EVALUATION OF CAD WITH FULL FIELD DIGITAL MAMMOGRAPHY IN THE NHS BREAST SCREENING PROGRAMME

NHSBSP Equipment Report 0910
December 2009

Authors

Paul Taylor,[a] Stephanie Wilson,[b] Henry Potts,[a] Louise Wilkinson,[c] Lisanne Khoo,[c] Rosalind Given-Wilson[c]

[a]Centre for Health Informatics and Multiprofessional Education, University College London, Highgate Hill, London
[b]Centre for HCI Design, City University, Northampton Square, London
[c]Duchess of Kent Breast Screening Unit, Blackshaw Road, London

Enquiries

Enquiries about this report should be addressed to

Paul Taylor
Centre for Health Informatics and Multiprofessional Education
University College London
Highgate Hill
London N19 5LW

Tel: +44 20 7288 3548
Email: p.taylor@chime.ucl.ac.uk

Published by

NHS Cancer Screening Programmes
Fulwood House
Old Fulwood Road
Sheffield
S10 3TH

Tel: 0114 271 1060
Fax: 0114 271 1089
Email: info@cancerscreening.nhs.uk
Website: www.cancerscreening.nhs.uk

Typeset by Prepress Projects Ltd, Perth (www.prepress-projects.co.uk)
Printed by Duffield Printers

CONTENTS

	ACKNOWLEDGEMENTS	iv
	EXECUTIVE SUMMARY	1
1.	INTRODUCTION	2
1.1	Objectives	2
2.	METHODS AND RESULTS	3
2.1	System tested	3
2.2	Observational study of readers working with FFDM with and without CAD	3
2.3	Impact on workload	6
2.4	The cost-effectiveness of CAD	7
2.5	Conformance with standards	10
3.	CONCLUSIONS	11
	REFERENCES	12

ACKNOWLEDGEMENTS

The authors are grateful to staff at the Duchess of Kent Screening Unit for their help with this evaluation.

EXECUTIVE SUMMARY

An evaluation of the use of computer aided detection with full field digital mammography was undertaken with the support of the NHS Breast Screening Programme. It included the following components

a. a study of the effect of computer aided detection (CAD) on workflow and the time taken to read a mammographic image
b. a systematic review and meta-analysis of studies comparing: single reading of images with and without CAD; double and single reading of images without CAD
c. an assessment of the cost-effectiveness of CAD in full field digital mammography (FFDM)
d. a review of issues relating to conformance with key standards.

The main points to emerge under each heading were as follows

a. Supplementing FFDM with CAD did not increase reading time. However the time taken to read each screening case was increased substantially by the use of FFDM rather than analogue film. Using CAD with FFDM was found to be simple and intuitive and the majority of readers would use it again. The addition of CAD had no demonstrable effect on the number of women recalled for arbitration or assessment; the number of cases examined was small, however, and only gross effects would have been detected.
b. A systematic review and meta-analysis was conducted of studies produced up to 2008 and has been published elsewhere.[1] Ten of these studies compared single reading of images with CAD and without it. They examined 410 000 subjects and found that CAD did not significantly increase sensitivity in cancer detection (1.04, 0.96–1.13). Seventeen studies compared single reading (SR) and double reading (DR) of images without CAD. They examined 1.2 million subjects and showed increased cancer detection rates with DR (1.10, 1.06–1.14). Recall rates increased for both CAD+SR (1.10, 1.09–1.12) and DR (1.17, 1.15–1.18) but decreased for DR with arbitration (0.94, 0.92–0.96). A later randomised controlled trial of CAD+SR compared with DR showed that they are equally effective at cancer detection.[2] The cost-effectiveness of CAD+SR with FFDM was compared with DR, focusing on the amount of time a radiologist needed to read cases and deal with the women recalled. An earlier study had shown that using CAD prompts with analogue film reading almost doubled reading time and demanded more radiologist time than did DR.[3]
c. For its analysis of CAD with FFDM the present report considered a range of possible increases in recall rates, in the time allocated for assessment, and in the number of CAD licences purchased. It demonstrated that it is possible to save radiologist time overall and reduce cost by using CAD+SR rather than DR.
d. Conformance with relevant standards – Integrating the Healthcare Enterprise (IHE) and Digital Imaging and Communications in Medicine (DICOM) standards – is discussed briefly where it relates to installing CAD with FFDM.

In summary, CAD with FFDM was found to be user friendly and did not affect radiologist workflow or reading time during film-screen reading. In the published studies, estimates vary of the extent to which using CAD increases recall rates; depending on these rates, and the time taken to assess a woman, CAD can save radiologist time during the screening and assessment process when compared with DR. Systematic review of published literature comparing CAD+SR with SR showed no significant increase in cancer detection rates and an increase in recall rates. When compared with SR, however, DR with arbitration significantly increased cancer detection and decreased recall rates.

1. INTRODUCTION

Computer aided detection systems for mammography have been available for a number of years. They have been widely adopted in the US, where evidence that they can improve sensitivity has encouraged their use to improve cancer detection by a single reader. DR is standard practice in the UK. Here, the argument for CAD has been not that it could improve on the sensitivity of DR, but that it could allow a single reader to achieve equal sensitivity. This would be desirable, for example, where centres are unable for practical reasons to perform DR, or where replacing DR with CAD+SR would be more cost-effective. The case for CAD in the UK therefore rests on whether

a. its sensitivity for cancer detection is equivalent to DR
b. recall rates can be kept acceptably low using CAD rather than a second reader
c. the logistics of CAD use are compatible with the NHS Breast Screening Programme (NHS BSP)
d. CAD is a cost-effective alternative to a second reader.

It is important to consider the logistics and economics of CAD use because the existing research into CAD focuses on its sensitivity and, as a result, relatively little is known about its cost-effectiveness or its impact on workflow. Analysis of these aspects is complicated by the fact that mammography as a specialty is in the process of adopting digital imaging technology, which dramatically simplifies the introduction of CAD.

1.1 Objectives

The purpose of this evaluation was to assess the logistics of using CAD with FFDM and to report on its suitability for use in the NHS BSP. We assessed

a. practical aspects of using the technology
b. the impact on radiologists' time
c. possible impact on cost-effectiveness
d. conformance with the relevant standards.

The practical aspects of the technology were addressed by means of an observational study of the system and by a questionnaire survey of image readers' attitudes to the system at the end of the study. The time taken to read films with and without CAD was measured during two periods after its introduction. The possible impact on cost-effectiveness was assessed through a systematic review of the published evidence and an economic evaluation. Conformance with relevant standards is discussed in relation to published statements on standards met by CAD systems. Each of these is explored in turn in the following section.

The study does not attempt to measure the impact of CAD on sensitivity; other trials do that. It does look at the impact of CAD on rates of referral for arbitration and rates of recall, but this was a small study and only large effects would have been detected. Finally, this was not a comparative study of different CAD systems: the market leading CAD system – R2 – was used throughout with a Hologic Selenia FFDM system.

2. METHODS AND RESULTS

This section describes the system that was assessed and the setting in which the assessment took place, before turning to the four objectives outlined in section 1.1.

2.1 System tested

In preparation for the study, a Hologic Selenium FFDM with CAD was installed in the South West London Breast Screening Unit in August 2007. It was used initially without CAD to allow readers to become familiar with the technology. Following problems with the reliability of the system, however, the machine was removed. A newer version of the Selenium was installed in November 2008. (See Figure 1.) Once again, data were collected initially without CAD, which was introduced a month later and remained in use until March 2009.

Before using CAD, image readers received the following training

- the manufacturer's standard training package (one session with an applications specialist and access to an e-package of instructions)
- a training roller viewer consisting of a set of cancer cases with CAD prompts, feedback on the radiological results of screen reading (including the assessment) and the definitive pathological findings following surgical treatment.

The South West London Breast Screening Unit employs a mixture of consultant radiologists and trained radiographers as film readers. It has adopted a protocol of DR with arbitration in which any film that either reader considers suspicious is referred for arbitration to two radiologists. The Unit uses an automated data entry system for analogue screening. It is integrated with the NBSS system but not with the FFDM system; readers therefore used paper forms to record the outcomes of digital films, rather than using two different computerised systems for a single task.

2.2 Observational study of readers working with FFDM with and without CAD

As part of this project, City University London undertook a small qualitative study to investigate the impact of CAD with FFDM on work processes. There were two aims for this element of the work

- to investigate any changes in workflow resulting from the introduction of CAD
- to investigate related logistical and usage issues.

In addition a member of the Unit carried out a questionnaire survey at the end of the project. This section summarises the qualitative study, its main findings and the survey results. It became clear during the course of the study that the move from analogue to digital reading had a significant impact on work processes and this, too, is briefly discussed.*

*Details of these elements of the study may be obtained from the first author.

2.2.1 The qualitative study

Qualitative data were collected on six occasions: three times during phase one of the study and three times during phase two. A total of five readers were observed in phase one using FFDM without CAD and six readers were observed in phase two using FFDM with CAD. Participants were a mixture of radiologist and radiographer film readers; they had varying levels of experience of reading mammograms and of using FFDM and CAD.

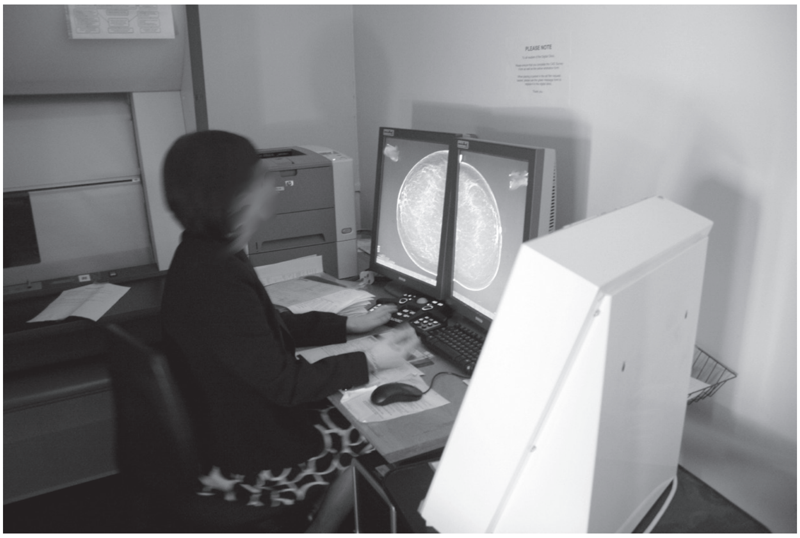

Figure 1 Arrangement of the digital machine, with the roller visible to the left of the reader and a light box to the right.

2.2.2 Findings

The technical work of reading mammograms is one part of a much longer screening process; the summaries that follow focus solely on the reading activity.

A. Summary workflow without CAD

1. Log in
2. Select cases to be read: check the clinic date on the paper sheet and select cases on the system (either by selecting all those screened on the given date or by using a combination of the date and the names of the first and last women in the clinic)
3. Read films for each woman in the clinic
   - 3.1 Display the case
   - 3.2 Check that the name on the screen matches the name on the packet
   - 3.3 Read the films: go through each viewing step in the pre-determined hanging order (eight steps in most cases). Where there are concerns, do one or more of the following: move backwards and forwards between views; compare views with old films; use the on-screen magnifier for a closer look; use the on-screen ruler (or a hand-held measuring device such as a pen) to compare distances in different views
   - 3.4 If there are extra views, look at them
   - 3.5 Complete the yellow paper form to report the outcome for this case. The results are not entered on the FFDM system
4. If it is a second reading complete the clinic list by writing down the names of women sent to arbitration.

In phase two (where CAD was used) readers were instructed to view the films for each case initially without CAD and then to review the CAD prompts and double check any prompted areas. The reader progresses through a series of image displays (the 'hanging protocol') at a single press of a button. The last hangings in the protocol show the CAD prompts, either on large or on small images or on both. These prompts were reviewed by all the readers who were observed and CAD markers were displayed in many of the cases. This was clearly an extra step in the workflow for every case that was read. In most cases, however, switching on CAD and looking at the resulting display was a very brief step and a technically undemanding part of the routine hanging sequence for viewing films. A more significant impact arose from readers' responses to the CAD markers. This response varied in duration from case to case: from a few seconds when the reader was completely confident that there were no abnormalities in the films to several minutes when the reader responded by investigating further.

B. Summary workflow with CAD

1. Log in
2. Select the cases to be read (as in A2 above)
3. Read films for each woman in the clinic
   - 3.1 Display the case
   - 3.2 Check that the name on the screen matches the name on the packet
   - 3.3 Read the films (as in A3.3 above)
   - 3.4 If there are extra views, look at them
   - 3.5 Switch to the CAD hanging protocol
   - 3.6 Respond to CAD by finishing the case, or by some combination of going back to look again, using the on-screen magnifier to examine what has been marked, or comparing views with old films
   - 3.7 Complete the yellow paper form to report the outcome for this case
4. If it is a second reading, complete the clinic list.

The CAD system is well designed and unobtrusive. Readers could readily and rapidly switch it on and off, and the markers were displayed in a manner that was subtle yet readable. As a result, although switching on CAD was an extra step in the workflow its impact on the overall time taken to read a case appeared almost negligible. Readers' responses to the CAD markers had a greater impact on the workflow and time taken.

The impact of digital reading on the workflow was greater than the introduction of CAD and disrupted working practice in several ways. One of the most fundamental changes associated with the move to digital concerns how old films were used. When using analogue, if old films were available they were displayed below the current ones. This allowed readers to make comparisons between old and new films without moving their eyes from the area of the roller viewer in front of them. It was more difficult to make these comparisons when using the digital set-up. At the start of the study (while readers were growing familiar with FFDM) soft copy versions of old films were loaded onto a roller viewer adjacent and at 90 degrees to the workstation. Prevalent screening cases (for which there were no prior films) had black films loaded onto the relevant space on the viewer to ensure that the number of cases on the roller matched the number on the workstation. This was designed to allow readers to move up one case on the roller for each case read on the workstation and so synchronise prior and current cases. In practice this proved awkward, and readers tended to lose synchronisation. They were required to move the case on the roller, select the next case on the workstation, look at the woman's packet and enter a result for each case, all of which could easily result in confusion. Later in the study, packets containing women's prior films in the order of the clinic were provided on the workstation and on a viewing box adjacent and at 90 degrees to the workstation. This made it easier to synchronise the packets and next cases on the workstation and readers could select prior films from the packet when they wished. With FFDM, readers appeared to look for abnormalities without reference to prior films and to refer to them only when concerned. On several occasions, second readers did not look at the old films because they could see on the report sheet that the first reader had already done so. Another way of making prior films available would be to digitise them and display them on the workstation. This has been shown to increase the number of occasions on which prior films are looked at and the time spent reviewing them. Digitisation of prior images was not possible during this evaluation, however.

In summary, the introduction of CAD added at least one step to the workflow for all readers; in some cases, responses to CAD markers added several minutes to the overall reading time. However the fact that readers spend extra time reviewing cases after turning on CAD does not necessarily mean that reading time is longer overall; they might be economising on their 'first look' in the knowledge that a second would be needed.

The technical work of reading one case takes relatively little time. This study has already suggested that digital reading takes substantially longer than analogue reading; it therefore seems probable that, for FFDM, CAD+SR is less time-consuming than DR.

2.2.3 Survey results

Seven readers completed questionnaires at the end of the study. These revealed some ambivalence towards CAD: although five agreed slightly that it was 'helpful overall', six did not feel it 'helped perceive additional abnormalities or change decisions to any degree'. CAD was unanimously felt to be better for calcifications than masses. Six of the seven readers would be willing to use it again. Asked for their views on the training needed to use CAD, respondents reported that the screening unit's own CAD training roller and participation in previous CAD studies were most helpful. All but one found the CAD manufacturer's training useful, while the majority felt more feedback or training rollers were unlikely to be helpful.

2.3 Impact on workload

2.3.1 Method

For this element of the study quantitative data were collected to assess the impact of CAD on (a) reading times and (b) the percentage of cases sent to arbitration and assessment. Although the numbers involved were too small to register any significant change in the detection rate, it was nevertheless felt important to track how many cancers were found.

Quantitative data were collected in two time periods.

- Between 22 August 2007 and 4 January 2008 a total of 73 clinics were observed. Each clinic was double read, giving a total of 146 observed clinics. There were missing data on timing in 30 of these, leaving a total of 116 complete measurements. During this period radiologists used the FFDM machine without CAD.
- Between 21 October 2008 and 14 March 2009 a total of 77 clinics were observed. Each clinic was double read, giving a total of 154 observations. Twenty-one of these had missing data; a further 14 were not considered because fewer than five sets of films were read (the time per case is unlikely to be measured accurately in such small clinics). Four others were considered

outliers: reading in 274 of the clinics in the study took under 1 min 45 secs per film; the four outliers were read in 2 mins 20 secs to 3 mins 55 secs. One of these was read with CAD, three without. Of the observations included in the data 91 used CAD, while 25 did not.

2.3.2 Results

Following a suggestion that second readers might behave differently from first readers, data from first and second readers were looked at separately.

Table 1 Timing data for cases read with and without CAD in 2007–8 and 2008–9. Data for first and second readers are presented separately

Reader	Period	Protocol	Number of cases	Time per case (secs)
First	2007–8	Without CAD	1711	62
	2008–9	Without CAD	185	63
	2008–9	With CAD	849	60
Second	2007–8	Without CAD	1871	57
	2008–9	Without CAD	216	53
	2008–9	With CAD	878	54

None of the observable comparisons shown in Table 1 (with CAD versus without CAD; first reader versus second reader; 2007–8 versus 2008–9) is statistically significant, and none of the possible interactions is significant.

Table 2 Outcomes for women screened using FFDM during the study period

Period	Protocol	Number of cases	Cases referred for arbitration (per 1000 women screened)	Cases referred for assessment (per 1000 women screened)	Cancers detected (per 1000 women screened)
2007–8	Without CAD	2087	156.7	49.3	6.2
2008–9	Without CAD	174	137.9	74.7	11.5
2008–9	With CAD	803	130.7	43.6	10

Analysis of the data presented in Table 2 excluded all clinics where CAD was used by one reader but not by the other. It included some clinics excluded from the analysis of timings (because timing data were incomplete or seemed unreliable) but where the use of CAD had been reliably recorded. No statistical differences were detected between clinics read with CAD and those read without. This was true for all of the outcomes measured: referral to arbitration, referral to assessment and cancers detected. It should be noted that the numbers recalled after screening because cancer is detected are small and an unreliable basis for assessing overall performance.

2.4 The cost-effectiveness of CAD

This section begins by examining the impact of CAD and of DR on recall rates and cancer detection rates, based on a systematic review and meta-analysis of the published evidence. It concludes with an analysis of CAD's cost-effectiveness.

2.4.1 Systematic review and meta-analysis

Bibliographic databases were searched for published evaluations of CAD and of DR. Only studies where CAD was introduced into routine screening and compared with SR were considered. Two researchers reviewed the evaluations of CAD and compared the results with those found in published evaluations of DR. A meta-analysis allowed a plausible range of values to be identified for the impact of CAD and of DR on the cancer detection rate and the screening recall rate.

Ten studies, with a total of 410 000 subjects, compared SR with and without CAD. A further 17, with a total of 1.2 million subjects, compared DR and SR without CAD.

Table 3 Results of meta-analyses of studies comparing single reading with (a) double reading, (b) double reading with arbitration and (c) single reading with CAD. Results are given as odds ratios with 95% confidence intervals in brackets. An odds ratio of less than 1 indicates that the intervention decreases rates, one greater than 1 that it increases rates. A confidence interval that crosses 1 indicates a non-significant (NS) effect

	Impact on cancer detection rates	Impact on recall rates
Double reading	1.10 (1.06–1.14)	1.17 (1.15–1.18)
Double reading with arbitration	1.08 (1.02–1.15)	0.94 (0.92–0.96)
Single reading with CAD	1.04 (0.96–1.13) NS	1.10 (1.09–1.12)

Meta-analysis found that DR increases both cancer detection and recall rates. However DR with arbitration increases detection rates and decreases recall rates. These are presented in Table 3 as odds ratios. They correspond to small changes in absolute risk: for example, detection rates increase by 0.44 extra cancers per thousand women screened (CI: 0.10–0.79) and recall rates fall by 0.2% (CI: 0.17–0.36). CAD does not have a significant effect on cancer detection rates and increases recall rates. Estimates of its impact on recall rates in both sets of studies vary considerably however. This systematic review is published elsewhere.[1]

2.4.2 Cost-effectiveness analysis

The calculation of CAD's cost-effectiveness is based on the balance of

- the cost of the time a radiologist who uses CAD+SR rather than DR spends or saves in screening and assessment
- the cost of CAD workstations.

This calculation takes no account of staff and consumable costs for extra assessment visits, or of the psychological effects on women of extra recalls.

The cost implications of introducing CAD were modelled for three screening centres: the median, top, and bottom of the inter-quartile range for the number of women screened in 2007/08. The same initial recall rate was used for each: 4.5%, the median national value in that year. When calculating the time saved by moving from DR to CAD+SR it was assumed that CAD has no impact either on the time taken to read films (59 secs) or on the number of cases sent to arbitration. Since the systematic review revealed wide variation in the impact of both CAD and DR on recall rates, a range of estimates were taken from Gilbert et al's account of a recent randomised controlled trial comparing DR with CAD+SR.[2] Three recall rates were used: the highest and the lowest from the three centres studied and the mean of the three. Estimates of the cost implication of additional recalls also vary. Their impact was therefore estimated under three different assumptions, with 20, 40 or 60 mins of radiologist time assigned to each extra case recalled for assessment. The results are shown in Table 4.

Table 4 Mean annual reduction in radiologist salary costs achieved by moving from DR to SR for a small, medium and large screening unit, assuming a 3%/15%/37% increase in recall to assessment and 20/40/60 min appointments for assessment. A minus sign indicates a negative saving, ie a cost increase

Size of unit	Relative increase in recall rate	Time taken per patient at assessment		
		20 mins	40 mins	60 mins
Small (12 000 films p.a.)	3%	31 516	30 701	29 910
	15%	28 002	23 567	19 262
	37%	21 522	10 410	–375
Medium (19 000 films p.a.)	3%	47 350	46 126	44 937
	15%	42 071	35 407	28 940
	37%	32 335	15 641	–563
Large (27 000 films p.a.)	3%	68 937	67 154	65 423
	15%	61 251	51 549	42 133
	37%	47 077	22 771	–820

CAD may be cost-effective as an alternative to DR if the mean annual cost is less than the amount saved through reduced use of radiologist time. The mean annual cost has been calculated here for a variety of possible configurations and the results are shown in Table 5. A comparison of Tables 4 and 5 reveals under what circumstances a CAD installation will be cost-effective. No installation will be cost-effective if the recall rate is significantly increased and recall appointments are long: for example, if four machines and four additional softcopy workstations are needed to handle throughput in a large screening unit, then either appointments need to be shorter than 40 minutes or the impact on recall rate needs to be below 15%. However there is clearly scope for cost-effective use of CAD.

Table 5 Mean annual costs of purchase plus maintenance for a variety of possible CAD installations

		Number of CAD systems purchased			
		1	2	3	4
Licences for additional softcopy viewing stations per system	0	7201	12 740	18 279	23 818
	1	10 780	19 898	29 017	38 135
	2	14 359	27 057	39 755	52 452
	4	21 518	41 374	61 230	81 085

The calculations assume a lifetime of seven years, take the Treasury target rate of 2% as their estimate for inflation, and use the recommended discount rate for future expenditure of 3.5%. Figures are presented as the average annual cost over the seven years. Figures for radiologist costs (£163 per hour in 2007/08) are taken from Personal Social Services Research Unit's *Unit Costs of Health and Social Care 2008*.[4] Calculations assume that all films are read by consultant radiologists and that there are no other costs associated with DR or with additional recalls. Costs quoted for R2 licences and maintenance were provided by Medical Imaging Systems Ltd. It is assumed that CAD implies no costs other than the initial purchase and a maintenance contract covering upgrades.

2.5 Conformance with standards

There are two sources to consider when evaluating imaging equipment. The first, IHE (Integrating the Healthcare Enterprise), is not itself a standards organisation but publishes technical profiles specifying the elements of established standards to be met to enable health providers to share information.[5] These elements are taken chiefly from the second source, DICOM (Digital Imaging and Communications in Medicine), a standard that specifies the messaging protocols and file format conventions required for inter-operability.[6] (Elements may also be used from Health Level Seven, HL7, a framework for the exchange, integration, sharing and retrieval of electronic health information.[7])

2.5.1 IHE

Digital mammography should allow a highly integrated workflow within which CAD processing is automatic and available instantly on request. The digital environment into which CAD fits includes a system to manage workflow, the CAD modality itself, the review workstation(s), the image manager (PACS) and printers. The IHE *Mammography User's Handbook*[8] identifies a number of issues relating to CAD.

a. Problems can arise if the CAD system is to be used with a review workstation from a different manufacturer. Digital mammography modalities may make images available separately 'for Processing' and 'for Presentation'. CAD will typically use the 'for Processing' images and forward its results to the review station for viewing with 'for Presentation' images. The 'for Processing' and 'for Presentation' images must be correlated; DICOM includes an attribute which allows the link to be made, but it is important to confirm that it is used.
b. Most PACS systems will allow CAD prompts to be archived. If they need this facility, users should check that the PACS system supports the DICOM *Mammography CAD SR* service class.
c. Most sites archive the 'for Presentation' images on PACS and jettison the 'for Processing' images. If the CAD results are to be regenerated, the 'for Processing' images must also be archived. ('Regenerated' here means that the images are reprocessed by CAD. As CAD algorithms are frequently updated the results may differ from those that would be obtained by retrieving archived prompts.)
d. If the Image Manager stores the 'for Processing' images then the system can be set up so that images are sent only to the Image Manager; this can then serve as the source of images for CAD processing.

2.5.2 DICOM

Customers buying CAD systems for use with PACS and Image Review workstations from other manufacturers will need to ensure that those systems comply with the DICOM standards. In particular, users will need to think carefully about their needs (in the light of the issues identified in section 2.5.1), check the DICOM conformance statements of the systems they plan to purchase, and ensure that the required DICOM service classes are supported.

The DICOM conformance statement for the R2 ImageChecker M5000 used in this evaluation confirms that it supports Digital Mammography X-ray Image Storage (for Processing and Digital Mammography X-ray Image Storage) as follows: for presentation classes and the verification class as a Service Class Provider; for Mammography CAD SR class as a Service Class User. To be integrated with a PACS system and allow the archiving of CAD prompts, therefore, the system would have to support Mammography CAD SR class as a Service Class Provider.

3. CONCLUSIONS

This evaluation found no evidence that CAD had a negative impact on the workflow of radiologists reading mammograms with FFDM. Neither the qualitative analysis nor the quantitative data on timings and referral rates indicated that CAD slows down readers or encourages them to recall more women for arbitration or referral.

Cases were double read in this study and the knowledge that each would be looked at by a second reader may have influenced first readers' response to CAD. Although the qualitative data collection took place some time after FFDM and CAD had been deployed, some participants had less experience with the system. With more exposure to it, readers would develop a clearer sense of CAD's strengths and limitations.

Using CAD with FFDM is clearly a much more straightforward process than using CAD with analogue films. This is because there is no need for digitisation or modification of the display technology. As a result, there is no discernible effect on reading time.

An earlier study by the authors, using analogue films, showed that reading time was 25 secs without CAD and 45 secs with CAD. This meant that very little time was saved by using CAD rather than DR; in fact, extra time was needed to manage the additional recalls.

The timings recorded in that study cannot be compared with the timings in this one because different recording processes were used. However the negative impact seen then on reading times was not registered in this evaluation. Nor was there any evidence here of an increase in recall rates. If this finding were robust, it would suggest that overall film reader time would be saved by replacing DR with CAD+SR.

Any decision on whether to adopt CAD as an alternative to DR must consider the possible impact on recall rates. This involves balancing the largely negative conclusions of our systematic review with the more positive conclusions of Gilbert et al's single study and the limited data collection exercise performed here. The review showed an increase in recall rates across published studies of 10% (CI 1.09–1.12) The overall recall rates in Gilbert et al were 3.4% for DR and 3.9% for CAD+SR; the difference between the rates was small but significant (RR 15%, $P<0.001$).

The impact of CAD on cancer detection remains unclear. The systematic review of published studies up to 2008 showed a non-significant increase of 4% (CI 0.96–1.13) whereas Gilbert et al found equivalence to DR in a single multicentre study. CAD seems to save radiologist time and costs unless the impact on recall rates is at the higher end of that identified in the studies seen so far, and if the additional recalls are especially costly in radiologist time. The decision to move from DR to CAD+SR should not depend solely on an analysis of the costs in radiologist time, however; it is important to consider the impact on the women attending screening of its sensitivity and specificity for cancer detection. Research suggests that a substantial fraction of women are prepared to trade the anxiety and inconvenience of an unnecessary mammogram for a higher probability of detecting cancer.[9] Where two approaches yield equal sensitivity at different rates of specificity, however, the more specific is clearly to be preferred.

REFERENCES

1. Taylor P, Potts HW. Computer aids and human second reading as interventions in screening mammography: two systematic reviews to compare effects on cancer detection and recall rate. *Eur J Cancer*, 2008, 44:798–807.
2. Gilbert FJ, Astley SM, Gillan MG et al. Single reading with computer-aided detection for screening mammography. *N Engl J Med*, 2008, 359:1675–1684.
3. Khoo LA, Taylor P, Given-Wilson RM. Computer-aided detection in the United Kingdom National Breast Screening Programme: prospective study. *Radiology*, 2005, 237: 444–449.
4. Curtis L. *Unit Costs of Health and Social Care 2008*, Personal Social Services Research Unit, University of Kent (available at http://www.pssru.ac.uk/uc/uc2008contents.htm).
5. *Integrating the Healthcare Enterprise* (available at http://www.ihe.net/).
6. *DICOM: Digital Imaging and Communications in Medicine* (available at http://medical.nema.org/).
7. *HL7: Health Care Level Seven* (available at http://www.hl7.org.uk/).
8. *HE Radiology: Mammography User's Handbook* (available at http://www.ihe.net/Resources/upload/IHE_Mammo_Handbook_rev1.pdf).
9. Ganott MA, Sumkin JH, King JL et al. Screening mammography: do women prefer a higher recall rate given the possibility of earlier detection of cancer? *Radiology*, 2006, 238:793–800.