



WHAT CAN WE ACTUALLY SEE USING COMPUTER AIDED DETECTION IN MAMMOGRAPHY?

Martina Džoić Dominković¹, Gordana Ivanac^{2,3}, Niko Radović² and Mislav Čavka^{2,3}

¹Department of Radiology, Orašje General Hospital, Orašje, Bosnia and Herzegovina;

²Department of Diagnostic and Interventional Radiology, Dubrava University Hospital, Zagreb, Croatia;

³University of Zagreb, School of Medicine, Zagreb, Croatia

SUMMARY – The main goal of this study was to compare the results of computer aided detection (CAD) analysis in screening mammography with the results independently obtained by two radiologists for the same samples and to determine the sensitivity and specificity of CAD for breast lesions. A total of 436 mammograms were analyzed with CAD. For each screening mammogram, the changes in breast tissue recognized by CAD were compared to the interpretations of two radiologists. The sensitivity and specificity of CAD for breast lesions were calculated using contingency table. The sensitivity of CAD for all lesions was 54% and specificity 16%. CAD sensitivity for suspicious lesions only was 86%. CAD sensitivity for microcalcifications was 100% and specificity 45%. CAD mainly 'mistook' glandular parenchyma, connective tissue and blood vessels for breast lesions, and blood vessel calcifications and axillary folds for microcalcifications. In this study, we confirmed CAD as an excellent tool for recognizing microcalcifications with 100% sensitivity. However, it should not be used as a stand-alone tool in breast screening mammography due to the high rate of false-positive results.

Key words: *Breast neoplasm; Image interpretation computer-assisted; Mammography; Computer aided detection (CAD); CAD sensitivity; CAD specificity*

Introduction

Breast cancer is the most common cancer in women and also one of the biggest healthcare issues. Early screening with mammography and advancements in therapy have significantly reduced mortality rate for breast cancer¹. In the United States, the mortality rate for breast cancer decreased by 31% in a ten-year period (1997-2007). Similar results have been reported from Sweden². The sensitivity and specificity of screening mammography have been shown to range from 78% to 99%³. According to the European guidelines for quality assurance in breast cancer screening and diagnosis, double reading in breast screening mammography is

recommended, meaning that breast images are independently reviewed by two specially trained breast radiologists. The sensitivity of mammography has increased by 5%-15% with the double reading practice¹.

In general, computer aided detection (CAD) is designed to detect abnormal breast tissue and to help radiologists discover early signs of breast cancer⁴. CAD is a computerized technique that assists radiologists with image interpretation. Although the use of CAD in mammography screening has been established in many countries, the contribution of CAD to breast screening mammography is still debatable due to a high rate of false-positive results⁵. Nevertheless, screening mammography with CAD has shown several advantages as well. These include detection of microcalcifications on mammograms, as well as a decreased rate of false-negative results in mammography.

The aim of this study was to compare the results of CAD reports in breast screening mammography with the results independently obtained by two radiologists

Correspondence to: *Martina Džoić Dominković, PhD*, Department of Radiology, Orašje General Hospital, 3rd Street, 76 270 Orašje, Bosnia and Herzegovina

E-mail: martina.dzoic@gmail.com

Received February 7, 2018, accepted April 12, 2018

(double reading) and to calculate the sensitivity and specificity of CAD for different types of breast lesions. Additionally, we identified lesions that CAD detected as false-positive.

Patients and Methods

Patients aged 50-69 years in Croatia undergo screening mammography every two years. Screening mammograms of 109 women were obtained from the archives of our Department during two-month period. The study was approved by the institutional Ethics Committee. The study was retrospective in design.

Standard mammography images were obtained using a full flat-panel detector mammography scanner (Mammomat Novation DR, Siemens, Erlangen, Germany). Craniocaudal (CC) and mediolateral oblique (MLO) views of the breasts were taken. The routine screening mammograms (2 CC and 2 MLO) from each patient were analyzed using CAD system (Carestream Health, Inc., New York). In total, we analyzed 436 mammography images (109 women x 2 breasts x 2 views) with CAD system following the CAD prompts, which marks masses with asterisks and microcalcifications with triangles. Two well-experienced breast radiologists with more than 10-year practice in mammography, independently reviewed the same mammogram images.

Each radiologist made his report and sorted the findings according to the Breast Imaging Reporting and Data System (BI RADS) category⁶. If the decision on the BI RADS category was not identical after the two reviews, the more experienced radiologist reviewed the image again and decided on the final BI RADS category. If a suspicious lesion was found, standard further analysis was performed to find out the outcome being either a malignant or benign lesion. We marked as benign only those lesions that were benign and without a change during 2-year follow-up according to the patient medical records, and malignant all the lesions that were confirmed as malignant by core biopsy or surgical biopsy.

The radiologists reported every change in breast tissue according to the BI RADS classification, as follows: calcifications (benign and suspicious – clusters of microcalcifications), architectural distortion/asymmetry, and lesions (morphologically suspicious, indeterminate or benign)⁶. Among other findings reported by

the two radiologists, we noted a number of axillary lymph nodes.

We counted the number of suspicious lesions and calcifications marked by CAD and compared them to the values reported by the radiologists for the same mammograms. Using a contingency table, we calculated the sensitivity and specificity of CAD for the suspicious, non-suspicious lesions and calcifications, as well as the positive predictive (PPV) and negative predictive values (NPV) of CAD.

Next, we excluded benign lesions and axillary lymph nodes from the total number of lesions reported by the radiologists to calculate the sensitivity and specificity of CAD for the lesions that were not typically morphologically benign. Furthermore, we excluded indeterminate lesions and architectural distortions/asymmetry from the total number of lesions in the radiologists' reports to calculate the sensitivity and specificity of CAD for the suspicious lesions only.

When CAD reported a false-positive result, we analyzed the images again to see if there was a connection between the CAD prompts for an image and the observations reported by the radiologists for the same mammogram. When no connection between the two reports was established, we tried to identify the structures that CAD marked as suspicious.

Results

The median age of the study women was 61 (interquartile range 54-64) years. In the 109 radiologists' reports, 89 lesions and 63 calcifications were observed. The lesions were classified into the following categories: suspicious, indeterminate and benign lesions, architectural distortion/asymmetry, and axillary lymph nodes. Analyzing 436 breast screening mammograms with CAD, 54% of all lesions reported by the radiologists were marked as suspicious (Table 1).

Ninety-five percent of the calcifications reported by the radiologists were benign and 5% were comprised of clusters of microcalcifications. CAD detected 62% of the total number of calcifications and 100% of microcalcifications reported by the radiologists (Tables 1 and 2).

The sensitivity of CAD for all types of breast lesions (benign, indeterminate and suspicious lesions, architectural distortions/asymmetry, and axillary lymph nodes) reported by the radiologists was 54%

Table 1. Number and percentage of different types of lesions and calcifications reported by CAD and double reading. CAD values were calculated with respect to the total number of lesions reported by the radiologist

| | Radiologists | CAD, n (%) | % of lesions reported by radiologists and marked by CAD |
|---------------------------------------|------------------|------------------|---|
| Suspicious lesions | 7 (8%) | 6 (12%) | 86% |
| Indeterminate lesions | 6 (7%) | 0 (0%) | 0% |
| Benign lesions | 43 (48%) | 22 (45%) | 51% |
| Architectural distortions/Asymmetry | 5 (6%) | 2 (4%) | 40% |
| Axillary lymph nodes | 28 (31%) | 19 (39%) | 68% |
| Total number of lesions | 89 (100%) | 49 (100%) | 54% |
| Benign calcifications | 60 (95%) | 36 (92%) | 60% |
| Microcalcifications | 3 (5%) | 3 (8) | 100% |
| Total number of calcifications | 63 (100%) | 39 (100%) | 62% |

CAD = computer aided detection

Table 2. Sensitivity, specificity, PPV and NPV of CAD for calcifications in breast tissues expressed as number and/or percentage

| | Radiologists | CAD | Sensitivity (CAD) | Specificity (CAD) | PPV (CAD) | NPV (CAD) |
|---------------------------------|--------------|----------|-------------------|-------------------|-----------|-----------|
| Calcifications (benign and MCC) | 63 (100%) | 39 (62%) | 62% | 44% | 44% | 62% |
| MCC | 3 (5%) | 3 (100%) | 100% | 45% | 6% | 100% |

CAD = computer aided detection; MCC = microcalcifications; PPV = positive predictive value; NPV = negative predictive value

Table 3. Sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) of CAD for different types of lesions in breast expressed as number and/or percentage

| | Radiologists | CAD | Sensitivity | Specificity | PPV | NPV |
|--|--------------|----------|-------------|-------------|-----|-----|
| All lesions ^a | 89 (100%) | 49 (54%) | 54% | 16% | 23% | 37% |
| Lesions without ALN ^{b,c} | 69 (100%) | 29 (42%) | 47% | 16% | 19% | 43% |
| Lesions without benign masses and ALN ^d | 18 (100%) | 8 (44%) | 44% | 16% | 6% | 71% |
| Suspicious lesions | 7 (100%) | 6 (86%) | 85% | 16% | 5% | 96% |

CAD = computer aided detection; ALN = axillary lymph nodes; ^aarchitectural distortion/asymmetry, morphologically suspicious, indeterminate or benign lesions, and axillary lymph nodes; ^barchitectural distortion/asymmetry, morphologically suspicious, indeterminate or benign lesions; ^caxillary lymph nodes; ^darchitectural distortion/asymmetry, morphologically suspicious, and indeterminate lesions

Table 4. Number and percentage of cancers detected with mammography and marked by CAD

| | MCC | Suspicious lesions | Indeterminate lesions | Architectural distortion/asymmetry | Total number of BCA |
|-------------------------------|----------|--------------------|-----------------------|------------------------------------|---------------------|
| Detection of BCA on mammogram | 3 (20%) | 7 (47%) | 2 (13%) | 3 (20%) | 15 (100%) |
| CAD | 3 (100%) | 6 (86%) | 0 (0%) | 2 (67%) | 11 (73%) |

CAD = computer aided detection; MCC = microcalcifications; BCA = breast cancer

and specificity 16% (Table 3). When benign lesions and axillary lymph nodes were excluded from the CAD analysis, the sensitivity of CAD for the lesions that did not demonstrate typically benign morphological characteristics (including suspicious lesions, indeterminate lesions, and architectural distortion/asymmetry) was 44% and specificity remained 16%. Finally, the sensitivity of CAD for suspicious lesions only was 86% and specificity 16% again (Table 3).

Fifteen (71%) of 21 lesions reported by the radiologists as BI RADS 3, 4, and 5 were confirmed carcinomas. They were reported as microcalcifications, suspicious lesions, lesions with indeterminate characteristics and architectural distortion or asymmetry. With CAD, 73% of the breast cancer cases were detected (Table 4).

In total, CAD marked 173 suspicious lesions in 436 mammograms. One hundred and thirty-five (78%) of those lesions were identified as some type of breast structures (e.g., real lesions or parts of breast anatomy). Forty-nine (36%) of those structures had already been recognized and reported by the radiologists, and 86 (64%) structures were marked as suspicious only by CAD. These structures included 55 (64%) glandular parenchyma, 9 (10%) connective tissues, 7 (8%) blood vessels, 5 (6%) axillary folds, 3 (3%) skin parts, 3 (3%) pectoral muscles, nipple in two (2%) cases, and benign calcifications also in two (2%) cases. Among 86 structures that were marked suspicious only by CAD, we could not identify 38 (22%) of them.

Analyzing the structures indicated by CAD as microcalcifications, 47 false-positive reports were obtained. Among these alleged microcalcifications, 19 (40%) were identified including blood vessel calcifications (28%), axillary folds (6%), benign lesions (2%), axillary lymph nodes (2%), and Cooper's ligaments (2%).

Discussion

The role of CAD in mammography was explored in several studies^{7,8}. The focus of these studies was to see how much influence CAD had on the reports made by radiologists, and whether this practice was time-efficient when it comes to making definitive decision in screening mammography. The authors also compared performance of single reading with CAD to double reading.

Single reading with CAD has demonstrated a significantly higher breast cancer detection rate compared

to double reading, but the recall rate for CAD readings has also been higher⁹. However, the sensitivity of radiologist reports has increased by 21% with the use of CAD⁷. In a study performed by Samulski *et al.*, the radiologists tried to identify cancers that had not been detected during first examinations. They increased the sensitivity from 25% to 35% with the use of CAD. In addition, the average time spent on one examination was not significantly higher when using CAD system⁸.

A study performed by Fenton *et al.* showed lower specificity in mammography screening when using CAD. Additionally, detection rate of breast cancers was not improved with CAD¹⁰. Another study also showed that screening performance was not improved with CAD on any metric assessed. There was no difference in cancer detection rate (4.1 in 1000 women screened with and without CAD). CAD did not improve intra-radiologist performance. Sensitivity was significantly decreased for mammograms interpreted with *vs.* without CAD in a subset of radiologists who interpreted mammograms both with and without CAD¹¹. On the contrary, another study suggests that single reading with CAD can be comparable to double reading in terms of efficiency and that the breast cancer detection rate is improved with CAD¹². In addition, Park *et al.* demonstrated that the breast cancer detection rate was not significantly different between symptomatic and asymptomatic groups of patients. They also concluded that the sensitivity of CAD was highest in fatty breasts in both groups of patients. In most of the cases, CAD detected invasive ductal carcinoma not otherwise specified (IDC-NOS) and ductal carcinoma *in situ* (DCIS); however, in this study, other types of breast cancer were not detected using CAD¹³. In general, CAD is not designed to detect benign lesions and should not be used as a stand-alone method in mammography screening. Nevertheless, we wanted to investigate the role of CAD as a 'third radiologist'. We also wanted to explore the type of lesions that are recognized by CAD, as well as the lesions that are not detected when using CAD.

According to our results, CAD recognized 54% of all masses reported by the radiologists. Next, CAD recognized 86% of suspicious masses, 51% of benign lesions, 40% of architectural distortions/asymmetry, and none of indeterminate masses.

Architectural distortion is "a distortion of the normal architecture with no definite mass visible"¹⁴. It is

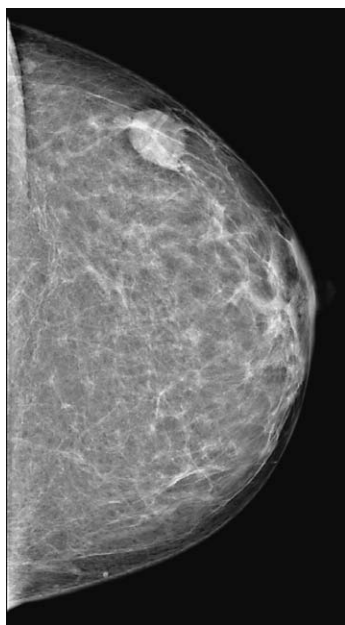


Fig. 1. Example of a mass that was not recognized by computer aided detection. The figure shows a regularly shaped mass that has sharp margins on one side and indistinct boundaries on the other side.

the third most common manifestation of breast cancer on mammography. Malignant lesions (IDC, invasive lobular carcinoma (ILC), DCIS) and benign lesions (surgical scars, radial scars, complex sclerosing lesions, fat necrosis, and intralobular fibrosis) can present as architectural distortions. This is not the first case where CAD was not able to make distinction between architectural distortions. In a study performed by Baker *et al.*, less than half of the architectural distortions were detected with CAD¹⁴. Radiologists should be aware of this problem when using CAD in mammography.

In another study, CAD showed 91% sensitivity for detecting breast cancer using mammography. In addition, 100% sensitivity for microcalcifications was noticed¹⁵. These results are in correspondence with our results. In total, CAD recognized 73% of breast cancers and 100% of the structures marked as microcalcifications.

In our study, the sensitivity of CAD increased to 86% when we excluded typically benign lesions, indeterminate lesions, and architectural distortion or asymmetry. This is due to the fact that CAD can only mark a suspicious mass in general, but it cannot specifically recognize benign lesions in a breast nor recognize indeterminate lesions that are not completely defined. For that type of distortions/masses, we need a well-trained eye of a radiologist who knows how to make distinction between these structures (Fig. 1).

The specificity remained the same (16%). This value is not high, which is not surprising considering the high rate of false-positive results with CAD. When analyzing the additional structures that were marked by CAD as suspicious, we noticed different breast changes and parts of breast anatomy. In most of the cases, CAD 'mistook' the lesions for glandular parenchyma, connective tissue and blood vessels, and microcalcifications for blood vessel calcifications and axillary folds.

In conclusion, we can confirm that CAD cannot be used as a stand-alone tool in screening mammography and it cannot replace a third radiologist on mammography examination. In some instances, such as detecting microcalcifications, CAD can be helpful as an additional tool, but given the high rate of false-positive results with CAD, caution is strongly recommended when interpreting the results.

References

1. Perry N, Broeders M, De Wolf C N, Törnberg S, Holland R, Von Karsa L. European guidelines for quality assurance in breast cancer screening and diagnosis. 4th edition. *Ann Oncol.* 2008;19(4):614-22. doi: 10.1093/annonc/mdm481
2. Brkljacic B, Miletic D, Sardanelli F. Thermography is not a feasible method for breast cancer screening. *Coll Antropol.* 2013;2:589-93.
3. Kolb TN, Lichy J, Newhouse JH. Comparison of the performance of screening mammography, physical examination, and breast US and evaluation of factors that influence them: an analysis of 27,825 patient evaluations. *Radiology.* 2002;225(1):165-75. doi: 10.1148/radiol.2251011667
4. Dromain C, Boyer B, Ferré R, Canale S, Delalogue S, Balleyguier C. Computer-aided diagnosis (CAD) in the detection of breast cancer. *Eur J Radiol.* 2013;82(3):417-23. doi: 10.1016/j.ejrad.2012.03.005
5. Azavedo E, Zackrisson S, Mejare I, Heibert Arnlin M. Is single reading with computer-aided detection (CAD) as good as double reading in mammography screening? A systematic review. *BMC Med Imaging.* 2012;12:22. doi: 10.1186/1471-2342-12-22
6. D'Orsi CJ, Sickles EA, Mendelson EB, Morris EA. *Breast Imaging Reporting and Data System: ACR BI-RADS Breast Imaging Atlas.* 5th edition. Reston, VA, USA: American College of Radiology, 2013.
7. Brem RF, Baum J, Lechner M, Kaplan S, Souders S, Naul LG, *et al.* Improvement in sensitivity of screening mammography with computer-aided detection: a multi-institutional trial. *AJR Am J Roentgenol.* 2003;181:687-93. doi: 10.2214/ajr.181.3.1810687.

8. Samulski M, Hupse R, Boetes C, Mus RDM, Den Heeten GJ, Karssemeijer N. Using computer-aided detection in mammography as a decision support. *Eur Radiol.* 2010;20:2323-30. doi: 10.1007/s00330-010-1821-8
9. Gilbert FJ, Astley SM, McGee MA, Gillan MG, Boggis CR, Griffiths PM, *et al.* Single reading with computer-aided detection and double reading of screening mammograms in the United Kingdom National Breast Screening Program. *Radiology.* 2006;241(1):47-53. doi: 10.1148/radiol.2411051092
10. Fenton JJ, Abraham L, Taplin SH, Geller BM, Carney PA, D'Orsi C, *et al.* Effectiveness of computer-aided detection in community mammography practice. *J Natl Cancer Inst.* 2011;103:1152-61. doi: 10.1093/jnci/djr206
11. Lehman CD, Wellman RD, Buist DSM, Kerlikowske K, Tosteson ANA, Miglioretti DL. Diagnostic accuracy of digital screening mammography with and without computer-aided detection. *JAMA Intern Med.* 2015;175(11):1828-37. doi: 10.1001/jamainternmed.2015.5231
12. Gilbert FJ, Astley SM, Gillan MG, Agbaje OF, Wallis MG, James J, *et al.* Single reading with computer-aided detection for screening mammography. *N Engl J Med.* 2008;359(16):1675-84. doi: 10.1056/NEJMoa0803545
13. Park CS, Jung NY, Kim K, Jung HS, Sohn KM, Oh SJ. Detection of breast cancer in asymptomatic and symptomatic groups using computer-aided detection with full-field digital mammography. *J Breast Cancer.* 2013;16(3):322-8. doi: 10.4048/jbc.2013.16.3.322
14. Baker JA, Rosen EL, Lo JY, Canalea S, Delalogue S, Balleyguier C. Computer-aided detection (CAD) in screening mammography: sensitivity of commercial CAD systems for detecting architectural distortion. *AJR Am J Roentgenol.* 2003;181:1083-8.
15. Murakami R, Kumita S, Hitomi T, Yoshida T, Sugizaki K, Kuwako T, *et al.* Detection of breast cancer with a computer-aided detection applied to full-field digital mammography. *J Digit Imaging.* 2013 Aug;26(4):768-73. doi: 10.1007/s10278-012-9564-5

Sažetak

ŠTO ZAPRAVO MOŽEMO VIDJETI PRIMJENOM RAČUNALNO POTPOMOŠNE ANALIZE KOD MAMOGRAFIJE?

M. Džoić Dominković, G. Ivanac, N. Radović i M. Čavka

Svrha ovoga istraživanja bila je usporediti rezultate računalno potpomognute analize (*computer aided detection*, CAD) u probirnoj mamografiji s rezultatima analize dva neovisna radiologa te utvrditi osjetljivost i specifičnost CAD-a za lezije u dojkama. Analizirali smo 436 mamograma pomoću CAD-a i usporedili rezultate s interpretacijom dva neovisna radiologa. Izračunali smo osjetljivost i specifičnost CAD-a za lezije u dojkama putem tablica kontingencije. Osjetljivost CAD-a za otkrivanje svih lezija u dojkama iznosila je 54%, a specifičnost 16%. Osjetljivost CAD-a za sumnjive lezije bila je 86%, a za mikrokalcifikacije 100% uz specifičnost od 45%. CAD je uglavnom pogrešno interpretirao žljezdani parenhim, vezivno tkivo i krvne žile kao tvorbe u dojkama, dok je kalcifikacije u krvnim žilama i aksilarni nabor miješao s mikrokalcifikacijama. Ovom studijom smo potvrdili da je CAD izvrstan alat za otkrivanje mikrokalcifikacija s osjetljivošću od 100%. No, ipak se ne bi trebao rabiti kao jedina metoda u probirnoj mamografiji dojki uzimajući u obzir količinu lažno pozitivnih rezultata koja je prilično visoka.

Ključne riječi: *Karcinom dojke; Računalno potpomognuta analiza slika; Mamografija; Računalno potpomognuta analiza (CAD); Osjetljivost CAD; Specifičnost CAD*