesions in this BI-RADS category, in fact, should be sent for biopsy, by the use of computerized descriptors reflecting the appearance of the lesions in the mammogram. The descriptors generated by the CAD device are similar to those used by the BI-RADS lexicon and are familiar to the radiologist.

2 Methods and Material

One hundred and nine cases with 113 lesions (17 masses, 96 clusters) were retrospectively culled from the archives of a university-affiliated facility. All the cases had been prospectively assigned BI-RADS 4A and forwarded for stereo-tactic biopsy. The mean age of the patients was 54.1 ± 8.6 years (range 33–72). The Institutional Review Board at the institution approved the use of these cases for the study, and did not require informed consent because the study was retrospective and patient anonymity was strictly enforced in all aspects of the study. Of the 113 BI-RADS category 4A lesions, 15 masses and 88 clusters proved to be benign at pathology.

The mammograms of the 109 cases were digitized at high resolution (600 dpi, 12 bit) by a prototype CAD device developed by Siemens CAD, Israel [13,14] and the digital images were displayed on the computer screen for further analysis. All 113 lesions were analyzed retrospectively by a radiologist using the CAD device with classification capabilities. The radiologist interactively defined an ellipse encompassing the lesion, on the digital image, and activated the classification algorithm.

For mass lesions the CAD device automatically extracted quantitative features that characterized the mass encompassed by the ellipse. These features characterized the masses by their shape, definition of margins and speculation. Speculation was considered to be a structure composed of lines radiating from a centroid, rather than a sawtooth border of a lesion with a distinct margin. Therefore, this analysis could also be