Comparison of Two Commercial CAD Systems for Digital Mammography

Stephanie Leon,¹ Libby Brateman,² Janice Honeyman-Buck,² and Julia Marshall²

BACKGROUND

In late 2006, an outpatient imaging center began its conversion to digital mammography (DM). Part of this transition was a plan to purchase software for computer-aided detection (CAD) to assist in analysis of the digital mammography images. The preparation to purchase one of two systems included a comparison of several specifications, including DICOM compatibility and the ability of the systems to fit well into the digital mammography program of the imaging center.

A small study was designed to determine whether one of the two different commercially available systems being considered [R2 Image-Checker (version 8.3.17) and iCAD Second Look (version 7.2-H)] was superior to the other for assisting in interpretation of digital mammography images from screening mammograms obtained with the newly installed General Electric Senographe DS unit. Patient images included in this study were given retrospective approval from the Institutional Review Board with waiver of informed consent for publication of the study results.

MATERIALS AND METHODS

The iCAD system was available at the facility for a portion of 2 days and was operated on the default “high” sensitivity setting as originally installed. (Other possible software choices were low and medium sensitivity.) Twenty-seven female patients received standard two-view digital screening mammograms (20 bilateral and seven unilateral) during this period. Only non-displaced breast implant

views were excluded. These 27 screening mammograms all happened to be diagnosed as benign. All “for processing” and “for presentation” images were stored on the PACS system belonging to the Department of Radiology. All 94 “for processing” digital mammography images were analyzed with the iCAD system during the 2-day period, and the CAD overlay information was stored separately from the digital mammograms in the PACS.

Because the mammography workstation could not show two different sets of CAD markings for a single patient exam, the patient study name and medical record number for the exam were changed to have a subtle difference from the original for each CAD assessment. In this way, each study was given a new Study-Instance-UID, and therefore, each study appeared to be a different patient to the workstation. No change was made to the image data.

These new “for processing” images were reanalyzed at a later time by the R2 system during the period in which it was installed, and its CAD overlay output was also stored separately. The R2 system did not have a sensitivity setting option. After both sets of CAD data were obtained and stored, the CAD

¹From the Department of Nuclear and Radiological Engineering, University of Florida, Gainesville, FL USA.

²From the Department of Radiology, University of Florida College of Medicine, Box 100374, Gainesville, FL 32610-0374, USA.

Correspondence to: Libby Brateman, Department of Radiology, University of Florida College of Medicine, Box 100374, Gainesville, FL 32610-0374, USA; tel: +1-352-2650291; fax: +1-352-2650279; e-mail: bratel@radiology.ufl.edu

Copyright © 2008 by Society for Imaging Informatics in Medicine

Online publication 13 August 2008

doi: 10.1007/s10278-008-9144-x

Table 1. Number of CAD Markers in Images

View	All markers		Masses		Calcifications	
	R2	iCAD	R2	iCAD	R2	iCAD
CC	25	26	12	21	13	5
MLO	27	23	15	19	12	4
Both views	52	49	27	40	25	9

markers for both systems were reviewed on a case-by-case basis at the digital mammography workstation used for mammography interpretation at the imaging center. This General Electric SenoAdvantage 2.1 workstation employed two 5 MP CRT monitors.

All comparisons of CAD markings were made subsequent to installation of the R2 system, which was also after the original diagnoses for this limited number of patients who had received their screening mammograms while the iCAD system was briefly installed. The CAD markings were assessed by the lead interpreting radiologist for mammography, the Board-certified director of mammography for the imaging center and for mammography at Shands Hospital at the University of Florida, in charge of resident teaching in mammography, and a Mammography Quality Standards Act-qualified mammographer with clinical experience of interpreting more than 1,000 mammograms per year for 19 years.

During the short duration of the iCAD installation, it happened that all 47 breasts were designated as BI-RADS® Category 2 by the interpreting radiologist. Therefore, all CAD markings indicated by either of the CAD systems were assumed to be markers associated with false-positive findings. Data used in the analysis for the two CAD systems included the total number of CAD markers, the number of masses marked, and the number of calcifications marked, independent of location. The data also included the number of masses and calcifications that the two systems identified in common (i.e., the same type of marker in the same location). Each of these categories was subdivided into craniocaudal (CC) and mediolateral oblique (MLO) views.

Finally, the lead interpreting mammography radiologist reviewed the markings from both systems and determined, for each image, whether one system was preferable to the other, in terms of the CAD markers presented, and which system had markings that appeared to be more intuitively logical. Because all CAD markers were considered to be false positives, this assessment was based on the anatomical features displayed in the mammography image,

such as whether there appeared to be some basis for the presence of each marker in the displayed image, or whether the marker location was spurious. Another criterion was whether one system displayed many more markers than the other in the image. Although it was clear that markers for calcifications were related to the presence of calcifications, the significance of the calcifications that were marked in terms of likelihood of malignancy was taken into account. It is interesting that none of the markers for masses appeared to be indicative of definable masses. There was no difference in the ease of use with the workstation.

RESULTS

The number of markers for each CAD system was tabulated by marker type (i.e., for calcifications or masses) and by mammographic projection (i.e., CC or MLO) for the 27 patients. When the markers were pooled for all marker types, all breasts and both views, CAD markers were absent in 12 out of 27 patient exams (44%) by R2 and in 11 patient exams (41%) by iCAD. When they were separated by marker type (calcifications or masses), markers for masses were absent in 12 patient exams by both systems; however, markers for calcifications were absent in 24 of 27 (89%) patient exams by iCAD compared with 14 (52%) patient exams by R2. At the other extreme, one patient had seven calcifications marked by both systems.

Table 1 presents the total number of CAD markers by projection and by marker type for the 2 systems. Differences existed in the number of CAD markers by CAD system in terms of calcifications and masses, with each system reasonably consistent between views. These differences disappeared when the total number of markers was pooled for each system.

The iCAD system marked more masses in 12 of the 27 exams (44%), compared with five (18%) for R2; however, the R2 system marked more calcifica-

Table 2. Average Number of CAD Markers Per Image

View	All markers		Masses		Calcifications	
	R2	iCAD	R2	iCAD	R2	iCAD
CC	0.53	0.55	0.26	0.45	0.28	0.11
MLO	0.57	0.49	0.32	0.40	0.26	0.09
Both views	0.55	0.52	0.29	0.43	0.27	0.10

tions in 11 of the 27 exams (41%) than iCAD, compared with 0 for iCAD. When the markers were pooled, the systems seemed more similar once again, with R2 markings more prevalent in 11 of the exams, and iCAD markings more prevalent in eight out of the total 27.

The total number of markers in each category was divided by the number of images (47 CC, 47 MLO, and 94 total) to obtain the average number of markers per image, in order to compare the results of this study with others. These data are summarized in Table 2.

Independent of location, the total number of markers was roughly equal for both systems, with the R2 system displaying an average of 0.55 markers per image and the iCAD system displaying 0.52 markers per image. The CAD markers generated by R2 were split evenly between masses (0.29/image) and calcifications (0.27/image), whereas the iCAD markers were weighted heavily toward masses (0.43/image) as opposed to calcifications (0.10/image).

It is interesting to note that in only a small fraction of the images did the two systems mark the same false-positive features: only 30 markers (representing 15 identified features) out of 101 total (29.7%) were shared between the two systems. The radiologist's assessments of which CAD system's markings were more logical were split almost evenly between R2 (10 votes), iCAD (9 votes), and "Neither" (eight votes). The differences were not significant ($P=0.896$, using χ^2 test for goodness of fit).

DISCUSSION

A 2007 study by Yang et al. found a mean number of 0.45 false-positive markers per image using only normal screening mammograms¹, which agrees well with the values of 0.55 and 0.52 markers per image (R2 and iCAD, respectively) found in this study. Another 2007 study using screening mammograms found that the R2 ImageChecker system produced 1.08 false positive markers per image and the iCAD Second Look system produced 1.41 false positives per image²; however, this study used digitized screen-film mammograms, so the results may not be directly comparable to those obtained using digital mammograms. Other recent studies from 2002 to 2006 have indicated a range of 0.27 to 0.61 false positives per image using CAD on digital mammograms^{3,4,5}, but these

studies were performed using diagnostic mammograms with abnormal findings.

The lead interpreting mammography radiologist was shown all CAD-marked images, but she retained her professional opinion that all of the marked features were benign. The BI-RADS[®] classification was not changed for any image after review of the CAD data; therefore, in this small study, use of CAD did not influence the rate of patient recall rates or biopsies.

The fact that the two systems rarely marked the same areas as suspicious lends support to the radiologist's assessment that the findings were all benign. Also, 14 of the 27 patients had an additional screening mammogram performed 12–14 months after the initial mammogram, and all of the recent mammograms were also negative or benign for these patients.

CONCLUSION

There were no important differences between the two systems, either in terms of the number of false positives or in the placement of the markers in appropriate locations, when the data were pooled as to the total number of CAD markings. However, the R2 system consistently marked fewer false-positive masses than the iCAD system, and the iCAD system consistently marked fewer false-positive calcifications than the R2 system. Given the similarities in findings between the two systems, the decision as to which CAD system to purchase was based on factors other than outcome or reader preference.

REFERENCES

1. Yang SK, Moon WK, Cho N, et al: Screening mammography-detected cancers: sensitivity of a computer-aided detection system applied to full-field digital mammograms. *Radiology* 244:104–111, 2007
2. Ellis RL, Meade AA, Mathiason MA, Willison KM, Logan-Young W: Evaluation of computer-aided detection systems in the detection of small invasive breast carcinoma. *Radiology* 245:88–94, 2007
3. Kim SJ, Moon WK, Cho N, Cha JH, Kim SM, Im JG: Computer-aided detection in full-field digital mammography: sensitivity and reproducibility in serial examinations. *Radiology* 246:71–80, 2008
4. Obenauer S, Sohns C, Werner C, Grabbe E: Computer-aided detection in full-field digital mammography: detection in dependence of the BI-RADS categories. *Breast J* 12:16–19, 2006
5. Baum F, Fischer U, Obenauer S, Grabbe E: Computer-aided detection in direct digital full-field mammography: initial results. *Eur Radiol* 12:3015–3017, 2002