

Lisanne A. L. Khoo, FRCR  
Paul Taylor, PhD  
Rosalind M. Given-Wilson,  
FRCR

Published online  
10.1148/radiol.2372041362  
Radiology 2005; 237:444–449

**Abbreviations:**

CAD = computer-aided detection  
CI = confidence interval

<sup>1</sup> From the Radiology Department, St George's Hospital, Blackshaw Rd, London SW17 0QT, England (L.A.L.K., R.M.G.); and Royal Free and University College Medical School, London, England (P.T.). Received August 3, 2004; revision requested October 8; revision received November 22; accepted December 27. Supported by the National Health Service Breast Screening Programme. **Address correspondence** to R.M.G. (e-mail: [rosalind.givenwilson@stgeorges.nhs.uk](mailto:rosalind.givenwilson@stgeorges.nhs.uk)).

The views and opinions expressed herein are those of the authors and do not necessarily reflect those of the Department of Health or the NHS Breast Screening Programme.

Authors stated no financial relationship to disclose.

**Author contributions:**

Guarantors of integrity of entire study, L.A.L.K., P.T., R.M.G.; study concepts/ study design or data acquisition or data analysis/interpretation, L.A.L.K., P.T., R.M.G.; manuscript drafting or manuscript revision for important intellectual content, L.A.L.K., P.T., R.M.G.; approval of final version of submitted manuscript, L.A.L.K., P.T., R.M.G.; literature research, L.A.L.K., P.T.; clinical studies, L.A.L.K., R.M.G.; experimental studies, R.M.G.; statistical analysis, P.T.; and manuscript editing, L.A.L.K., P.T., R.M.G.

© RSNA, 2005

# Computer-aided Detection in the United Kingdom National Breast Screening Programme: Prospective Study<sup>1</sup>

**PURPOSE:** To evaluate prospectively the recall and cancer detection rates with and without computer-aided detection (CAD) in the United Kingdom National Health Service Breast Screening Programme.

**MATERIALS AND METHODS:** The study had appropriate ethics committee approval. Informed consent was not required; however, patients were informed that their mammograms might be used in research efforts, and all patients agreed to participate. Mammograms obtained in 6111 women (mean age, 58.4 years) undergoing routine screening every 3 years were analyzed with a CAD system. Mammograms were independently double read. Twelve readers participated. Readers recorded an initial evaluation, viewed the CAD prompts, and recorded a final evaluation. Recall to assessment was decided after arbitration. Sensitivities were calculated for single reading, single reading with CAD, and double reading, as a proportion of the total number of cancers detected by using double reading with CAD.

**RESULTS:** A total of 62 cancers were detected in 61 women. CAD prompted 51 (84%) of 61 radiographically detected cancers. Of 12 cancers missed on single reading, nine were correctly prompted; however, seven prompts were overruled by the reader. Sensitivity of single reading was 90.2% (95% confidence interval [CI]: 83.0%, 95.0%), single reading with CAD was 91.5% (95% CI: 85.0%, 96.0%), and double reading without CAD was 98.4% (95% CI: 91.0%, 100%). Cancer detection rate was 1%. Recall to assessment rate was 6.1%, with an increase of 5.8% because of CAD. Average time required, per reader, to read a case was 25 seconds without CAD and 45 seconds with CAD.

**CONCLUSION:** CAD increases sensitivity of single reading by 1.3%, whereas double reading increases sensitivity by 8.2%.

© RSNA, 2005

Although it has been proved that mammography significantly reduces death from breast cancer in women aged 50 years and older (1), mammography is known to be an imperfect screening modality. Errors occur in all areas of radiologic reporting, and estimates of error rates are between 5% and 50% (2). Between 25% and 33% of visible cancers may be missed at mammography because of the naturally complex structure of breast tissue, the difficulty of distinguishing signs of malignancy from variations of normal appearance, and the fatigue and variability of observers (3–6). Double reading has been shown to improve sensitivity, particularly when combined with arbitration (3). Double reading, however, increases the requirement for radiologic staffing, and there is a shortage of radiologists, particularly breast radiologists, in the United Kingdom (7).

Computer-aided detection (CAD) systems can potentially be used to address staffing problems and improve error rates in mammogram reading. The two main commercially available systems (ImageChecker, R2 Technology, Los Altos, Calif; Second Look, iCAD, Nashua, NH) have similar levels of sensitivity of up to 98.2% for microcalcifications, 88.7% for soft-tissue lesions, and 90.0% for malignant lesions overall (8,9). It has been argued

that CAD, performing at this level of sensitivity, has the potential to replace the second reader (10).

To accurately assess the potential of CAD, realistic studies of its effect on reader decision-making are needed. Thus, the purpose of this study was to evaluate prospectively the recall and cancer detection rates with and without CAD in the United Kingdom National Health Service Breast Screening Programme.

## MATERIALS AND METHODS

### Subjects

This study was approved by the ethics committee. Our ethics committee did not require individual consent because our study used approved widespread technology and the established routine technique of double reading. Nevertheless, all patients screened at the South West London Breast Screening Service are informed that this screening unit is involved with training and research and that their records may be used for these purposes. Patients have the opportunity to refuse use of their records and imaging studies for research or training. No patients in our study declined to participate.

Women were enrolled through the South West London Breast Screening Service, which is a part of the United Kingdom National Health Service Breast Screening Programme. Women aged 50–64 years are invited to undergo mammographic screening every 3 years. Women older than 64 years may refer themselves. The examination consists of two views of each breast. Between March 21, 2003, and January 9, 2004, inclusive mammograms obtained in 6111 women (age range, 45–94 years; mean age, 58.4 years) were read with CAD. During this period, another 13 391 mammograms were read on rollers; CAD was not available on these rollers.

### Study Design

Mammograms obtained in these 6111 women were digitized and analyzed with the ImageChecker system, version 5.0 (R2 Technology). Women requiring more than the standard four views for coverage of the whole of both breasts could not be analyzed with this system, and they were excluded from the study. Mammograms were loaded in batches on a roller viewer and accompanied by previously obtained round screening mammograms, when available. Twelve film readers participated in this study; seven readers were consultant radiologists (including R.M.G.), and five readers

were radiographers trained in reading screening mammograms. All mammograms were independently double read by at least one consultant radiologist. All readers met the training requirements of the National Health Service Breast Screening Programme: They read at least 5000 screening mammograms per year; they were aged 34–56 years, with a mean age of 43 years; and their experience in mammography ranged from 4 to 23 years, with a mean experience of 11 years.

Each reader viewed current and available prior mammograms for each patient; at this time, an opinion about any visible abnormalities and whether to recall the patient for further assessment was recorded. CAD prompts for the current mammograms were then displayed on an adjacent liquid crystal display screen. The reader then reassessed the prompted areas before recording a revised assessment.

After double reading, women were considered healthy or assigned for (a) discussion, with a view to recall the patient for technical repeat of mammography; (b) comparison of current and prior mammograms, which were not available in the unit; or (c) arbitration, with a view to recall the patient for further clinical assessment.

Women in the first two categories were excluded from the study, since CAD prompts would not necessarily be available when images were reviewed. At arbitration, cases were discussed by an additional two consultant radiologists who reviewed current and previously obtained (if available) mammograms, CAD prompts, and proformas completed by the first two readers before they decided whether to recall patients for assessment. All of the radiologists (including R.M.G.) who participated in the study participated in arbitration meetings. Final outcomes were recorded after assessment.

A summary of all the women excluded from the study is presented in Table 1. There were two cancers in these women. One cancer was in a woman who required more than four views for coverage of her breast. The other cancer was in a woman who had an expedited appointment; therefore, although CAD prompts are available, there is no reader analysis data for CAD.

### Statistical Analysis

Both the proportion of women recalled to undergo assessment and the cancer detection rate were assessed. The recall to arbitration rate was recorded. The assessments recorded by radiologists before and after they inspected the prompts allowed them to determine the number of

**TABLE 1**  
Women not Processed with CAD or Excluded from CAD Analysis

| Reason for Exclusion         | No. of Women |
|------------------------------|--------------|
| More than four film images   | 405 (6.6)    |
| Technically inadequate       | 46 (0.8)     |
| Technical recall discussion  | 300 (4.9)    |
| Location of prior mammograms | 138 (2.2)    |

Note.—Data in parentheses are percentages.

women in which they changed their assessment. The proportion of such women was calculated for the total number of women recalled to arbitration or assessment and for the total number of cancers detected. Each case was double read; thus, the number of observations is double the number of women, minus any exclusions or women with missing data.

Relative sensitivity was calculated for each of three protocols (ie, single reading, single reading with CAD, and double reading) by dividing the number of cancers detected with each protocol by the total number of cancers detected. The data obtained are not a measure of the true sensitivity because until 3-year follow-up data are obtained, the total number of missed cancers cannot be ascertained. Thus, sensitivity is expressed as a proportion of the total number of cancers detected with double reading with CAD. Exact 95% confidence intervals (CIs) for the relative sensitivities were calculated. Where the CIs for two sensitivities overlapped, the difference was assumed to not be statistically significant.

For double reading, cancer was considered detected if either reader detected it at step 1; cancer was considered missed if neither reader detected it until step 2 (Figure). A cancer detected by a reader was considered detected at single reading if the reader detected it at step 1; otherwise, it was deemed missed. If a cancer was not detected by a reader at step 1 but was detected at step 2, it was deemed detected by that reader for single reading with CAD; otherwise, it was deemed missed.

Timings were obtained for reading 18 rollers with and without CAD. These timings were obtained after the main data collection period was finished; therefore, they reflect normal workflow and are unaffected by the additional data collection requirements of the study. The timings were measured as the mean timing per case per reader for each roller. Estimates

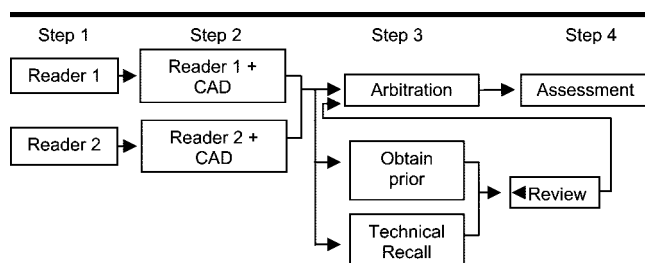


Diagram shows sequence of steps in women whose images were read with the ImageChecker system (R2 Technology).

**TABLE 2**  
Number of Readings Attributed to CAD at Each Point in the Process

| Assessment  | Total Observations | No. Attributed to CAD |
|---|--------------------|-----------------------|
| Recalled to arbitration                             | 1639               | 89 (5.7)              |
| Recalled to assessment after arbitration discussion | 635                | 36 (6.0)              |
| Detected cancers                                    | 118                | 2 (1.7)               |

Note.—Data in parentheses are the percentage increase that is attributed to CAD.

**TABLE 3**  
Relative Sensitivities of Screening Protocols

| Screening Protocol                              | Proportion of Readings | Relative Sensitivity (%) | Missing Patients   |
|---|------------------------|--------------------------|--|
| Single reading                                  | 110/122                | 90.2 (83, 95)            | Two separate single entries excluded because of data recording discrepancy   |
| Single reading with CAD                         | 108/118                | 91.5 (85, 96)            | Two separate single entries excluded because of data recording discrepancy<br>Two readings; no CAD data as more than four views<br>Two readings without CAD information as appointment was expedited |
| Double reading                                  | 61/62                  | 98.4 (91, 100)           | None   |
| Double reading with CAD and symptomatic history | 62/62                  | 100                      | None   |

Note.—Data in parentheses are 95% CIs. Women excluded from analysis are detailed as missing patients.

were also obtained for the time spent on arbitration per reader by monitoring time taken and number of cases arbitrated over a 3-week period. Patients recalled for assessment were given 1-hour appointment times; thus, 1 hour was used as an estimate of the cost in clinician time of recalling a patient for assessment.

Radiologic and gross pathologic details of detected cancers were recorded, as were prompt rates and whether cancers were correctly prompted. A prompt was deemed correct if, after review by two authors (R.M.G. and L.A.L.K.), the cancer was judged to be marked in one or both views with the appropriate mass or calci-

fication marker. A sample of 613 noncancerous cases was selected at random and reviewed by an author (L.A.L.K.) to provide an estimate of the false prompt rate.

Recall and cancer detection rates were also measured for the period of the study on the mammograms read on rollers that were not equipped with CAD viewers. The procedure followed was the same as that used for mammograms, which involved double reading of current and available prior mammograms followed, when appropriate, by arbitration and recall. Twelve readers read these mammograms, including 10 readers who participated in the study.

We used the StatPages.net tool (accessed

October 21, 2004), which is available at <http://members.aol.com/johnp71/confint.html>, for statistical analysis.

## RESULTS

### Effect of CAD on Individual Readers

Table 2 shows the number of cancers diagnosed with CAD at each point in the process. These data represent the number of occasions the reader recalled the case only after viewing the CAD prompts as a proportion of the total number of observations.

### Relative Sensitivities

Table 3 shows the relative sensitivities of three screening protocols. The measure of sensitivity is relative to that of double reading with CAD. Sensitivity of single reading is 110 of 122 single readings, or 90.2% (95% CI: 83.0%, 95.0%); in 12 readings, a reader indicated that a patient was healthy before viewing the prompts, but the patient was recalled (either the reader changed her mind after viewing the prompts or the patient was recalled by the other reader or on the basis of symptoms), and cancer was confirmed. Sensitivity of single reading with CAD is 108 of 118, or 91.5% (95% CI: 85.0%, 96.0%) since there were 10 occasions when a reader indicated a mammogram was normal after prompting; however, the woman was recalled by the other reader, and cancer was confirmed. All but one of the cancers were seen by one of the readers before looking at the prompts; thus, sensitivity in double reading is 61 of 62, or 98.4% (95% CI: 91.0%, 100%).

### Prompt Rates

The ImageChecker system (R2 Technology) prompted 51 of 61 (84%) cancer cases. Of the 12 instances in which a reader did not recall the cancer before prompting, nine cancers were correctly prompted. Action was taken in two correct prompts (one mass and one calcification). All seven prompted cancers in which the reader took no action were mass prompts. The false prompt rate was 1.59 per case. One of the prompted cancers in which no action was indicated by either reader after CAD prompting was in a patient who was already scheduled for recall, as she was symptomatic at the time of screening.

## Timing

Mean time per reader per case was 25 seconds for reading without CAD (hence 50 seconds for double reading) and 45 seconds for reading with CAD. Ignoring the possible confounding factor of non-independence between readers, a Mann-Whitney test was performed with these data ( $z = 4.46$ ,  $P < .001$ ) (18 rollers in each group). Each additional arbitration requires 2.2 minutes of radiologist time, and each additional assessment appointment requires 1 hour. The extrapolated overall time effect of single reading with CAD relative to double reading is presented in Table 4. For our study group of 6111 women, single reading with CAD took an extra 9.43 hours when compared with double reading.

## Gross Pathologic Data

Gross pathologic results for cancers allocated to CAD are given in Table 5. Of the two women in whom reader opinion changed after CAD, one had a 30-mm-diameter grade 3 invasive ductal carcinoma with associated lobular carcinoma in situ, and two of three lymph nodes were positive; the other woman had a 7-mm-diameter intermediate and high-grade ductal carcinoma in situ, without an invasive component.

## Comparison with Mammograms Read without CAD

More women were allocated to arbitration in the group of patients whose mammograms were read with CAD (13.8%; 95% CI: 13.0%, 15.0%) than in the group of patients whose mammograms were read on rollers without CAD viewers (10.5%; 95% CI: 9.1%, 12.0%). Significantly more women were also recalled for assessment in the CAD group (6.1%; 95% CI: 5.5%, 6.7%) than in the non-CAD group (5.0%; 95% CI: 4.6%, 5.4%). There was, however, no significant difference in cancer detection rates in the study sample (10 per 1000; 95% CI: 7.0%, 13.0%) compared with that in the other cases (8.4 per 1000; 95% CI: 7.0%, 10.0%).

## DISCUSSION

Prospective studies of CAD have used two different designs. Some researchers have asked radiologists to record their assessments before and after viewing CAD prompts and reported the proportion of cancers detected only after consulting the prompts. Previously reported in-

**TABLE 4**  
Extrapolated Comparison of Time Costs of Single Reading with CAD Relative to Double Reading

| Comparison                                  | Stage of Evaluation      |                      |                    | Overall Time Cost (min) |
|---|--------------------------|----------------------|--------------------|-------------------------|
|   | Film Image Reading (min) | Arbitration (min)    | Assessment (min)   |                         |
| Per case                                    | −0.09                    | +2.18                | +60                | NA*                     |
| Per study group<br>(12 222 single readings) | −1100                    | +189.65 <sup>†</sup> | +2040 <sup>‡</sup> | +1130                   |

Note.—Recall to arbitration and assessment values reflect additional benign cases attributable to CAD. Negative time entry denotes radiologist time saved for single reading with CAD relative to double reading; positive time entry denotes time penalty.

\* NA = not applicable.

<sup>†</sup> 87 single readings.

<sup>‡</sup> 34 single readings.

**TABLE 5**  
Pathologic Data for 62 Cancers Detected in Patients Allocated to CAD

| Tumor Type  | No. of Cancers<br>(n = 62) |
|---|----------------------------|
| Ductal carcinoma in situ alone                          | 13*                        |
| Ductal carcinoma in situ with lobular carcinoma in situ | 1                          |
| Invasive ductal carcinoma                               | 35                         |
| Invasive tubular carcinoma                              | 4                          |
| Invasive lobular carcinoma                              | 7                          |
| Invasive lobular and ductal carcinoma                   | 1                          |
| Large necrotic tumor, likely breast primary             | 1                          |

\* Size range, 3–40 mm. High grade cancer,  $n = 6$ . Other grades of cancer,  $n = 7$ .

creases in cancer detection with this design range from 12% (11) to 19.5% (12). Other researchers have compared the overall cancer detection rate in two distinct sets of women, with CAD used in only one set of women. Gur et al (13) evaluated findings over a period in which CAD was available for some but not all of the women seen. They found no difference in cancer detection rate with or without CAD. Cupples (14) found an increase of 13% in cancer detection after the introduction of CAD. The ideal way to evaluate the relationship between CAD prompting and reader behavior would be with a randomized controlled trial of CAD; to our knowledge, however, no such report has been published. Our study was designed to replicate the study of Freer and Ulissey (12).

In our study, there was an increase of recall both to arbitration and to assessment (5.7% and 6.0% increase, respectively) because of CAD in the study group. Studies have shown that the effects of CAD range from slight improvements in specificity to increases in recall rate of up to 19% (11–13,15,16). In the 36 single readings (31 women) recalled to assessment because of CAD, there were

two cancers; thus, the positive predictive value for recall to assessment on the basis of a CAD prompt is 5.6%. This compares unfavorably with an overall positive predictive value of 19% for malignancy of recall cases in the unit.

Although the number of increased cancers (above single reading) detected with CAD was small (1.5%, two of 118 readings), the total is similar to that in other prospective studies. Bandokar et al (17) found two extra cancers that were prompted by CAD in 4089 screening cases (12%), and Young et al (11) found three extra cancers that were prompted by CAD in 12 082 screening mammograms (6.2%). These percentages are larger than those in our study because fewer total cancers were detected. This reflects differences between screening practice in the United States and that in the National Health Service Breast Screening Programme. In the United Kingdom, the women screened are older (mean age, 58.4 years in this study), the screening interval is longer (3 years), and double reading is routinely used. In the study of Freer and Ulissey (12), the median age was 49 years, the screening interval was 13 months, and single reading was practiced.



## Why Correct CAD Prompts Are Dismissed

In retrospective studies, the sensitivity of CAD in women with a cancer that was missed has been measured, and it has been assumed that this would translate into increased cancer detection rates. Brem et al (18) found that CAD prompted 65% of 123 missed cancers. They assume that these cancers would have been identified if prompted; therefore, they estimate a 21.2% increase in radiologist sensitivity. Warren Burhenne et al (19) used a similar method and estimated that the false-negative rate could be cut by 77%. These studies, however, assume that readers will respond to every correct CAD prompt. The findings of this study, which are in line with the findings of other studies, show that readers reject some true prompts (15,20,21).

It is important to understand this phenomenon. The study design ensured that readers turned on the screen and looked at the CAD prompts. The seven correct prompts that were ignored or overruled were all for masses. There were few false prompts for these cases (three false prompts for seven sets of mammograms). Previous studies have shown that a high level of confidence that a mammogram is normal is associated with increased likelihood that a correct prompt will be ignored or overruled (21). It has been postulated that greater confidence can be placed on calcification prompts than on other prompts (22).

It is possible that CAD will have more influence on detection and search errors than on decision errors. In this study, despite the high cancer detection rate of 1%, 99% of women are healthy. With a false prompt rate of 1.59 per case for 6050 healthy women and a sensitivity of 84% for 62 cancers, readers will have to dismiss 180 false prompts for every true prompt. This low specificity may be the major factor explaining why readers are more likely to ignore correct prompts than respond to them. It has been postulated that failure to act on nonspecific but CAD-marked findings prospectively in patients with subtle cancer does not constitute negligence (23).

## Timings and Resources

Average time required for single reading of a set of mammograms was 25 seconds without CAD and 45 seconds with CAD. Thus, reviewing CAD prompts and the original mammograms nearly doubles the time a single reader takes to read

a set of mammograms. Time saved per reading, by adopting CAD with single reading rather than double reading, is 5 seconds per case. However, when extra time spent on arbitration and assessment is considered, there is a net loss of 9.43 hours of radiologist time over the study period.

Use of CAD resulted in increased demand on resources, including a fee of \$4800 (in U.S. dollars) per month for the lease of the ImageChecker system (R2 Technology). Also, a part-time member of the clerical staff sorted mammograms, put them through the digitizer, and matched them with a bar code reader to enable them to be loaded on the CAD roller. In the United States, the financial calculation would be different because of increased reimbursement for the use of CAD, which may range from \$15 to \$28 per examination (24).

## Study Design

There are some potential shortcomings in the design of our study. The outcome of patients with negative mammograms is not confirmed. To establish the true false-negative rate, 3 years of follow-up will be needed. Although the outcome of such follow-up will be interesting, it does not affect the primary measure of cancers diagnosed with CAD.

A number of women were excluded from analysis because they were designated for technical recall, reading was delayed while old mammograms were obtained, or multiple mammograms were obtained at initial screening. Only two cancers occurred within this group of patients, and it does not appear that exclusion of these women significantly changed the measurement of cancers in the CAD group.

A further question is whether training and experience with the CAD system were adequate for readers during the study. The majority of readers in this study were trained to use this system by the manufacturer applications specialist; this training is given to new purchasers of the system. This training was cascaded by those who were present to the few readers who were not present at initial training. Readers participating in this study also took part in previous archive studies of CAD within our unit. During the study, they were using CAD in their routine daily screening work for more than 9 months. Participation in arbitration discussion, assessment, and gross pathologic review meetings provided readers with feedback on the outcome of

women undergoing screening during this period. No evidence of an increasing or decreasing trend in terms of responsiveness to prompts was seen during this time.

The study is not a straightforward comparison of single reading with CAD against double reading, since the extra step of arbitration applied to all women in whom recall was considered. This should not affect the main outcome measure of sensitivity, as the function of arbitration is to improve specificity.

In a number of cases, readers changed their mind to recall to arbitration discussion on the basis of CAD prompting (89 single readings in 78 women). Only 31 women (36 single readings) had findings that were deemed suspicious at arbitration, and these women were recalled and assessed. Only two women had cancer. It will be interesting to follow up the rest of these women over 3 years to see whether any cancers have developed that might also have been potentially diagnosed with CAD.

Recall (6.1% vs 5.0%,  $P < .001$ ) and cancer detection (1.0% vs 0.84%,  $P = .2$ ) rates were higher in the study group than in patients with mammograms that were read during the same period on rollers not equipped with CAD. The overwhelming majority of the extra cancers were detected before viewing CAD prompts. Although there was no systematic difference between the women whose films were read with and without CAD, there was a difference between the readers reading CAD and non-CAD cases, with two readers reading only non-CAD cases.

## Conclusion

Double reading with CAD served as the reference standard; the relative sensitivity of single reading is 90.2% (95% CI: 83.0%, 95.0%), the relative sensitivity of single reading with CAD is 91.5% (95% CI: 85.0%, 96.0%), and the relative sensitivity of double reading is 98.4% (95% CI: 91.0%, 100%). Single reading with CAD is significantly slower than single reading without CAD. Our estimates suggest that use of single reading with CAD as an alternative to double reading would increase the time required. CAD prompted 84% of cancers in this study; however, readers ignored or overruled 78% of correct prompts on missed cancers.

**Acknowledgments:** The participation of staff at the Duchess of Kent Screening Unit, St George's Hospital, London, England, is gratefully acknowledged, as is the contribution of Henry Potts, PhD, who advised on statistical analyses.

## References

1. International Agency for Research on Cancer Prevention. Breast screening. In: IARC handbooks of cancer prevention. Vol 7. International Agency for Research on Cancer Prevention, World Health Organization. Lyon, France: IARC Press, 2002.
2. Fitzgerald R. Error in radiology. *Clin Radiol* 2001;56:938–946.
3. Blanks RG, Wallis MG, Moss SM. A comparison of cancer detection rates achieved by breast cancer screening programmes by number of readers, for one and two-view mammography: results from the UK National Health Service Breast Screening Programme. *J Med Screen* 1998;5(4):195–201.
4. Moss S, Blanks R. Calculating appropriate target cancer detection rates and expected interval cancer rates for the UK NHS Breast Screening Programme: Interval Cancer Working Group. *J Epidemiol Community Health* 1998;52(2):111–115.
5. Beam CA, Layde PM, Sullivan DC. Variability in the interpretation of screening mammograms by US radiologists: findings from a national sample. *Arch Intern Med* 1996;156:209–213.
6. Blanks RG, Wallis MG, Given-Wilson RM. Observer variability in cancer detection during routine repeat (incident) mammographic screening in a study of two versus one-view mammography. *J Med Screen* 1999;6(3):152–158.
7. Field S. Breast screening issues: Royal College of Radiologists Survey July 2003. Royal College of Radiologists [newsletter] 2004;76:13–15.
8. R2 Technology. Clinical data. R2 Technology. Available at: <http://www.r2tech.com/prf/prf001.html>. Accessed July 15, 2004.
9. iCAD. Clinical trial summary. iCAD. Available at: [http://www.cadxmed.com/clinical\\_data/trial\\_summary.html](http://www.cadxmed.com/clinical_data/trial_summary.html). Accessed July 15, 2004.
10. Malich A, Marx C, Facius M, Boehm T, Fleck M, Kaiser WA. Tumour detection rate of a new commercially available computer aided detection system. *Eur Radiol* 2001;11(12):2454–2459.
11. Young WW, Destounis SV, Bonaccio E, Zuley ML. Computer aided detection in screening mammography: can it replace the second reader in an independent double read? preliminary results of a prospective double blinded study (abstr). *Radiology* 2002;225(P):600.
12. Freer TW, Ulissey MJ. Screening mammography with computer-aided detection: prospective study of 12,860 patients in a community breast center. *Radiology* 2001;220(3):781–786.
13. Gur D, Sumkin JH, Rockette HE, et al. Changes in breast cancer detection and mammography recall rates after the introduction of a computer-aided detection system. *J Natl Cancer Inst* 2004;96(3):185–190.
14. Cupples TE. Impact of computer-aided detection (CAD) in a regional screening mammography program (abstr). *Radiology* 2001;221(P):520.
15. Taylor PM, Champness J, Given-Wilson RM, Potts HW, Johnston K. An evaluation of the impact of computer-based prompts on screen readers' interpretation of mammograms. *Br J Radiol* 2004;77(913):21–27.
16. Thurfjell E, Thurfjell MG, Egge E, Bjurstam N. Sensitivity and specificity of computer-assisted breast cancer detection in mammography screening. *Acta Radiol* 1998;39(4):384–388.
17. Bandokar P, Birdwell RL, Ikeda DM. Computer-aided detection (CAD) with screening mammography in an academic institution (abstr). *Radiology* 2002;225(P):458.
18. Brem RF, Baum J, Lechner M, et al. Improvement in sensitivity of screening mammography with computer-aided detection: a multi-institutional trial. *AJR Am J Roentgenol* 2003;181(3):687–693.
19. Warren Burhenne LJ, Wood SA, D'Orsi CJ, et al. Potential contribution of computer-aided detection to the sensitivity of screening mammography. *Radiology* 2000;215(2):554–562.
20. Field S, Garvican L. A pilot evaluation of the R2 ImageChecker system and user's response in the detection of interval breast cancers on previous screening films. *Clin Radiol* 2001;56(10):833–837.
21. Taylor P, Given-Wilson R, Champness J, Potts HW, Johnston K. Assessing the impact of CAD on the sensitivity and specificity of film readers. *Clin Radiol* 2004;59(12):1099–1105.
22. D'Orsi CJ. Computer-aided detection: there is no free lunch. *Radiology* 2001;221:585–586.
23. Ikeda DM, Birdwell RL, O'Shaughnessy KF, Sickles EA, Brenner RJ. Computer-aided detection output on 172 subtle findings on normal mammograms, previously obtained in women with breast cancer detected at follow-up screening mammography. *Radiology* 2004;230:811–819.
24. Elmore JG, Carney PA. Computer aided detection of breast cancer: has promise outstripped performance? *J Natl Cancer Inst* 2004;96(3):162–163.