Clinical image quality in mammography

Patsy Whelehan
Senior research radiographer, University of Dundee
Honorary specialist radiographer, NHS Tayside

Introduction
Image quality in mammography encompasses several interlinking domains. Certain determinants of overall image quality are physical and unchanged, such as the detector's pixel size. Features of the image such as exposure, contrast and noise can be considered as being determined by an interaction between the system's technical specifications and the operator, with the operator perhaps exerting the minority influence in the context of automatic exposure control. A final group of parameters can be considered as primarily determined by the operator – the radiographer or assistant practitioner – in concert with the client/patient. These include positioning and presence or absence of blur caused by patient motion.

While physical properties and technical parameters can be evaluated by objective testing, overall quality of the final clinical image, and its suitability for radiological interpretation, can currently only be assessed by more subjective means. This article will draw on the available literature to evaluate the importance of mammographic image quality, and discuss methods for assessing that quality.

The clinical importance of good image quality in mammography

It is intuitively obvious that cancers will be missed on mammography if image quality is sufficiently poor, but empirical research to demonstrate this and quantify the problem is not plentiful. Several studies have addressed the contribution of poor image quality to false negative screening mammography. Of those cited, the association between image quality failures and false negative mammography or delayed diagnosis was the primary focus of the study in only two and the study cohorts overlapped in two. All but one of the studies were carried out before the implementation of digital mammography.

Work by Taplin and Buist involved statistical analyses to investigate factors affecting the odds ratio of a cancer being diagnosed in the interval after a negative screen rather than at screening. In these two studies, image quality was assessed by one radiologist on a five-point scale (collapsed to pass, borderline or fail), addressing the following aspects: positioning, compression, exposure, noise, sharpness, contrast and artefacts. In Buist's cohort of 73 women with cancer who were aged 40-49 at screening, approximately 20% of the excess odds of having an interval cancer diagnosed within 12 months of a negative screen was explained by poor image quality. In Taplin's larger cohort, not restricted to younger age, mammographic sensitivity was about 18% lower when positioning criteria were classified as failed. Positioning failure significantly increased the odds of an interval cancer versus a screen-detected cancer, in univariable and multivariable analysis. Associations were not statistically significant for other image quality parameters.

Rauscher's study investigated the influence of image quality on breast cancer stage at diagnosis in 268 patients diagnosed either at or within two years of screening, and aimed to identify differential effects according to socio-demographic variables. Image quality assessment was similar to that used by Buist and Taplin but Rauscher et al summed the scores for the criteria and then dichotomised them whereby higher image quality was defined by scores of 4 or 5 on a scale of 1-5. They also considered the more technical image quality parameters, such as exposure and contrast, separately from those deemed to be most heavily dependent on the skill of the mammographer. The key finding from this study was that higher scores on the ‘technologist-associated’ factors, such as positioning, were associated with earlier stage at diagnosis.

The remaining studies asked what proportions of false negative screenings were attributable to mammography image quality failures. These studies mostly included low numbers of false negative lesions (n=11-19 in four studies and 115 in one) and only descriptive statistics were appropriate to report the incidence of image quality failure in these collections. The largest of this group of studies included 286 screen-detected cancers which, in informed retrospect, had been visible on a prior screen. One hundred and fifteen of these were subsequently classified as detectable based on a majority decision of five readers in simulated screening conditions. Subjective assignment of reasons why the cancers might originally have been missed was carried out by two expert radiologists and image quality failures were cited as one of the likely factors in 53 of the 116 cases. Positioning and compression were the most common types of deficiency. The other four studies all concerned false negative interval cancers (n=62 in total). One or two or three radiologists subjectively attributed causes to the false negative results in these studies. Combining the findings, 11 of 62 false negatives were attributed to image quality deficits of various kinds, including positioning.

In summary, published evidence regarding the contribution of clinical image quality deficiencies to false negative mammography is of mixed quality and limited quantity. However, the evidence does support the intuitively obvious position that good mammographic technique is essential to maximising cancer detection, and indicates that poor technique may account for a substantial proportion of missed cancers in breast screening. Furthermore, there are suggestions that factors which depend most heavily on the skill of the mammographer are the most influential, perhaps partly because they are the most subject to variance.

Assessing clinical image quality in mammography

There are several important roles for rigorous assessment of mammographic image quality. These include ensuring high standards of mammography to minimise missed cancers, use in education, training and competency assessment, and research use. Regarding research use, the image quality assessments in the studies discussed above were somewhat unsystematic in several, and where more formal scales were employed, no firm evidence was offered for their reliability.

Image quality assessment in clinical mammography prac-
tice, designed to support high standards and thereby effective cancer detection and diagnosis, is subject to European guidelines and UK NHS Breast Screening Programme guidelines. These consist of largely qualitative and subjective criteria, such as "appropriate exposure", "inframammary angle clearly demonstrated". Since at least 1995 it has been recognised that such criteria risk being defined too vaguely, for example employing adjectives which lack any acceptable limits or valid measurement techniques. Unfortunately, these problems are still yet to be resolved. Quality assessment tools have been developed which aim to enable clear categorisation between images which meet or do not meet groups of criteria. Probably the best known is the PGMI system, which stands for perfect, good, moderate, inadequate. A three-point scale using similar criteria has also been described. Both these scales retain some adjectives which are open to subjective judgement and therefore susceptible to observer variability, eg "appropriate", "adequate" and "wide". Observer variability has been shown to be a problem even when categorising images dichotomously as acceptable or not. Figures 1 and 2 illustrate some of the judgement criteria which have to be made in applying the PGMI system. Recent work by Boyce and colleagues aims to address observer variability and thereby improve reliability and validity in the PGMI scale by clarifying and refining the definitions of the criteria, while Mercer and colleagues have conducted preliminary work, drawing on classical test theory from the field of psychology, proposing a new scale.

Emerging technologies in image quality assessment

The advent of digital mammography has several implications for image quality assessment. In the context of breast screening on mobile units where previously it was not usually possible to view the mammograms during the client’s appointment, changing to digital imaging meant that any repeats required could be conducted immediately rather than at a second appointment. However, difficulties can arise from the limited time available to the mammographer for evaluating the images, and from inferior viewing conditions in the examination room compared to the reporting environment. Detecting movement blur in the lighting conditions and with the monitor specifications which predominate in mammography examination rooms can be particularly challenging. Upon implementation of digital screening, it was observed that there was an increased rate of repeats later deemed unnecessary. In response to these problems, the new capabilities associated with digital technology are being harnessed to develop automated methods for detecting positioning faults. This may involve computer detection of anatomical landmarks on the images and the application of measurements to enable established positioning criteria to be evaluated. In addition, it may be possible to employ methods utilising modulation transfer function or fractal dimensions for immediate detection of blurring. However, the development of automated clinical image quality evaluation tools relies on human observers to provide the reference standard; therefore, the importance of improving the validity and reliability of observer-based classification schemes is undiminished.

References

Figure 1
Would you repeat this image? Is the pectoral muscle too narrow?

Figure 2
Would you classify this image as good or moderate in the PGMI system? If the inframammary angle is not clearly shown, and/or the nipple is not in profile, it is moderate.