ORIGINAL ARTICLE

Validating Image-box—a new approach to multicentre radiology reviews using a web-based image review system: initial methodology

Penelope Moyle [1], Masako Katako [2], Kevin Wheeler [3], Oliver Morrish [4], Susan Gerty [3], Matthew Wallis [1], Diana Eccles [3], Ruth Warren [2]

1. Cambridge Breast Unit, Cambridge University Hospitals, Cambridge, United Kingdom. 2. Department of Radiology, Cambridge University Hospitals, Cambridge, United Kingdom. 3. Southampton General Hospital, University of Southampton, Hampshire, United Kingdom. 4. East Anglian Regional Radiation Protection Service, Cambridge University Hospitals, Cambridge, United Kingdom.

Correspondence: Penelope Moyle. Address: Cambridge Breast Unit, Cambridge University Hospitals NHS Foundation Trust Box 97, Hills Road, Cambridge, CB2 0QQ, United Kingdom. Telephone: 44-148-0363-671. Fax: 44-122-3217-847. Email: pennymoyle@doctors.org.uk

Received: August 25, 2011 **Accepted:** October 10, 2011 **Published:** December 1, 2011

DOI: 10.5430/jbgc.v1n1p24 **URL:** http://dx.doi.org/10.5430/jbgc.v1n1p24

Abstract

Objective: To validate the use of Image-box— a web-based image review system for large scale multicenter trials.

Methods: As part of the multicentre trial "Prospective study of Outcomes in Sporadic versus Hereditary breast cancer (POSH) study" mammograms were collected and digitized. Software was created allowing digitized mammograms to be viewed from any location with internet access. Two validation studies were performed. Firstly phantom studies using line-pairs (1-20 lpmm-1) to assess spatial resolution and the CDMAM 3.4 phantom were used to assess the visibility of gold discs for contrast detail. Secondly a comparison was made between the scoring of 77 patients' original diagnostic analogue mammograms from 29 hospitals and the corresponding web images. Films and web images were scored by one experienced breast radiologist according to the BIRADS classification. At least 8 weeks elapsed between scoring images from the same patient.

Results: The original analogue spatial resolution was 20 lpmm-1 and web resolution of the same image reduced at 8.9 lpmm-1. Contrast detail assessment demonstrated analogue and web images with near concordance and the image quality factor, IQFinv for the web was reduced but within the 95% confidence interval of the original mammograms. BIRADS final assessment showed good agreement between the analogue and the identical web images (Kappa 0.82).

Conclusion: The overall reduction in spatial resolution did not adversely affect the quality of the diagnostic image on Image-box. This may be due to its functionality, specifically the high magnification zoom which enhances diagnostic interpretation. Image-box is a valuable tool to review mammograms as part of a multicentre trial. It addresses the logistical challenges for large breast imaging reviews and has further potential for education and self-assessment.

Key words

Multicenter, Breast cancer, Image-box, Web-based imaging

1 Introduction

There is growing interest in using radiological imaging as a part of large-scale multicentre trials especially in breast cancer research. There are logistical challenges of assembling images from many sources, involving expert radiologists to review the images within an acceptable timescale that will not compromise patient care in the donor centers. The "Prospective study of Outcomes in Sporadic versus Hereditary breast cancer (POSH) study" is a multicenter cohort study incorporating a systematic review of radiological and pathological characteristics of young onset breast cancer cases. Image-box (a web-based image review system) was designed specifically for the POSH study to address the logistical challenges for a large volume, multicentre retrospective review. This initial methodology paper validates the process of image-box.

The POSH study is a prospective cohort study recruiting 3,000 women aged 40 years or younger at breast cancer diagnosis. The overall aim of the POSH study is to establish whether genetic status influences the prognosis of primary breast cancer independently of known prognostic factors [1]. The aim of the radiology review within this study is to elucidate whether there are radiological features of breast cancer in women less than 40 years which may be predictive of genetic, pathological factors or outcomes, and which may impact on patient care or screening. The radiology review will analyze 2000 breast cancer patients less than 40 years from 123 hospitals within the United Kingdom (UK).

Our aim was to produce a web based image review system of mammograms which could be viewed from any location with internet access to try and avoid the social and financial costs of moving radiologist to one area to review these images. It had to be easy to navigate, secure, validated for diagnostic image quality, and post process the acquired information ready for statistical interpretation. The design and diagnostic validation of Image-box will therefore be described and the advantages and disadvantages of a web-based system will be illustrated using examples from the POSH study (figure 1).

2 Patients and methods

2.1 Patients

All patients included in this study were participating in the Prospective study of Outcomes in Sporadic versus Hereditary breast cancer (POSH) study for which the study protocol is published1. Ethical approval was obtained (approval number MREC /00/6/69) and all participants in the POSH study gave informed consent including the use of their medical, pathology and imaging records. All patients included were younger than 41 years at diagnosis of invasive breast cancer.

2.2 Software

Image-box is available on line http://152.78.9.156/imagebox with a demonstration mode available (username: demo password: password) to further illustrate this text. The process of obtaining images ready to be securely viewed on the web takes a number of stages which will be described.

2.3 Preparation of images

The original diagnostic mammograms of recruited patients are requested from participating hospitals within the POSH study. Original diagnostic analogue mammograms are accepted but no copied films are included as this degrades the quality of the image. The images are sent to our institution by secure courier where they are digitalized, anonymised and given a unique study number.

2.4 Digitalization

Digitalization of the analogue mammograms is by an Array laser 2905 film digitizer (Array Corporation USA, Hampton, NH, USA) which gives a pixel size of 3600 by 4800 and dicom resolution of 1mm=20pixels, equivalent to 12bit. The dicom

images are converted to bitmap images as this format is the most suitable for uploading to the web. This conversion reduces the resolution to 8 bit. Digital mammograms are anonymised, assigned their unique POSH study number then converted to bitmap (8 bit) format and sent by compact disc (CD) to the POSH study centre for uploading to the web.

The formatted and anonymised images are uploaded into Image-box where there is further compression of the images to 550 by 900 pixels.

2.5 Image-box

Image-box is a secure website which is accessed only by password to registered users associated with the radiology review. The bank of patient images allocated a unique study number can be viewed and assessed on line at any time and on any computer with internet access. Image viewing can be assisted by a variety of function tools including: Scroll bar, measurement calipers, zoom and expand function to assess the image in a larger size and in more detail and a navigation function to move around the image and to scroll to the next set of mammography images (all patients have 4 standard mammographic views). Details are best appreciated on the website.

2.6 Recording the results of the image review

The assessment of the mammograms and ultrasounds uses the BI-RADS classification [2, 3]. BI-RADS is a standardized lexicon which includes terms for describing the morphology and distribution of breast lesions. It also describes the radiologist's level of suspicion about a mammographic abnormality.

An integrated scoring system is part of Image-box with the radiologist ticking a web scoring card which is unique to each patient. This helps reduce clerical error with separate papers or database systems for scoring (figure 2). When assessing the mammograms there is a lesion report if an abnormality is found and an overall general report. The general report is always entered even if the mammograms are considered normal as the assessment of density is an important risk factor in breast cancer [4].

The results are automatically uploaded onto an excel spreadsheet in a binary format once the 'save case' button has been pressed. It is then available for future statistical analysis. Once a case is finished a new image set is automatically allocated to the reader.

2.7 Validation of Image-box

Image quality is paramount is any imaging review, especially mammography. Subtle lesions and micro calcification need high resolution images for diagnostic accuracy and correct interpretation. To assess the image quality of the web images compared to the original analogue mammography, two validation studies were performed:

1) Phantom study

2) Comparison of original film verses web images film by a radiologist assessing and scoring the films.

2.8 Phantom validation study

Two test objects used for quality assurance in mammography were imaged and scored. A test pattern consisting of a range of line-pairs (1-20 lpmm-1) was used to assess spatial resolution (Figure 3) and the CDMAM 3.4 contrast detail phantom (Artinis Medical Systems B.V., Zetten, The Netherlands) was used to assess the visibility of gold discs of various diameter and thickness (Figure 4). Three images of the line-pair test object aligned at 45° to the midline were taken in contact with the Bucky at 28kVp, Mo/Mo, 20mAs. Eight images of the contrast-detail phantom were taken with it placed in the center of 4cm of PMMA and exposed at 27kVp, Mo/Rh with the automatic exposure control terminating exposures at a mean of 91.7 mAs (representing a mean glandular dose of 1.7mGy to a 60 mm breast).

Exposures were taken on a Siemens MAMMOMAT 3000 Nova (Siemens AG, Earlangen, Germany) and images were captured using Kodak Min-R EV film with a Min-R intensifying screen. The resultant films were then digitized, compressed and uploaded to the web-based viewing system as previously described. The web phantom images were viewed on a Dell desktop computer with the screen size of 34cm by 27cm and maximum screen resolution in ambient lighting. These are generic desktop computers used at our institution.

The spatial resolution images were viewed and scored by looking for the last resolvable line-pair group. The contrast detail images were scored using the methodology detailed in the manual for the phantom [5] resulting in a threshold gold thickness for each detail diameter and an overall image quality factor given by

$$IQFinv = 100 \Big/ \sum_{i=1}^{16} CiDi$$

where Ci is the threshold thickness of the disc (16 in total) with diameter Di. All images were scored independently by two observers in an area with low ambient light levels.

2.9 Validation of assessing original diagnostic analogue films verses web images

The second validation study involved 77 patients with original diagnostic analogue mammograms (2 standard views for each breast; cranio-caudal and medio-lateral views) from 29 hospitals. An experienced consultant radiologist read and scored the diagnostic films for each breast (154 breasts) according to the BIRADS classification. A 5-category final assessment was given to each breast (table 1).

Of the 154 breasts reviewed 48 demonstrated a mass lesion and 23 demonstrated calcification. The specific BIRADS descriptors for mass margin and calcification were recorded (table 2).

The original films were then digitized, anonymised and up loaded onto Image-box as previously described. The web images were read and scored again by the same radiologist a minimum of 8 weeks later which reduced reading bias from remembering the original cases. The radiologist used a Dell desktop computer with the screen size of 34cm by 27cm with maximum screen resolution in reduced ambient lighting which again is the generic computer used in our institution.

3 Results

3.1 Phantom validation study

The spatial resolution measured on the film images was 20 lpmm-1, the maximum score on the test object. The resolution on the web versions of the same images were 8.9 lpmm-1. The same result was scored on each image by both observers.

Figure 5 demonstrates the scores of the contrast detail images. The mean threshold thickness determined by each observer for each disc diameter on each image is shown. The difference in image quality factor, IQFinv for the two viewing methods is shown in figure 6. Error bars in both figures indicate the 95% confidence interval.

3.2 Validation of assessing original diagnostic analogue films verses web images

The final assessment showed good agreement between the analogue and the identical web images (weighted Kappa 0.82, Table 3). Of the 48 masses, there was intraobserver agreement on the margin description in 71% and of the calcification description 52%.

Table 1. Final assessment using the BIRADS lexicon

BIRADS Score	Overall mammographic features of the assessed breast
1	Normal
2	Benign
3	Probable benign
4	Suspicious abnormality
5	Highly suggestive of malignancy

Table 2. BIRADS descriptors for mass margin and calcification

BIRADS margin description	BIRADS calcification description
Circumscribed	benign
Micro lobulated	Punctuate/ powder
Obscured	Granular/ crushed stone
Indistinct	Casting linear/branching
speculated	

Table 3. Mammogram agreement between web and film images (77 pts, 154 breasts)

	No. of breasts
5-category final assessment of the breast	122/154 (79%)-Kappa 0.82
Mass margin (5 different categories)	34/48 (71%)
Calcification type (4 different categories)	12/23 (52%)

4 Discussions

There is a significant reduction in spatial resolution when images are digitized and this is an expected consequence of the finite pixel size of digital images. This is recognized in the European guidelines [6] that require film-screen imaging systems to achieve spatial resolutions of >12 lpmm-1 (when measured on top of 45mm of PMMA). Although the resolution on the web versions was 8.9 lpmm-1 there are no such requirements for digital mammography (although a resolution approaching the Nyquist limit should be expected) and there are no requirements on processed images.

The threshold thickness phantom for the detection of detailed diameters is similar for analogue and image-box mammograms and variations between the two are within the error of the measurement. The overall score of each method as indicated by the IQFinv value show a slight decrease in detail detectability when the web-based system is used however this again is within the error of the measurement. Therefore from the two phantom studies described, analogue mammograms can be processed for web reading without losing diagnostic capability.

Observer variability in the interpretation of mammograms is well described [7]. Berg et al. demonstrated substantial intraobserver disagreement with five radiologists using BIRADS classification [8]. In describing the borders of a mass and focal asymmetry, intraobserver agreement ranged from 29%-57% (mean 40% concordance). They demonstrated microcalcification intraobserver agreement from 14%-71% (mean 49%); however the overall final assessment of the breast had the highest concordance at 65%. Our study demonstrates higher intraobserver. Concordance in the overall assessment (weighted Kappa 0.82) and with the margin description (71%). There was a lower agreement with micro calcification (52%) but this still is comparable to the literature.

Intraobserver variability is dependent on the subtly of the lesion and microcalcifications can be difficult to identify and interpret. Even allowing for the reduced line pairing quality these results demonstrate that subtle microcalcifications can is still be identified. Pijnappel et al. [9] demonstrated microcalcification intraobserver variability has the lowest agreement with a kappa of 0.54 (standard error 0.07) in analogue film to which our data is comparable in a processed web image. Penedo [10] demonstrated lossy compression of digital mammographic data at 80:1 with JPEG2000 or the object-based SPIHT algorithm can be performed without decreasing the rate of detection of clusters of micro calcifications and masses. We have not compressed to this extent using Bitmap images although our cohort was analogue not digital mammograms.

Therefore, we conclude that the overall reduction in line pairs has not adversely affected the quality of the diagnostic image on Image-box. This may be due to the functionality of Image-box, specifically the high magnification zoom which will enhance diagnostic interpretation.

One limitation of this study is having one radiologist interpreting the imaging however this is also strength of this approach during this initial evaluation and validation of the use of web images for reporting radiological characteristics since intraobserver variation, when applied to a standard set of images, may vary to a great extent depending on reader experience and skill. Reproducibility is more likely when one has fixed a defined threshold for suspicion which results from extensive clinical experience. The radiologist reporting the images has more than 20 years breast radiology interpreting both analogue and digital mammograms so variation in experience is internally controlled for. A time interval of more than 8 weeks between reading the two mammogram sets was to reduced memory bias that might lead to over-estimation of concordance between film and web images. All the images in the study were analogue which was as expected as this is a retrospective study and most UK centers are in a transition to digital imaging.

New techniques are emerging for clinical management of patients with mammograms performed at remote sites. The digitized and anonymised mammograms being transmitted across low-level communication systems to the central site for radiology review [11]. The main advantage of Image-box is the accessibility of the radiological images linked to a comprehensive research database. We are using Image-box as a research tool, but its further application can be in education and self-assessment. Self-assessment will become increasingly important to the radiologist in the context of revalidation. As part of the quality assurance programme for the National Health Service Breast Screening Programme (NHSBSP), each film-reader can participate in a voluntary self-assessment scheme (Personal Performance in Mammographic Screening, PERFORMS) which consists of a number of recent challenging breast screening cases that are amassed nationally and distributed bi-annually [12]. This self-assessment tool travels around the U.K as a computer with cases and database stored on its hard drive. Only when the PERFORMS computer is at the Radiologists institution can the radiologist access the cases. Image box has the advantage of all images stored at one central server ready to be accessed at any point with internet access and is not geographically tied. Accessible high quality self-assessment will certainly be a necessity for the future.

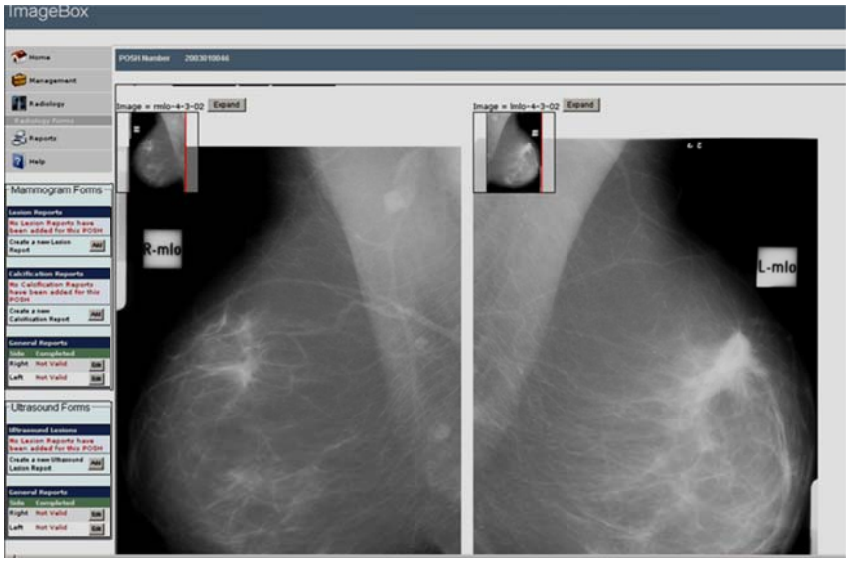

Figure 1. An example page of Image-box

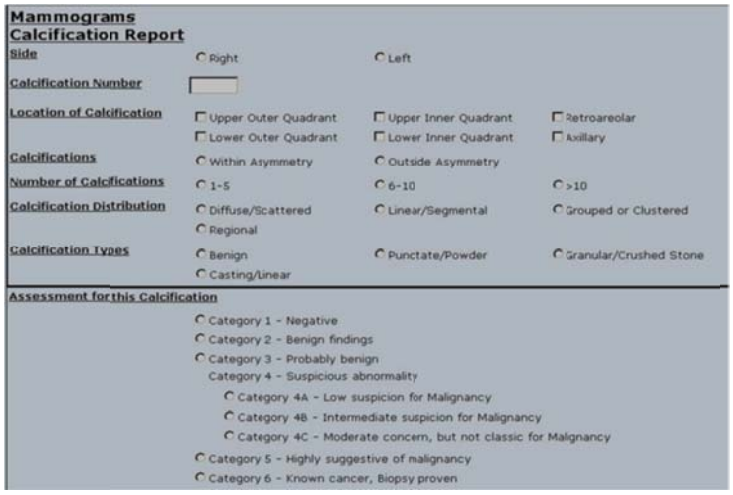

Figure 2. An example of the integrated scoring system in Image-box

Figure 3. A line- pairs test pattern to assess mammographic spatial resolution

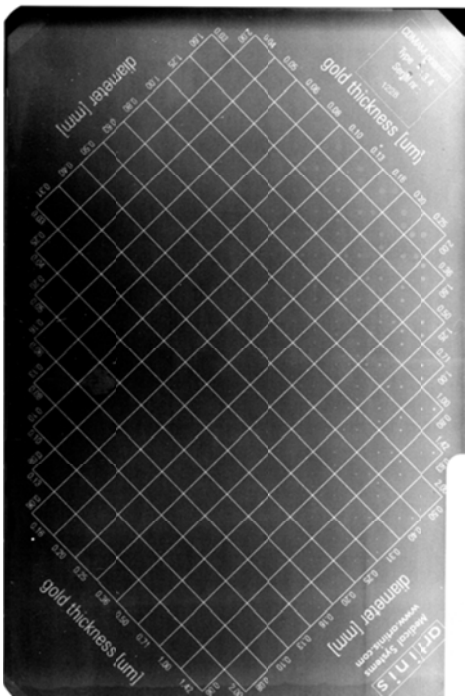

Figure 4. CDMAM 3.4 contrast detail phantom to assess the visibility of gold discs of various diameter and thickness

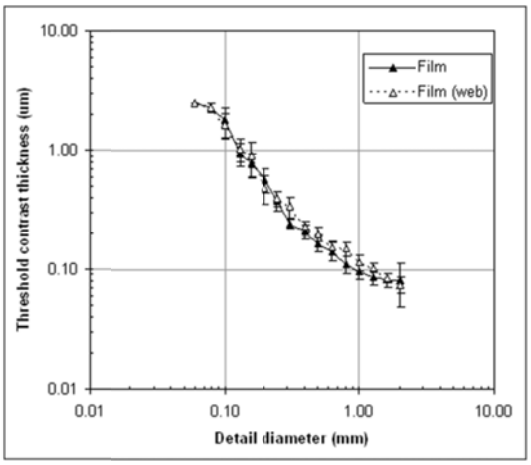

Figure 5. Variation in threshold thickness of a gold disc with detail diameter for original film and web-based images

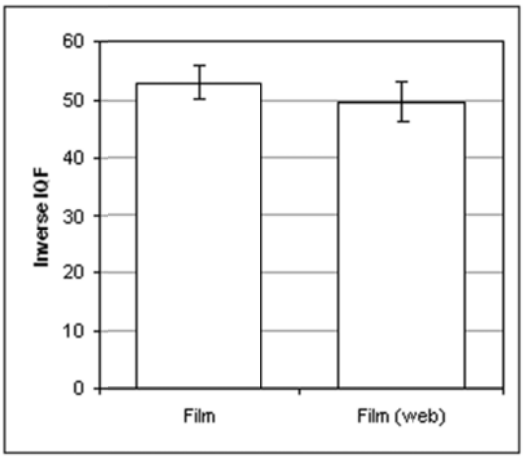

Figure 6. IQFinv for original film and web-based images

5 Conclusions

This study describes the concept, process and validation of Image-box to see if is a feasible tool for multicentre trials in breast imaging. We conclude analogue mammograms can be processed to the web without loss of diagnostic information and intraobserver variability of the BIRADS scoring system between analogue and web images show good concordance, in keeping with the literature. Radiologists are therefore able to access high quality mammograms quickly and easily with an integrated scoring system linked with the patient's images. Image-box will proceed as web radiology review as part of the POSH study, with hopes it will contribute to other large multicentre trials. It has potential future roles in education and self-assessment.

Conflict of interest statement

We declare that we have no conflict of interest.

References

[1] Eccles D, Gerty S, Simmonds P, Hammond V, Ennis S, Altman DG; POSH steering group. Prospective study of Outcomes in Sporadic versus Hereditary breast cancer (POSH): study protocol. BMC Cancer. 2007;15(7):160.

[2] Reston,VA, American College of Radiologists. BI-RADS: mammography. Breast imaging reporting and data system: BI-RADS atlas, 4th ed. American College of Radiology. 2003.

[3] Reston,VA, American College of Radiologists. BI-RADS: ultrasound. Breast imaging reporting and data system: BI-RADS atlas, 4th ed. American College of Radiology. 2003.

[4] Ding J, Warren R, Girling A, Thompson D, Easton D; Mammographic Density, Estrogen Receptor Status and Other Breast Cancer Tumor Characteristics. Breast J. 2010 Apr 12 [Epub ahead of print].

[5] Thijssen M & Bijkerk K, Manual Contrast-Detail Phantom Artinis CDMAM type3.4 Artinis Medical Systems B.V. 2006.

[6] European Guidelines for Quality Assurance in Mammography Screening, European Commission Forth Edition. 2006.

[7] Ciatto S, Houssami N, Apruzzese A, Bassetti E, Brancato B, Carozzi F et al; Reader variability in reporting breast imaging according to BI-RADS assessment categories (the Florence experience).Breast. 2006;15:44-51. PMid: 16076556. http://dx.doi.org/10.1016/j.breast. 2005;04:019

[8] Berg WA, Campassi C, Langenberg P, Sexton MJ; Breast Imaging Reporting and Data System: inter- and intraobserver variability in feature analysis and final assessment. AJR. 2000;174:1769-77. PMid: 10845521.

[9] Pijnappel RM, Peeters PH, Hendriks JH, Mali WP; Reproducibility of mammographic classifications for non-palpable suspect lesions with microcalcifications. Br J Radiol. 2004;77:312-14. PMid: 15107321 http://dx.doi.org/10.1259/bjr/84593467

[10] Penedo M, Souto M, Tahoces PG, Carreira JM, Villalón J, Porto G, et al. Free-response receiver operating characteristic evaluation of lossy JPEG2000 and object-based set partitioning in hierarchical trees compression of digitized mammograms. Radiology. 2005;237:450-57. PMid: 16244253. http://dx.doi.org/10.1148/radiol.2372040996

[11] Leader JK, Hakim CM, Ganott MA, Chough DM, Wallace LP, Clearfield RJ, et al. A multisite telemammography system for remote management of screening mammography: an assessment of technical, operational, and clinical issues. J Digit Imaging. 2006;19:216-25. PMid: 16710798. http://dx.doi.org/10.1007/s10278-006-0585-9

[12] Scott HJ, Gale AG; Breast screening: PERFORMS identifies key mammographic training needs. British Journal of Radiology. 2006;79,S127-33. PMid: 17209118. http://dx.doi.org/10.1259/bjr/25049149