TITLE: Full Field Digital Mammography versus Computed Radiography for Breast Cancer Screening: A Clinical and Cost-Effectiveness Review

DATE: 14 August 2008

CONTEXT AND POLICY ISSUES:

Breast cancer is the most common form of cancer that affects Canadian women.\(^1\) The Canadian Cancer Society estimates that in 2008, 22,400 women will be diagnosed with breast cancer and 5,300 women are anticipated to die from the disease.\(^1\) Early diagnosis of breast cancer can reduce mortality, improve the prognosis of individuals with the disease, and reduce the amount of surgery and chemotherapy needed for treatment.\(^2,3\) Population-based screening for breast cancer can aid in the early diagnosis of the disease.\(^2,3\)

Recommendations for population-based screening of Canadian women differ according to age group.\(^4\) In women aged 40 years to 49 years, it is recommended that they have a clinical breast examination by a trained healthcare professional at least every two years and discuss their risk of breast cancer and the benefits and risks of mammography with their physician.\(^4\) In women aged 50 years to 69 years, it is recommended that they have a clinical breast examination by a trained healthcare professional at least every two years and a mammogram every two years.\(^4\) Women aged 70 years and older should discuss with their physician how frequently they should be tested for breast cancer.\(^4\)

A number of techniques are available for breast cancer screening. Film screen mammography (FSM) is most commonly used for population-based screening for breast cancer in asymptomatic women.\(^2,3\) FSM involves using a x-ray system to image the breast and record the image on film.\(^3\) The system is specifically designed for breast imaging and uses low-energy radiation and high-contrast, high-resolution film to capture information about the morphology, anatomy and gross pathology of the breast.\(^3,5\) FSM has been shown to reduce breast cancer mortality by up to 29%, has high spatial resolution and low cost, but is not without limitations.\(^2\) Up to 20% of palpable breast cancers may not be visible with FSM. Further, there may be...
difficulties with detection of lesions in dense glandular tissue. Technical limitations of FSM include film processing, management and storage, and film artifacts.

Digital mammography, in which digital detectors are used instead of x-ray film, can also be used to image the breast for breast cancer screening. The digital detectors absorb x-ray photons and convert them to an electric charge, which is converted to an image. In full-field digital mammography (FFDM), a computer processes the signal and displays the image on a high resolution computer monitor, which can be viewed directly by a radiologist. This allows the radiologist to manipulate the image to some extent and correct for over-exposure or under-exposure, which could potentially reduce the need for recall examinations due to poor image quality. Further, FFDM systems eliminate some of the technical problems associated with x-ray film handling and may have better contrast resolution than film screen mammography, which could be beneficial in imaging dense breast tissue. Images produced using FFDM can also be printed to generate hard-copies. Computed radiography (CR) is an indirect digital mammography technique that captures the image on a reusable plate that is then scanned by a reader to produce the digital image. As with FFDM, the image can either be viewed on a monitor or printed and viewed. Often CR can be used in conjunction with existing FSM systems, which can potentially reduce costs relative to FFDM. Both CR and FFDM may be limited in availability.

There are advantages and disadvantages to both FSM and digital mammography (CR and FFDM). In deciding whether to adopt FFDM or CR for breast cancer screening, both the effectiveness and cost relative to film screen mammography are important to consider. This report will review the evidence of clinical and cost-effectiveness of digital mammography techniques relative to FSM, which could potentially help in decision-making at the level of the healthcare system.

RESEARCH QUESTIONS:

1. What is the clinical effectiveness of full field digital mammography compared to computed radiography and/or film screen mammography for breast cancer screening?

2. What is the cost-effectiveness of full field digital mammography compared to computed radiography and/or film screen mammography for breast cancer screening?

METHODS:

A limited literature search was conducted on key health technology assessment resources, including PubMed, EMBASE, The Cochrane Library (Issue 3, 2008), University of York Centre for Reviews and Dissemination (CRD) databases, ECRI, EuroScan, international HTA agencies, and a focused Internet search. Results include articles published between 2003 and July 2008, and are limited to English language publications only. A limited hand search was also conducted. Filters were applied to limit the retrieval to systematic reviews, meta analyses, health technology assessments and economic studies.

HTIS reports are organized so that the higher quality evidence is presented first. Therefore, health technology assessment reports, systematic reviews, and meta-analyses are presented first. These are followed by economic evaluations.
SUMMARY OF FINDINGS:

The literature search identified two health technology assessments, three systematic reviews and four economic analyses in which FFDM was compared to digital mammography.

Health technology assessments

In 2006, the California Technology Assessment Forum reviewed the safety and efficacy of full-field digital mammography for breast cancer screening and diagnosis. In order to be included in the assessment, screening studies had to compare the results of digital mammography (either FFDM or CR) with FSM, include at least 250 patients and report test characteristics (i.e. sensitivity, specificity) based on histologically confirmed cancer diagnoses, ideally with at least one-year follow-up from the mammogram to ensure that negative results represent true negatives. Five screening studies met the criteria for the review. One study was based in the United States (Lewin et al., 2002), two were based in Norway (referred to as the Oslo I and Oslo II studies), one was based in Japan (Yamada, et al.), and one was based in both the United States and Canada (referred to as the Digital Mammographic Imaging Screening Trial or DMIST study). Four of the five studies exclusively used FFDM, the specific device being the Senographe 2000D, while the DMIST study used four different FFDM devices (Senographe 2000D, Senoscan, Digital Mammography System, and Selenia FFDM) and one CR device (CR System for Mammography). The sample size of the studies ranged from 480 to 42,760. Where reported, the average age of the screening population was similar, but the age-based criteria for selection into the study differed. An age cut-off of 40 years and older was used in Lewin et al. and 50 years and older in the Oslo I study. In the Oslo II study, women were enrolled who ranged in age ranges from 45 to 69 years, while Yamada et al. enrolled women between the ages of 50 to 69 years. In the DMIST study, there was no restriction on age. In the DMIST study, Olso I study and Lewin et al., all women received both FFDM and FSM, while in the Oslo II study, women were randomized to one of the two imaging techniques. Details of the methodology used in Yamada et al. were not clearly reported in the review, but it appeared that all women received both imaging techniques.

No attempt was made to pool or meta-analyze data across studies. A definition of terms can be found in the Appendix. Lewin et al. found that FFDM was less sensitive, was more specific, had a lower recall rate than FSM, and had a similar area under the receiver operator characteristic (ROC) curve. In the Oslo I study, FFDM was less sensitive, less specific, and had a higher recall rate than FSM. In the Oslo II study, no significant differences were found in the breast cancer detection rate, recall rate, and positive predictive value of the two diagnostic techniques. In the DMIST study, the area under the ROC curve did not differ between digital mammography (FFDM and CR combined) and FSM, nor did the sensitivity and specificity of the two techniques differ. In women under the age of 50, those with dense breasts, and those who were premenopausal or perimenopausal, digital mammography was superior to FSM according to the area under the ROC curve. In women under the age of 50, sensitivity was greater with digital mammography than with FSM. Specificity of the two imaging techniques did not differ in these sub-groups. The authors of the review felt that methodological limitations of Yamada et al. precluded interpretation of its results, but stated that the recall rate was higher with digital mammography.

The authors stated that in the absence of appropriate attention to training and the work environment, digital mammography could potentially be less sensitive and less specific FSM, but with careful training and quality control, digital mammography should be equivalent to FSM.
The authors of the report concluded that the DMIST study definitively demonstrated that digital mammography was equivalent to FSM for breast cancer screening in asymptomatic women, although FFDM was shown to be better than FSM in certain subgroups of women. This was the only study of the five studies that included Canadian women. The authors considered the DMIST study to be the most methodologically sound and seemed to draw conclusions mainly from its findings.

In Canada it is recommended that women between the ages of 50 to 69 years have screening mammography every two years. Three of the five studies enrolled women younger than 50 years of age, an age group where routine screening mammography is not recommended in Canada. Thus, it is not clear if the results of these studies would be entirely relevant or generalizable to the Canadian screening population. Further, the DMIST study found that digital mammography was superior to FSM in women under the age of 50 years. Again, given Canadian screening recommendations, it is not clear if this observation would be relevant in Canada. Finally, if the prevalence of breast cancer in Canada differs from those countries in which the technologies were evaluated, some measures of diagnostic performance may differ from those observed in the studies.

In 2007, the ECRI Institute issued an emerging technology evidence report on FFDM for breast cancer screening. In order to be included in their review of clinical effectiveness, studies had to be published in full (i.e., not as abstracts) and in English. Further, the studies had to enroll a large number of women from the general population, use both FFDM and FSM in all participants, and have multiple radiologists interpret the results. Three studies were identified that met the criteria for the review, including the DMIST study, Oslo I, and Lewin et al. The authors of this review felt that these three studies did not provide reliable sensitivity, specificity, and positive predictive value statistics. This was because, due to ethical considerations, biopsies to definitively diagnose breast cancer were only performed in women whose mammogram was suggestive of the disease, and therefore results may be subject to verification bias. They felt that the cancer detection and recall rates were the most appropriate for comparing the clinical utility of the two technologies and only reported on these outcomes. Overall, in the DMIST study, detection rates with FFDM and FSM did not differ, but were higher with FFDM in women under the age of 50, those with dense breasts and those who were premenopausal or perimenopausal. The detection rate of the two technologies did not differ in the Oslo I study or in Lewin et al. The recall rate with FFDM was higher than FSM in two of the three studies. In terms of clinical effectiveness, it was concluded that there were no statistically significant differences in overall breast cancer detection rates between with FFDM and FSM in breast cancer screening in the general population, but that FFDM was more effective at detecting cancers in women under the age of 50 years, those with dense breasts, and pre-menopausal or peri-menopausal women. It was also concluded that the higher recall rates associated with FFDM could add to the costs associated with its use.

Systematic reviews and meta-analyses

In 2006, the Blue Cross Blue Shield Association Technology Evaluation Center updated a previous assessment of FFDM in both screening and diagnosis of breast cancer. The objective of the assessment was to answer the question of how FFDM compares to film screening mammography for cancer detection and avoidance of follow-up testing. In order to be included in the review, studies had to compare the results of FFDM to FSM on the same patients or in two groups of patients for either screening or diagnosis of breast cancer. The studies also had to report on cancer detection based on histology in at least some of the patients. Studies were excluded if they used simulated breast tissue or had been included in the
previous assessment of FFDM. Three studies were identified, two of which were screening studies (the DMIST study\textsuperscript{13} and the Oslo II study\textsuperscript{14}). From these studies, it was concluded that FFDM was as accurate as FSM in all women, but more accurate in those who were under the age of 50 years, had dense breasts, and were premenopausal or perimenopausal.

In 2005, Elmore \textit{et al.}\textsuperscript{19} systematically reviewed breast cancer screening and examined evidence about new methods of breast cancer screening, including FFDM. The inclusion criteria for the review were not outlined in the report. Lewin \textit{et al.},\textsuperscript{12} Oslo I,\textsuperscript{15} and Oslo II\textsuperscript{14} studies were included in the review, the methodology and results of which have been previously described. The authors did not draw any conclusions specific to FFDM.

In 2004, Irwig \textit{et al.}\textsuperscript{20} systematically reviewed the accuracy of new technologies that had been proposed for breast cancer screening, one of which was FFDM. In order to be included in the systematic review a number of quality and applicability criteria had to be met. As well, the study had to report a measure of true or false positive rate. Lewin \textit{et al.}\textsuperscript{12} was the only study identified in which FFDM was assessed, the results of which have been previously described. From this study, the authors of the review concluded that FFDM demonstrated an incremental improvement in sensitivity relative to FSM, but felt that this may not translate into an absolute benefit in practice. Further, overall, the authors concluded that larger and better quality studies were needed.

**Economic evaluations**

Tosteson \textit{et al.} (2008)\textsuperscript{21} assessed the cost-effectiveness of digital mammography for breast cancer screening in women in the United States who were aged 40 years or older in 2000, using a computer-based model. The model included performance and resource use data from the DMIST study\textsuperscript{13} and incorporated data pertaining to the natural history of breast cancer, breast cancer detection, breast cancer treatment, and competing-cause mortality. The economic analysis compared a number of different strategies: film mammography for women of all ages; digital mammography for women of all ages; age targeted screening (digital mammography for women <50 years of age and FSM for women ≥ 50 years of age); age and density targeted mammography (digital for women <50 years of age or those ≥ 50 years of age with radiographically dense breasts and film for those who do not fall into these categories). A subgroup analysis was also performed for women age 65 years or older (the Medicare population). Costs and quality adjusted life years (QALYs) were simulated for each screening scenario. Screening patterns were assumed to be similar to those observed for mammography in the United States. A life-time time horizon was used and the analysis was performed using societal and Medicare perspectives.

In the population aged 40 and older, the cost per QALY gained with all-digital mammography screening was $331,000 (95% CI: $268,000 - $403,000) compared to all-film mammography screening. When age targeted digital mammography screening was used, the costs per QALY gained was $26,500 (95% CI: $21,000 – 33,000) relative to all-film mammography. For age and density targeted digital mammography, the cost per QALY gained was $84,500 (95% CI: $75,000 - $93,000). In women aged 65 years and older, density-targeted digital mammography cost $97,000 per QALY gained (95% CI: $77,000 - $131,000) relative to all-film mammography. All-digital mammography was dominated by all-film mammography in that it was more costly and equally as effective in women aged 65 years and older. The authors identified a number of limitations in their model, including the use of Medicare standardized payment amounts to estimate mammography cost, failure to include any additional costs from determining breast density prior to screening, a somewhat arbitrary cut-point for the age targeted screening (the
age of 50 was used as a surrogate marker for menopause), and they did not evaluate the cost-effectiveness of screening strategies that combined film and digital mammography. Overall, the authors concluded that age targeted digital mammography screening was the most efficient approach to provision of digital mammography screening in the U.S. population. Given that the model used US-based inputs (e.g., costs, screening patterns), the results are not clearly generalizable to the Canadian population.

In 2008, the National Health Service Centre for Evidence-based Purchasing assessed the cost-effectiveness of FFDM and CR relative to FSM imaging in breast cancer screening. In the analysis it was assumed that all technologies were clinically equivalent, so only costs were considered. This was a United Kingdom-based analysis. A number of assumptions about staffing, salaries, image storage, workflow and workload were made in order to estimate costs. Costs for the technologies were estimated over a seven year life-cycle (referred to as whole life costs). Different cost estimates were made for static screening sites and mobile screening sites, as a result of the differences in equipment requirements and workflow. Staff, consumables, purchase, and service costs were included, as well as resource costs specific to the type of screening site. The sensitivity analyses were also performed in which capital costs, staff costs, service charge and utility costs, consumable costs, life-cycle and appointment time were varied.

For static screening sites, it was found that the cumulative seven-year costs for FFDM and CR were similar (around £2.0 to £2.2 million), but lower than FSM (around £2.7 million) for an average workload. Further, FSM was found to be the least cost-effective and CR to be the most cost-effective technique for all likely workloads at a static screening site. Similarly, for mobile screening sites, seven-year cumulative costs were around £2.7 million with FSM and around £2.0 million for both FFDM and CR. FFDM had a slightly higher cost than CR, which also had the lowest cost across all workloads. Sensitivity analyses indicated that the cost of consumables had the greatest effect on the relative cost-effectiveness of the technologies. Variations in the other factors did not significantly impact the results of the analysis. The authors concluded that CR and FFDM were cost-effective relative to FSM and that CR was the most cost-effective technology for screening sites, but only by a small margin compared to FFDM. The authors pointed out a number of limitations in their evaluation, including that the assumptions used in costing may not hold for all circumstances and that they did not include costs that would be incurred in changing over from FSM to FFDM or CR and costs associated with system breakdowns. Due to differences in Canada’s and the United Kingdom’s healthcare system (such as salaries and work-flow) and costs associated with the technologies themselves, it is not clear if these results would be generalizable to the Canadian context.

Ciatto et al., (2006) assessed the differential costs of full-field digital mammography and FSM using real world expenditures from a centre in Italy. Purchase costs of the equipment, supplies (e.g., film, film developer), work time of a radiographer and radiologist, and costs associated with archiving were included in the assessment. Differential costs were compared for two average workload levels: 5000 tests or 10,000 tests per year for each mammography unit. The perspective of the analysis was not stated, but it appeared to be that of the health-care payer. The cost per test with FFDM was €24.22 and €9.91 greater than FSM under the 5000 and 10,000 test scenario, respectively. The differential cost of full-field mammography increased by about €1 to €2 if all patients’ images were laser printed. The authors concluded that as there was no evidence that FFDM was superior to FSM in clinical effectiveness, the additional expense associated with FFDM may not be justified. Given that this study was based upon Italian costs, it is not clear that these results would be generalizable to the Canadian health care system. Further, it was not clear if the technologies were used for screening or diagnosis of breast cancer.
In 2004, the National Health Service Breast Cancer Screening Programme compared the cost of FFDM to FSM in breast cancer screening. Direct costs (in 2002 values) were included in the model (equipment, consumables, storage and staff workload). The data sources for the costs were two pilot studies of two different FFDM systems and were calculated for screening of 10,000 women annually. FSM was estimated to cost £113,700 per year for screening of 10,000 women, compared to £231,578 for FFDM. However, if FFDM images were not produced in hard copy, the cost dropped to £157,623 annually. Further, it was estimated that if throughput was increased the number of FFDM machines could be reduced by 20%, and if hard copy images were not produced, the annual cost of screening 10,000 women with FFDM would be reduced to £128,345. Overall, the authors concluded that the initial cost of FFDM was higher than FSM, but longer term costs may be similar if hard copy FFDM images were not printed and fewer FFDM systems were required. However, they also emphasized that the true costs of running FFDM systems remained uncertain. The authors pointed out a number of limitations in the analysis, including uncertainty in estimates of equipment, maintenance contracts and archiving costs. The authors felt that the true costs associated with FFDM have been underestimated, due to uncertainty in the costs of picture archiving and communication systems and radiographic information systems. Again, due to differences in Canada’s and the United Kingdom’s healthcare system (e.g. salaries, work-flow) and costs associated with the technologies themselves, it is not clear if these results would be generalizable to the Canadian context.

**Limitations**

While a number of health technology assessments, systematic reviews, and economic analyses were identified that compared digital mammography to FSM, there are still limitations to this body of evidence. One of the reviews of clinical effectiveness was published prior to several of the large screening studies. While their conclusions were based upon the best available data at the time, subsequently published studies were not consistent with earlier studies. As such, the relevance of the conclusions of these early reviews is questionable. One review did not draw any conclusions at all, so did not provide particularly useful information. Further, there were no comparisons of the clinical effectiveness of FFDM and CR; thus, this question remains unaddressed. As well, in the largest and most methodologically rigorous study included in the reviews of clinical effectiveness (the DMIST study), there was no distinction made between FFDM and CR. These technologies were simply analyzed together as the digital mammography group. It is not clear if this was appropriate to do so. The most recent study of clinical effectiveness of the imaging technologies was published in 2005. Thus, it is possible that advances in the imaging technologies may have occurred in the past three years. In terms of cost-effectiveness, only one study (based in the United Kingdom) compared the costs of CR to FFDM and it is not clear in the results would be generalizable to Canada.

**CONCLUSIONS AND IMPLICATIONS FOR DECISION OR POLICY MAKING:**

Overall, FFDM and FSM appear to be equivalent in terms of clinical effectiveness for breast cancer screening in the general, asymptomatic female population. In fact, this assumption was applied to the majority of the economic analyses, which chose only to compare the relative costs of the technologies. However, based upon evidence of the largest screening study, FFDM may be superior in several subgroups of women including women under the age of 50 years, those with dense breasts, and pre-menopausal or peri-menopausal women. The relevance of these findings are not entirely clear given that, in Canada, general screening is not recommended in women under the age of 50 years and in terms of logistics, it is not clear how
women with dense breasts would be identified prior to screening. As such, while age-targeted screening for breast cancer appeared to be cost-effective\textsuperscript{21} in the United States, this observation may not be particularly relevant for decision-making in Canada. Based upon single year cost estimates in the general screening population (i.e, not in any particular subgroup),\textsuperscript{9,23,24} FFDM would appear to be more costly. However, United Kingdom-based seven-year cost estimates favour both FFDM and CR over FSM at static and mobile screening sites.\textsuperscript{22} Thus, it is difficult to draw conclusions about the relative costs of the technologies as it varied when different time horizons were used. Conclusions about the clinical effectiveness of FFDM relative to CR could not be drawn based upon the available literature, but the technologies appeared to be relatively similar (slightly favouring CR over a seven-year period) in terms of cost.

Although it appears that the clinical effectiveness of the various techniques for breast cancer screening are similar, further research relevant to the Canadian guidelines for screening would be helpful to aid in decision making. In addition, research comparing CR to FFDM is lacking and these studies would be required to determine which of these methods should be used for breast cancer screening. The costs of the different techniques, including equipment and staffing costs, will have to be considered when deciding which method to use for breast cancer screening in Canada.

**PREPARED BY:**
Ron Pohar, BScPharm
Melissa Severn, MLIS
Health Technology Inquiry Service
Email: htis@cadth.ca
Tel: 1-866-898-8439
REFERENCES:


APPENDIX 1: Definitions

Sensitivity: Proportion of people with the target disorder who have a positive test (true positives).

Specificity: Proportion of people without the target disorder who have a negative test (true negatives).

Positive Predictive Value: Proportion of people with a positive test who have the target disorder.

Negative Predictive Value: Proportion of people with a negative test who do not have the target disorder.

Area under the receiver operating characteristic curve (AROC): measure of the accuracy of a diagnostic test in terms of discrimination or the ability of the test to correctly classify those with and without the disease. A value of 1 would mean that the test is perfect.

Diagnostic Accuracy or Probability: The proportion of all tests that have given the correct result.