# HEALTHCARE TECHNOLOGIES AND PROFESSIONAL VISION

ROB PROCTER, MARK HARTSWOOD, ROGER SLACK, JAMES SOUTTER, ALEX VOß
*School of Informatics, University of Edinburgh, UK*

MARK ROUNCEFIELD
*Department of Computing, University of Lancaster, UK*

**ABSTRACT:**
This paper presents some details from an observational evaluation of a computer assisted detection tool in mammography. The use of the tool, its strengths and weaknesses, are documented and its impact on reader's 'professional vision' (Goodwin 1994) considered. The paper suggests issues for the design, use and, importantly, evaluation of new technologies in everyday medical work, pointing to general issues concerning trust – users' perception of the dependability of the evidence generated by such tools and suggesting that evaluations require an emphasis on the complex issue of what technologies afford their users in everyday work.

## INTRODUCTION:

This paper presents findings from an ethnographic evaluation of a computer aided detection (CAD) tool in mammography in order to consider some issues of the relationship between innovative healthcare technologies and 'professional vision' [1]. As Goodwin notes, it is through coding, highlighting and 'producing and articulating material representations' which participants build and contest professional vision - socially organised ways of seeing and understanding. In investigating the impact of technological interventions on everyday 'professional vision' in mammography we are concerned to investigate the effects on reader performance, including how the results provided by these systems are made sense of and deployed in everyday work. Our interest then is the important topic of how new and developing technologies fit into or transform existing professional practices. We use observational studies to examine how technologies and the practices they inform or resource, get absorbed, modified or changed through processes of 'innofusion' and 'domestication' [2] whereby the technology is effectively 'made at home in the world'.

Professional vision concerns being a competent practitioner, distinguishing - at a glance or through various artful practices - between degrees of 'normal' and 'abnormal'. In this paper the actual practices of 'reading' mammograms are highlighted and analysed as aspects of 'professional vision'. Our interest lies in explicating the varied worldly techniques whereby objects are identified, viewed, interpreted and categorised as 'suspicious' or 'not suspicious' - explained away or seen as in need of further examination before adequate explanation can be produced. Our interest is in the interaction between the CAD (computer assisted detection) technology and these 'ways of seeing' and whether and in what ways readers come to trust or

distrust the machine, how the technology impacts on ‘incongruity procedures’, calculation and interactional practices..

**MAMMOGRAPHY:**

Breast cancer accounts for one-fifth of deaths from cancer among women in the UK. The goal of the screening programme, based upon mammography, is to achieve a reliable cancer detection rate. In breast screening one or more X-ray films (mammograms) are taken of each breast by a radiographer and examined for evidence of abnormality. Evidence suggests that readers’ performance can be improved by computer-aided detection of target features [3]. The principle is to apply image analysis techniques to identify target features in each mammogram and draw these to the reader’s attention through the use of prompts. The prompts act as an attention cue, and so counteract the effects of variability in concentration and, more generally, make the visual search pattern more systematic and complete. A number of prompt-based, computer-aided detection tools have now been developed but the practical realisation of improved reader performance is not easy [4; 5].

**THE CAD SYSTEM IN USE:**

The system whose trials we observed consisted of a scanning and processing unit with a LCD display and a film viewer with two built in CRT displays to show the prompts. The system targets ill-defined and speculated lesions and micro-calcifications. Calcification clusters are marked by a shaded triangle. Ill-defined lesions are marked with an asterisk and a circle is drawn around either prompt to indicate confidence in detection. The system does not perform a comparison between left and right views (i.e. asymmetry). In the first phase of the trial the system was not used directly by the readers - instead booklets of mammograms were prepared, using historical film, containing sets of prompted cases. Readers were shown the appropriate film - CC’s and Obliques, (but not previous films or notes) - and asked to indicate areas of concern on the booklet and make a decision as to whether the case should be recalled. Before the tests each reader was given an explanation of how the system worked, emphasising that the machine was for detection not diagnosis and that it 'spotted' masses and calcifications and the appropriate prompts. They were also advised that the sensitivity of the machine had been set such that there would inevitably be a lot of false prompts; and warned that since this was a test set there would be more cancers than in 'normal', everyday readings.

This paper is concerned with only a part of the reading process - which we have commented on in more detail elsewhere [6]. In this instance, the actual practices of 'reading' and the interaction with the CAD systems prompts are highlighted and analysed as aspects of 'professional vision'. Our interest lies in explicating the varied worldly techniques whereby objects are identified, viewed, interpreted and categorised as 'suspicious' or 'not suspicious' - and the role that computer aided detection plays, or might play, in this process.

**CAD IN ACTION: STRENGTHS AND WEAKNESSES OF THE TECHNOLOGY:**

In the first phase of the trial the system was not used directly by the readers - instead booklets of mammograms were prepared containing 60 prompted cases. During the tests readers were shown the appropriate film - CC’s and Obliques, but not previous films (or any notes) - and asked to indicate areas of concern on the booklet mammograms and make a decision as to whether the case should be recalled. As part of the ethnographic evaluation

readers were observed doing the various test sets and then asked about their experiences of using the prompts. Readers were also taken back to cases identified in the test set where they had appeared to have had difficulty or spent a long time making their decision and asked to talk through any problems or issues to do with the prompts and their decisions. Although there were inevitably some variations in how readers approached a reading and the test, the fieldwork extract below gives some idea of the process observed:

Simplified fieldwork extracts:

Case 1: Gets blank film to mask area of the film ("so I can concentrate on it" .. these are set up differently from the way I usually look at them .. so I have to train my eye each time ."). Using magnifying glass. Marking on booklet. Looking from booklet to scan. Homing in on an area - "I'd say its benign"

Case 2: Using blank film. Takes film off roller and realigns. Magnifying glass. Looking from booklet to film. "I'd not recall .. what the computer has picked up is benign .. it may even be talcum powder."

Readers identified a number of strengths of the CAD system in supporting their work through its ability to identify subtle signs - signs that some readers felt they might have missed. The technology thereby stimulated interaction by prompting readers to re-examine the mammogram - making them look twice (and in some instances affecting their decision). As one reader said:

"Those micros that the computer picked up .. I might have missed it if I was reading in a hurry .. I'd certainly missed them on the oblique.." "This one here the computer certainly made me look again at the area.." "I thought they were very useful, they make me look more closely at the films .. I make my own judgement .. but if the prompt is pointing things out I will go and look at it again but I make my own judgement .. it might bring up abnormalities that I haven't seen whether they're benign or .. nothing but its still nice to have a prompt so that I can look again.."

Or another reader:

".. a few little densities have been marked by CAD .. its taking me back to this which I've already just had a look at .. I think .. I think probably we'd have another look at it .. really.. CAD's pulling me toward it.. its picked it up .. I think we'd better have another look .. its really an asymmetry.."

Admitting that readers, as humans, were liable to have 'off' days, consistency was also seen as a strength of the technology, in a phrase strangely redolent of the 'Terminator':

" .. its just the fact that its more consistent than you are .. because it's a machine.."

The system was also regarded as useful in compensating readers' individual deficiencies and reminding them of 'good practice'. Readers frequently express the opinion that they are better at 'spotting' some cancers - as having skills or deficiencies in noticing particular types of object within films. This was another area where the CAD prompts were seen as useful, as both compensating in some (consistent) way for any individual weaknesses of the reader and as a reminder of 'good practice':

"My approach tends to be to look for things that I know I'm not so good at ... there are certain things that you do have to prompt yourself to look at, one of them being the danger areas."

"I do ..I have areas where I know I'm weak at seeing .. you know ones that you've missed .. one is over the .muscle there .. its just because the muscle is there .. if you don't make a conscious effort to look there you tend not to see that bit of breast . and the other area is right down in the chest wall - breast and chest wall area .. because in older women the cancers tend to be in the upper outer quadrant so I look in that area very carefully .. it depends on the type of breast really .. with some breasts they're very fatty and you can just look and not see something but in other breasts .. so I try to look at the whole film, because I know if I just glance at it and don't make that conscious effort I don't look .. and also these rather big breasts where all the breast tissue gets scrunched up .. that's sometimes difficult.."

" I'd made up my mind about where the cancer was .. but I was looking at all these other areas ..because one has to look at the other breast from experience because one has to look for the second cancer that maybe difficult to see .. and also you're looking for multi-focal cancer .."

Finally the prompts were seen as useful in confirming readers in their detection and (strangely) diagnosis, both in terms of the presence and absence of anything 'seeable' in the mammogram. This is illustrated in the following comments; firstly about the presence of suspicious features in the mammogram:

"CADs marked that and I agree"

".. calcification on the left .. it looks very suspicious .. yeah, CAD's marked it as well .. bring it back.."

And secondly, concerning the absence of prompts supporting the readers initial judgement that there is nothing suspicious or worthy of recall:

"..CAD's not marked them .. and I don't think they're anything either.."

".. I don't think that's anything at all .. its glandular tissue .. there's a little asymmetrical density here but you often get that .. interestingly CAD's not marked it .. so it obviously thought so too .. it's a bit of benign calcification.."

".. glandular tissue on the right .. but I don't think there's anything there … (looks at booklet) .. CAD isn't interested either.."

".. on first look I don't see anything .. (looks at booklet) .. and not does CAD .. on second look I still don't see anything.."

Amongst the weaknesses identified by readers was the distracting appearance of too many prompts:

"this is quite distracting .. there's an obvious cancer there (pointing) but the computer's picked up a lot of other things.." "there's so many prompts .. especially benign calcifications .. you've already looked and seen there are lots of benign calcs.."

The system was also seen to regularly prompt the 'wrong' things, that is benign features or artefacts of the film process:

".. what the computer has picked up is benign .. it may even be talcum powder.." 'I'm having trouble seeing the calc its picked up there ..(pointing) . I can only think its an artefact on the film (a thin line at the edge of the film)"

At the same time the system was seen to be missing obvious prompts that raised wider issues to do with getting used to, trusting and 'understanding' the machine:

"That's quite a suspicious mass on the CC ..I'm surprised it didn't pick it up on the oblique.." (Points to area) "I'm surprised the computer didn't spot it .. its so spiky .. I'd definitely call that back.." "I'm surprised the computer did'nt pick that up .. my eye went to it straight away.."

Getting used to the machine - making it at home in a working process - is important. For as readers adjusted to the strengths and (particularly) weaknesses of the technology there was some evidence of ways in which it might come to be incorporated into their professional repertoire. As one reader commented:

.."..I've seen that before.. marking near the muscle and the edge of the film .. CAD's marked something that's not real.."

The prompting algorithms in the system cause it both to prompt features that are not cancers, as well as miss features that may be obviously cancers because they are either under 10mm or over 20mm - ie. Large, obvious cancers might be missed. In addition, normal features in the breast such as calcified arteries or crossing linear tissues can be prompted as micro-calcifications while other normal features such as ducts and tissue radiating from the nipple or inadvertent crossing of parenchymal tissue can produce a prompt for a cancerous mass.

(looking at booklet) .. ".. I've been told to look in places I think are normal really .. (looks again) .. that are normal.."

There is a sense in which might be said that the machine works too well, providing not just too many prompts but prompting features that a skilled reader would not accept as 'promptable'. Of course it was designed to work in exactly this way, though readers engrossed in their work

soon forgot this. There is also a sense in which readers had too high expectations of the technology - expecting it to prompt for abnormalities - such as asymmetry - that were not part of the technology's repertoire. In part this was a feature of the technology that the readers (at least in this trial) effectively 'forgot' but which might be incorporated into the readers' 'recipe knowledge' in time.

## MAMMOGRAPHY AND PROFESSIONAL VISION:

'The relevant unit for the analysis of the intersubjectivity at issue here is thus not these individuals as isolated entities but (. . .) a profession, a community of competent practitioners, most of whom have never met each other but nonetheless expect each other to be able to see and categorize the world in ways that are relevant to the work, tools, and artifacts that constitute their profession' [1: 615].

The situated practical actions of reading of mammograms constitute the radiologists professional vision as a "socially organized ways of seeing and understanding events that are answerable to the distinctive interests of a particular social group" [1: 606 ]. Here we are concerned with Goodwin's (1994) notion of 'professional vision' and the role of technology - like CAD - in supporting, sustaining or undermining it.

"Discursive practices are used by members of a profession to shape events in the domains subject to their professional scrutiny. The shaping process creates the objects of knowledge that become the insignia of a profession's craft: the theories, artifacts, and bodies of expertise that distinguish it from other professions." (ibid.)

The discursive practices identified and discussed by Goodwin are:

- coding practices, i.e. ways in which phenomena are transformed into graspable concepts or categories in terms of the profession's discourse,
- highlighting, i.e. the marking of some phenomena as more salient than others, again in terms of the framework of the profession's discourse,
- producing and articulating material representations of real phenomena.

Such professional vision is always selective vision, "the ability to see relevant entities" [1: 626]. The coding, highlighting and representational practices are directed at a selection of the phenomena observable. Goodwin (1994) documents how archaeologists learn to look for small differences in the colour of dirt on patches of soil, because these may represent traces of archaeologically relevant items. In mammography readers exercise a combination of perceptual skills to find what may be faint and small features in a complex visual environment, and interpretative skills to classify them appropriately -- i.e., as benign or suspicious. Several types of mammogram features are early indicators of breast cancer, for instance:

- Micro-calcification clusters are small deposits of calcium visible as tiny bright specks.
- Ill-defined lesions are areas of radiographically-dense tissue appearing as a bright patch that might indicate a developing tumour.
- Stellate lesions are visible as a radiating structure with ill-defined borders.
- Architectural distortion may be visible when tissue around the site of a developing tumour contracts.
- Asymmetry between left and right mammograms may be the only visible sign of some features.

Professional vision is a way of looking and seeing - a technique for making work relevant features available, making them stand out from the rest of the field and accounting for just

what they are in and as a part of this or that domain of scrutiny. In mammography readers learn how to find the features on the mammogram and interpret them. In their everyday work readers employ a number of techniques for making features visible: this includes using the light box or magnifying glass; adopting particular search patterns:

"Start at top at armpit..come down ..look at strip of tissue in front of armpit..then look at bottom .. then behind each nipple .. the middle of the breast.."

Readers also attempt to 'get at' a lesion by measuring with rulers, pens or hands from the nipple in order to find a feature in the arc; comparing in the opposite view; aligning scans; looking 'behind' the scans; 'undressing lesions' by tracing strands of fibrous tissues into and out of the lesion area and so on. The concept of 'undressing' a suspicious object is an important aspect of the practical constitution of professional vision in reading. Overlapping breast tissue can often mimic the appearance of stellate lesions or distortions in the film. Readers do not necessarily take the 'spiky' appearance of an object as being indicative of a recallable object. Rather, they will 'trace' the lines through the object: if the lines are 'drawn to' the object as opposed to find-able through it then this is an instance of a suspicious or malignant object. This 'undressing' or 'picking the lesion to bits' involves divorcing the object from other components that apparently make it up and examining if the constituents are found continued relevantly elsewhere in the proximity. This figure ground relationship is made in the embodied practices already mentioned such as tracing the lines of suspicious objects with pens and fingers. Readers also make use of 'worldly interpretations' [7] of the significance of the object - through ideas about 'territories of normal appearance' and 'incongruity procedures' [8] to spot abnormalities within the films they read. Thus the positioning of an object in a particular area of the breast renders it more suspicious than if it had been elsewhere.

Professional vision also (and perhaps paradoxically) involves the ability *not* to see things - or, perhaps more accurately the ability to see and dismiss things (in fact the vast majority of things) seen on the mammogram as artefactual or irrelevant. The capacity to eliminate certain aspects from the category of potentially relevant objects called 'data' is as much an ingredient of professional vision as the inclusion of other aspects into that category. Professional vision teaches us to see certain things, and simultaneously not see other things in the same professionally relevant way,

However, we have some reservations as to the notion of 'professional vision' as presented by Goodwin (though one does not have to endorse Goodwin's presentation of 'professional vision' to see its value as an analytic tool, especially for design). In some ways Goodwin's articulation of 'professional vision' reifies 'ordinary vision' (in spite of Goodwin's avowed intentions). We refused to be seduced by the notion of 'professional' in 'professional vision'. To be sure a category owned knowledge exists in domains such as mammography but it is learnable because it is irredeemably grounded in the everyday. This is not to say that it is only common sense but that it has an anchorage in features of everyday working life - some of which we have documented here. Our interest in this paper is not to critique the idea of professional vision but to use the concept to examine the interaction between the technology and these 'ways of seeing' and what technology can afford readers. The prompting system should not be taken to make things less uncertain – decisions still have to be made and these fall to the readers. The prompts are docile in that their character is simply to prompt as opposed to say what should be done. In the above fieldwork we see that readers attempt to

ascertain what a prompted feature is. That a prompt occurs is a meaningful thing, but what to do about it is still a reader's matter. The system still requires the professional vision of the reader to render prompts as what they accountably are. Professional vision in this sense is both a form of calculation work and an aspect of orienting to the 'rules' of reading and interpreting mammograms in a way that the CAD system does not, even though it may appear to. It might (arguably) be suggested that both the CAD system and the reader are involved in 'calculation' work? In the case of the system its algorithm shapes whether or not it 'sees' something that merits prompting. For the reader, however, the issue points to the routine work of deploying and displaying a system of rational calculability, making a system of calculability operate in order that the routine achievement and display of 'proper calculation' be displayed. Such calculation work is shaped by an awareness of everyday working contingencies such as talcum powder or dense tissue. It requires an understanding of how the 'formulas' for breast cancer might be applied in specific cases? - a sensitivity towards the criteria of correctness and requirements for its satisfaction. What are the determinants of its applicability? What are the requirements of making it work? The point is to arrive at some efficient and reasonable, defeasable estimation of 'how things stand'

- 'if you know they are on HRT for instance you might accept patches in one (film) where you wouldn't accept in another'

## CONCLUSION - EVALUATING CAD TOOLS IN MEDICINE:

We are interested in investigating the impact of technological interventions on everyday medical work. In this paper, as well as extending our understanding of reading practices and considering usability issues for deployment we are also concerned to investigate some of the effects of CAD tools on reader performance, including readers understanding of the CAD tool's behaviour. This illustrates the general point of how, in areas such as medicine, the introduction of computer-based detection and diagnosis aids highlights and problematises how the results provided by these systems are made sense of and deployed in everyday work. A particular interest is in evaluating the effectiveness of the CAD system in identifying and prompting for what might be characterised as 'difficult cases'. An argument for the potential of CAD systems can be made on the diversity of the contributions brought to bear on decision-making by the reader and CAD system acting in concert. This may be particularly relevant if the CAD system can assist with what readers might regard as 'difficult cases'. If prompting by the CAD system only occurs for cases the reader is very unlikely to overlook then some might suggest (controversially) that the lack of any obvious performance gain makes such tools of questionable value.

The paper highlights some important questions for the design, use and evaluation of new technologies in everyday medical work. Medical systems are often found wanting in everyday, real time, real world use (e.g.,[ 9]). How readers use prompts to inform their decision-making, and how they make sense of the system's behaviour, may be important for maximising effectiveness [5; 10]. In this case readers rationalised 'false' prompts by devising explanations or accounts of its behaviour that were grounded in the properties of the mammogram image - eg that it was talcum powder, or an artefact of the developing process. This points to general issues concerning trust – users' perception of the dependability of the evidence generated by such tools -- and how trust is influenced by users' capacity for making sense of how the tool behaves. Consequently, we suggest evaluations require an emphasis on the complex issue of what technologies afford their users and not on issues or approaches that

elide the lived work of using technologies. Our very preliminary evaluation of the CAD system raises a number of questions concerning the appropriateness of quasi-clinical trials for technological innovations. In particular, the artificial character of the trial, divorced form the lived reality of everyday reading work and the various affordances of the worksite casts some doubt on the value of such tests for the deployment of the technology in a (very different) real world setting. As Heathfield [11; 12] argues we have inadvertently created:

"a catch 22 situation whereby we cannot move forward with information technology in health care because of the lack of evaluation, yet our failure to deploy systems in routine clinical practice and allow them to mature means that we have nothing of any substance to evaluate."

## REFERENCES


[1]. Goodwin, C. (1994). Professional Vision. American Anthropologist. 96; 606-633.

[2]. Williams R., Slack, R. and Stewart, J. (2000). Social Learning in Multimedia, Final Report to European Commission, DGXII TSER, University of Edinburgh.

[3]. Hutt, I. (1996). The Computer-Aided Detection of Abnormalities in Digital Mammograms. Unpublished Ph.D. Thesis, Manchester University.

[4]. Hartswood, M., Procter, R., Williams, L. and Prescott, R. (1997). Subjective responses to prompting in screening mammography. In Proceedings of the Medical Image Analysis and Understanding Workshop, Oxford, July.

[5]. Hartswood, M., Procter, R. and Williams, L. (1998). Prompting in practice: How can we ensure radiologists make best use of Computer-Aided Detection Systems? In Karssemeijer, N. et al. (Eds.) Proceedings of the Fourth International Workshop on Digital Mammography, Nijmegen, Netherlands, June 7th-10th. Kluwer Academic Publishers.

[6]. Hartswood, M, Procter, R. and Slack, R. (2001). Performance Management in Breast Screening: A Case Study of Professional Vision and Ecologies of Practice. Submitted to the Journal of Cognition, Technology and Work.

[7]. Driessen, J. (1999) Worldly Interpretations of a Suspicious Story. Ethnographic Studies.

[8]. Sacks, H. (1972) 'Notes on Police Assessment of Moral Character'. In Sudnow, D. (ed) Studies in Social Interaction New York: Free Press. P. 280-93.

[9]. Hartland, J. (1993). The Use of Intelligent Machines for Electrocardiograph Interpretation. In G. Button, (Ed.) Technology in Working Order. London: Routledge.

[10]. Hartswood, M., Procter, R., (2000). Computer-aided Mammography: A Case Study of Error Management in a Skilled Decision-Making Task. Topics in Health Information Management, Vol. 20(4), p. 38-54.

[11]. Heathfield, H and Buchan I. (1996) Letters: Current evaluations of information technology in health care are often inadequate' in BMJ 1996: 313: pp1008.

[12]. Heathfield, H. A. and Wyatt, J. (1993). Philosophies for the design and development of clinical decision support systems. Methods of Information in Medicine, 32(1):1-8.