J HK Coll Radiol 2003;6:195-200

ORIGINAL ARTICLE

Comparison of Spot Digital and Conventional Mammography in the Evaluation of Microcalcifications

KKT Pak,[1] JLF Chiu,[1] CS Kwok,[2] SC Wong,[1] SY Yiu,[1] SCH Chan[1]

[1]Department of Radiology and Imaging, Queen Elizabeth Hospital, and [2]Department of Optometry and Radiography, The Hong Kong Polytechnic University, Hong Kong

ABSTRACT

Objective: *To compare the diagnostic accuracy of spot digital and conventional screen-film mammography and the influence of clinical data on the characterisation of microcalcifications.*

Patients and Methods: *Twenty nine patients with 30 groups of microcalcifications were studied. Two radiologists assessed the biopsy-proven calcifications on conventional screen-film cone-magnification views and digital spot views. Ratings on breast density, visibility, and likelihood of malignancy were estimated on a 4-point confidence scale individually by each radiologist. Both radiologists were asked to give a second malignancy rating for conventional screen-film mammograms with patients' clinical data. Receiver operating characteristic methodology was applied to evaluate the results. The classification accuracy was quantified by area under the receiver operating characteristic curve. Statistical differences in the area under the receiver operating characteristic curve values for the effects of digital mammography and clinical data were estimated.*

Results: *The average area under the receiver operating characteristic curve values were higher for evaluation of microcalcifications on conventional screen-film images. Prior knowledge of patient's clinical data did not improve the diagnostic accuracy on conventional screen-film mammography. But the differences in area under the receiver operating characteristic curve values for both comparisons were not statistically significant ($p > 0.05$).*

Conclusion: *The diagnostic accuracy of spot digital is comparable to that of cone magnification views on conventional screen-film mammography. The effects of intra- and inter-reader variability should be considered. Consistency in malignancy rating based on lesion characterisation may be affected by variation in interventional threshold among the readers.*

Key Words: *Mammography, Microcalcifications, Receiver operating characteristics*

INTRODUCTION

Breast cancer is the third leading cause of cancer deaths among women in Hong Kong. The age-standardised incidence rate increased from 30.9 to 34.6 per 100,000 standard female population in 1983 to 1993 to 35.0 to 40.8 during 1994 to 1999, with the largest significant increase observed among women aged 50 to 59 years. The age-standardised incidence rate and mortality of female breast cancer in Hong Kong are approximately 40% to 60% and 35% to 50% of those in western countries, respectively.[1]

Conventional screen-film mammography (CSFM) has been widely accepted as a screening modality for the past 20 to 30 years in western countries. However, there is insufficient evidence to justify population-based breast cancer screening by mammography for women in Hong Kong and other Asian populations with low breast cancer prevalence.[2] With the development of digital mammography in recent years, the limitations of traditional mammography can be mitigated as each part of the breast imaging chain, including image acquisition, image storage, and display, are separated and optimised. Furthermore, digital processing allows the manipulation of image contrast, which improves lesion conspicuity.[3]

Correspondence: *Dr KKT Pak, Department of Radiology and Imaging, Queen Elizabeth Hospital, 30 Gascoigne Road, Jordan, Kowloon, Hong Kong.*
Tel: (852) 2958 5662; Fax: (852) 2958 6048;
E-mail: karen_pak@hotmail.com

Submitted: 3 July 2003; Accepted: 30 November 2003.

There are ongoing clinical trials and testing for all types of digital mammographic equipment in the USA, which include 2 major clinical trials funded by non-industrial sources — the International Digital Mammography Development Group Digital Mammography Pilot Study, and the Department of Defense Full-field Digital Mammography Screening Trial. However, the initial Food and Drug Administration (FDA) guidelines for manufacturers of digital mammography systems to obtain FDA approval were flawed since the level of agreement required for digital mammography and screen-film mammography was not attainable because of intra- and inter-reader variability. Consequently, the FDA revised its guidance document in February 1999, stating that digital mammographic trials for FDA approval must be based on the truth regarding breast cancer status and not on the direct agreement with findings at screen-film mammography. Methods such as receiver operating characteristic (ROC) analysis must be used to compare the performance of this new technology to screen-film mammography.[4] At the time of writing, there were already 4 FDA-approved full field digital mammography systems available in the market, and accredited by the American College of Radiology (ACR).[5]

ROC analysis is an efficient way to display the relationship between sensitivity and specificity for tests that have continuous outcome. The area under the curve (AUC) measures the ability of the test to correctly classify those with and without the disease.[6] Several studies have thus been conducted worldwide to compare the speed and accuracy of conventional screen-film and digital mammography using the ROC analysis.[7-12]

In this study, we used ROC analysis to compare the diagnostic accuracy of radiologists in the evaluation of microcalcifications viewed on digital spot images and cone-magnification views on conventional mammography. As one of many factors that might affect the decision-making process, the influence of pertinent clinical data (patient's age, past and family history of breast cancer, prior imaging findings) was also investigated.

PATIENTS AND METHODS

Patient Selection

Twenty nine patients with 30 sets of conventional screen-film mammograms performed at the Queen Elizabeth Hospital were included in this study. All mammograms contained indeterminate or suspicious microcalcifications in which percutaneous stereotactic large core biopsies were recommended for histological diagnosis using 14 gauge needles. All biopsies were performed on the digital biopsy and spot imaging system in the Radiography Clinic of the Hong Kong Polytechnic University. Retrieval of calcifications was documented on either specimen radiographs or pathological examination.

Equipment

Conventional screen-film cone-magnification views on cranio-caudal and medio-lateral/lateral projections were performed on Senographe DMR (General Electric Medical Systems, Milwaukee, USA). Digital spot images were obtained using the digital biopsy and spot imaging system Opdima®, which was an add-on upright stereotactic unit on the Mammomat 3000 (Siemens Medical Systems, Erlangen, Germany) at the Hong Kong Polytechnic University. The digital images were acquired with a CCD camera with pixel depth of 12 bits, matrix size 1024 x 1792 pixels and a minimum spatial resolution of 10 lp/mm. The digital images were displayed and processed on a MagicView 1000 workstation with high resolution (2048 x 2560 pixels) and high luminance monitors. The system provided roam-and-zoom functions with the click of a mouse and variable windowing setting. Reading of both spot screen-film and digital soft copy images were performed in a dark environment which was suitable for interpretation of mammograms. The spot screen-film images were hung on a multipanel mammographic light box with suitable masking of extraneous light. The screen-film images were examined with a magnifying glass and strong light where necessary.

Study Design

Two radiologists with similar mammographic interpretation experience (approximately 5 years) were asked to evaluate the microcalcifications present on the cone-magnification views obtained by CSFM and the pre-biopsy digital spot views taken not more than 2 weeks apart. Image quality was assessed on a 4-point rating scale, as follows: 1 = fatty, 2 = scattered fibroglandular, 3 = heterogeneously dense, and 4 = extremely dense, for breast density; and 1 = very poor, 2 = poor, 3 = good, and 4 = very good, for visibility of calcifications. The likelihood of malignancy for both sets of images were indicated by each radiologist on a 4-point confidence rating scale, as follows: 1 = benign, 2 = probably benign, 3 = suspicious, and 4 = highly suspicious based on the criteria of the Breast Imaging Reporting and Data System (BI-RADS). Both radiologists were asked to

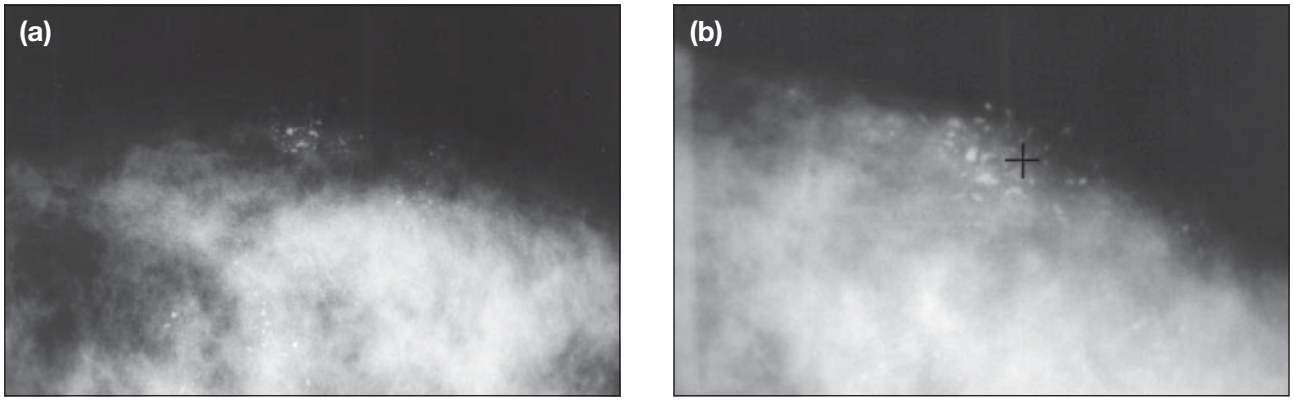

Figure 1. Suspicious clustered pleomorphic and branching calcifications on (a) spot screen-film image and (b) digital spot view. Biopsy revealed high-grade ductal carcinoma in situ.

give a second rating for the likelihood of malignancy with knowledge of patient's clinical data — age, family history or known history of breast cancer, clinical symptoms, and relevant prior mammogram findings for CSFM.

Data Analysis

Histological diagnosis was used as the gold standard. Benign conditions were confirmed by lesion stability on follow-up mammogram performed after 6 months. The confidence ratings of each radiologist obtained from reading CFSM and digital spot images were analysed using ROC methodology (ROCKIT program available through the Internet from Metz CE, The University of Chicago, Chicago, USA).[13] The diagnostic accuracy was quantified by area under the ROC curve, AUC. The Dorfman-Berbaum-Metz method using the LABMRMC 1.0B software (Beta version 3) was applied to calculate the statistical significance of difference between the averaged areas under the ROC curves that are estimated for the 2 diagnostic tests using 'jackknifing' and analysis of variance (ANOVA) methods.[14] It was also determined whether there was significant inter-reader variability in confidence ratings with and without clinical data for each reader.

RESULTS

Overall Results

Thirty mammograms from 29 patients were reviewed in this study. The mean age of the women was 47 years. Stereotactic core biopsies of the microcalcifications were performed between September 2000 and June 2002. Microcalcifications were confirmed at specimen radiography or biopsy pathology in all cases. Seven malignant or suspicious lesions were diagnosed on core biopsy in the 30 cases (23.3%). Five cases of ductal carcinoma in situ (intermediate to high grade) were diagnosed on core biopsy, which were confirmed on subsequent surgical excision (Figures 1a and 1b). There was 1 mucocele-like lesion, which revealed ductal carcinoma in situ at excision with hookwire localisation. Another biopsy revealed atypical apocrine changes, with high grade ductal carcinoma in situ confirmed after surgery (Table 1).

The rest of the lesions had a benign diagnosis or showed no evidence of malignancy on core biopsy. Stability of

Table 1. Patients' ages and biopsy results.

Age (years)	Biopsy results
49	No malignancy
55	Fibrocystic disease
43	Sclerosing adenosis
47	Proliferative fibrocystic disease
47	Atypical lobular hyperplasia with pagetoid spread
48	High-grade ductal carcinoma in situ
45	Fibrocystic disease
43	Intermediate-grade ductal carcinoma in situ
41	Sclerosing adenosis
46	Moderate- to high-grade ductal carcinoma in situ
47	No malignancy
53	No malignancy
60	Fibrocystic disease
52	Benign
	Benign
38	Mucocele-like lesion (excision revealed high-grade ductal carcinoma in situ)
38	High-grade ductal carcinoma in situ
47	No malignancy
50	No malignancy
68	Intermediate-grade ductal carcinoma in situ
49	Non-proliferative fibrocystic disease
47	No malignancy
41	Fibrocystic disease
53	Atypical apocrine changes (excision revealed high-grade ductal carcinoma in situ)
45	Proliferative fibrocystic changes
48	Proliferative fibrocystic changes
26	No malignancy
48	No malignancy
49	No malignancy
37	No malignancy

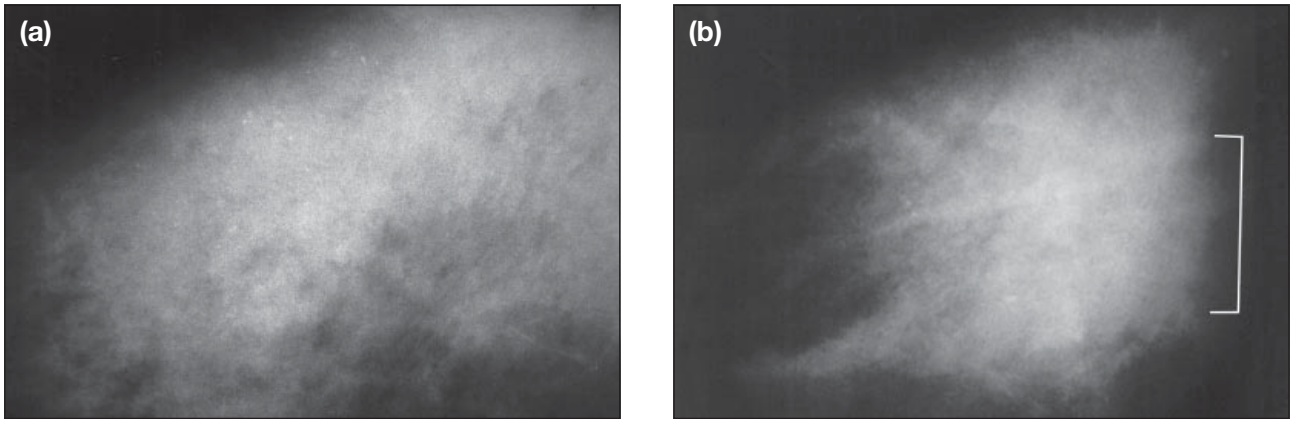

Figure 2. Extremely dense breast with segmental faint indeterminate clustered microcalcifications on (a) spot screen-film image and (b) digital spot view. Biopsy revealed fibrocystic changes with focal moderate epitheliosis and adenosis.

Table 2. Areas under the receiver operating characteristic curves for evaluation of microcalcifications on conventional screen-film mammography (CSFM) and digital spot mammography (DSM), with and without clinical data.

Radiologist	Area under the curve	
	CSFM	DSM
A + B	0.830 ± 0.065	0.741 ± 0.087
	With clinical data	Without clinical data
A	0.676 ± 0.097	0.789 ± 0.096
B	0.830 ± 0.103	0.869 ± 0.084

All values are expressed as mean ± standard error.

the lesions were confirmed on follow-up mammogram performed 6 months later.

Ratings on Image Quality

Subjective ratings by both readers for breast density and visibility of calcifications on CSFM and digital spot mammography (DSM) showed that most of the patients in this group had scattered fibroglandular to extremely dense breasts (1% fatty, 56% scattered fibroglandular, 35% heterogeneously dense, and 8% extremely dense) [Figures 2a and 2b], which agreed with the reported ethnic differences in mammographic densities comparing Caucasian and Asian women.[15] Visibility rankings were more favourable for digital spot images, with 70% of cases being considered 'good to very good' on DSM compared with 56.7% for CSFM.

Receiver Operating Characteristic Analysis

The AUC of the ROC analysis and the corresponding ROC curves were shown in Table 2 and Figures 3 and 4, respectively. Higher AUC values were observed by both radiologists when assessing microcalcifications on CSFM than on digital mammography, but the mean difference was not statistically significant (p = 0.04; 95% confidence interval, -0.12, 0.23). Prior knowledge of the patient's clinical data did not seem to improve the diagnostic accuracy, but the mean difference in AUC values with and without clinical data was statistically insignificant (p = 0.07; 95% confidence interval, -0.06, 0.21).

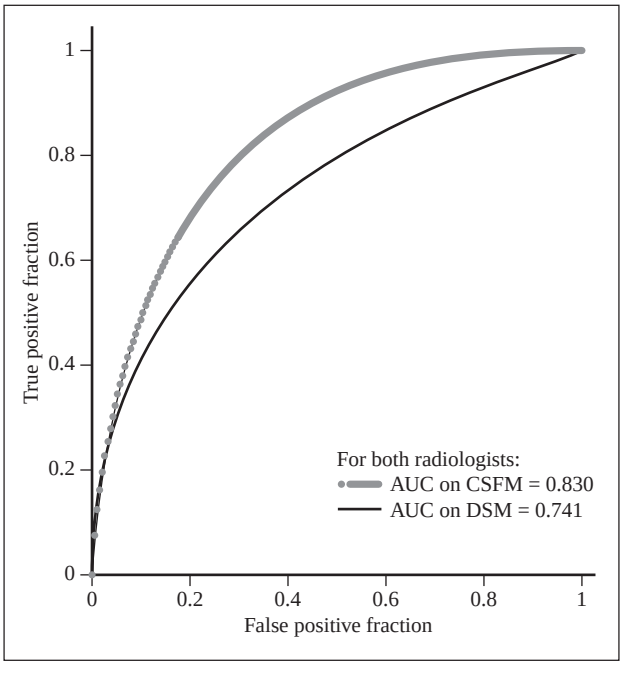

Figure 3. Receiver operating characteristic curves and areas under the curves (AUC) for both radiologists on conventional screen-film mammography (CSFM) and digital spot mammography (DSM).

DISCUSSION

Approximately 10% to 20% of palpable breast cancers are not visible on traditional screen-film mammography due to insufficient contrast between normal and abnormal breast tissue. As the mammographic features of early-stage breast cancers are not very specific, the need for high sensitivity leads to biopsy of many low-suspicion lesions. The positive predictive values of mammographic signs are often below 30%.[16] The limitations of

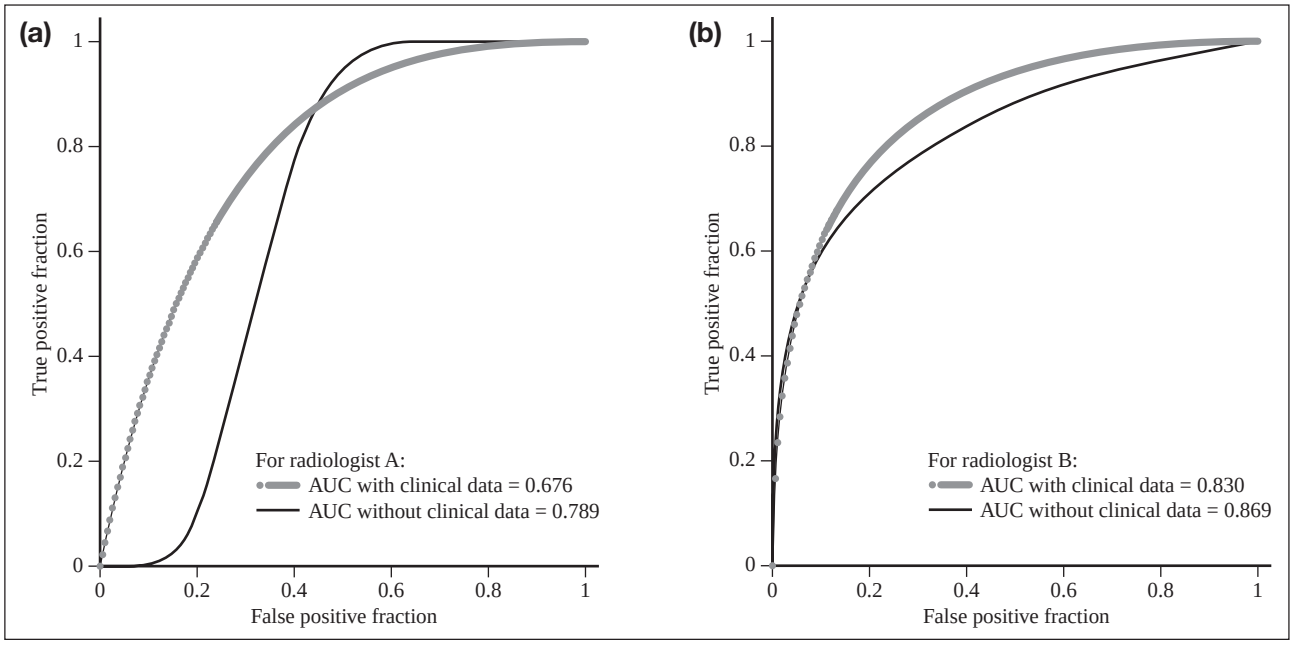

Figure 4. Receiver operating characteristic curves and areas under the curves (AUC) for 2 individual radiologists (a and b) on conventional screen-film with and without clinical data.

mammography in detection of breast cancer in Asian women are further exacerbated by relatively dense parenchymal pattern. Sensitivity, specificity, and AUC performance for populations of patients with dense parenchymal patterns have been documented as being lower than those for a more general population of patients.[17] Digital mammography, however, offers potential advantages over the standard use of radiographic film with its large dynamic range that allows examination of all areas of the breast, despite the varying density. There are thus great enthusiasm and hope for improved detection of early breast cancer by digital mammography with its enhanced contrast resolution, optimised image processing and display, as well as digital storage and transfer.

The average AUC values for the ROC curves generated from CSFM and DSM in this study are 0.83 and 0.74, respectively. The reported average AUC of radiologists based on BI-RADS categories are in the range of 0.74 to 0.95.[18] Digital mammography enhanced the visibility of lesions for dense breasts in our study, as indicated by the higher visibility ratings obtained on DSM (70.0%) compared with CSFM (56.5%). However, it is interesting to note that improved lesion detection did not result in better diagnostic accuracy. We attempt to explain this discrepancy by study design and intra- and inter-reader variability in the characterisation of microcalcifications.

Firstly, the viewing of soft copy display of DSM images was relatively new to both readers as only conventional screen-film mammography was available at Queen Elizabeth Hospital. Better diagnostic accuracy for DSM might be obtained if a pilot reader study with training cases was performed for the readers to become more familiar in viewing digital mammograms and to provide positive feedback for evaluation of lesions on the malignancy rating scales.

It should be noted that our data is generated from a biased data set, which did not include lesions in which biopsy was not performed, thus not representing a random sample of the lesions in the patient population. The positive predictive value for stereotactic biopsy of microcalcifications in this study was 23.3%, slightly higher than 18.9% for the same procedure among patients attending the Department of Radiology and Imaging at the Queen Elizabeth Hospital (data derived from the audit of interventional breast procedures at the Queen Elizabeth Hospital, with a case mix of 70% problem solving to 30% screening mammograms performed in the year 2002). There was also a selection bias in choosing cases with faint calcifications in parts of the breast deemed difficult to biopsy. The subtlety of lesions in the case sets, which may function as a predictor of case difficulty, may nullify the benefits of improved contrast resolution provided by digital mammography.

Secondly, only cone magnification views of CSFM and spot digital images were given for malignancy rating. As decision making is a complex process, the simple

approach of using a malignancy rating scale from spot views of the calcifications may not be the preferred methods by radiologists in their usual clinical practice. The inter-reader variability in mammographic interpretation could be substantial for both feature analysis and management, albeit the readers in this study had a similar amount of experience.

Radiologists may not consider all the features systematically and tend to depend on a limited number of conspicuous features in their decision making process. Variability exists in the management of indeterminate lesions as a function of interventional thresholds. The greatest inconsistency was noted in the relatively low-suspicion clustered amorphous calcifications (malignancy rating between 'probably benign' and 'suspicious'). Furthermore, high breast density was shown to increase disagreement 2-fold in both lesion detection and final assessment.[19]

Intra-reader variability was also noted when interpreting CSFM images with and without clinical data. Although it is reasonable to be cautious in the evaluation of indeterminate lesions in patients with known risk factors of breast cancer, the lower accuracy with known clinical data should remind us to be more objective and consistent in interpreting screening mammograms of high-risk patients, so as to avoid unnecessary biopsy and anxiety in this group of patients.

Finally, other factors including lesion type, location of the lesion within the breast, preference for soft-copy display by radiologists, and potential for computer-aided diagnosis to improve lesion characterisation on digital mammography were not investigated in this study. Future studies involving larger randomised samples are required to explore the potential role and full benefits of digital spot mammography in the management of breast cancer in the Chinese population.

ACKNOWLEDGEMENTS

The authors wished to thank Dr KH Lai for his advice and assistance in the software application for statistical analysis in this study. A research grant funded by the Hong Kong Polytechnic University for this study is also gratefully acknowledged.

REFERENCES

1. Tsang T, Cheung KF. Incidence and mortality trends of female breast cancer, 1983-1999. Public Health Epidemiol Bull 2002; 11:57-62.
2. Leung GM, Lam TH, Thach TQ, Hedley AJ. Will screening mammography in the East do more harm than good? Am J Public Health 2002;92:1841-1846.
3. Pisano ED, Parham CA. Digital mammography, sestamibi breast scintigraphy and positron emission tomography breast imaging. Radiol Clin North Am 2000;38:861-869.
4. Pisano ED. Current status of full-field digital mammography. Radiology 2000;214:26-28.
5. Center for Devices and Radiological Health page. Food and Drug Administration website. Information for mammography facility personnel, inspectors, and consumers about the implementation of the Mammography Quality Standards Act of 1992 (MQSA). http://www/fda.gov/cdrh/mammography/digital.html.
6. Goldin J, Sayre JW. Review: a guide to clinical epidemiology for radiologists: Part II Statistical analysis. Clin Radiol 1996; 51:317-324.
7. Yip WM, Pang SY, Yim WS, Kwok CS. ROC curve analysis of lesion detectability on phantoms: comparison of digital spot mammography with conventional spot mammography. Br J Radiol 2001;74(883):621-628.
8. Chan HP, Helvie MA, Petrick N, et al. Digital mammography: observer performance study of the effects of pixel size on the characterization of malignant and benign microcalcifications. Acad Radiol 2001;8:454-466.
9. Becker L, Taves D, McCurdy L, Muscedere G, Karlik S, Ward S. Stereotactic core biopsy of breast microcalcifications: comparison of film versus digital mammography, both using an add-on unit. AJR Am J Roentgenol 2001;177:1451-1457.
10. Brettle DS, Ward SC, Parkin GJS, Cowen AR, Sumsion HJ. A clinical comparison between conventional and digital mammography utilizing computed radiography. Br J Radiol 1994;67:464-468.
11. Pisano ED, Cole EB, Kistner EO, et al. Interpretation of digital mammograms: Comparison of speed and accuracy of soft copy versus printed-film display. Radiology 2002;223:483-468.
12. Pisano ED, Cole EB, Major S, et al. Radiologists' preferences for digital mammographic display. The International Digital Mammography Development Group. Radiology 2000;216: 820-830.
13. Metz CE. ROC methodology in radiologic imaging. Invest Radiol 1986;21:720-733.
14. Dorfman DD, Berbaum KS, Metz CE. ROC rating analysis: generalization to the population of readers and patients with the jackknife method. Invest Radiol 1992;27:723-731.
15. Maskarinec G, Meng L, Ursin G. Ethnic differences in mammographic densities. Int J Epidemiol 2001;30:959-965.
16. Chan HP, Sahiner, B, Helvie MA, et al. Improvement of radiologists' characterization of mammographic masses by using computer-aided diagnosis: an ROC study. Radiology 1999; 212:817-827.
17. Cole EB, Pisano ED, Kistner EO, et al. Diagnostic accuracy of digital mammography in patients with dense breasts who underwent problem-solving mammography: effects of image processing and lesion type. Radiology 2003;226: 153-160.
18. Venta LA, Hendrick RE, Adler YT, et al. Rates and causes of disagreement in interpretation of full-field digital mammography and film-screen mammography in a diagnostic setting. AJR Am J Roentgenol 2001;176:1241-1248.
19. Berg WA, Campassi C, Langenberg P, Sexton MJ. Breast Imaging Reporting and Data System: inter- and intraobserver variability in feature analysis and final assessment. AJR Am J Roentgenol 2000;174:1769-1777.