Quantification of the UK 5-point breast imaging classification and mapping to BI-RADS to facilitate comparison with international literature

¹K TAYLOR, DCR, MSc, ¹P BRITTON, MBBS, FRCR, ¹S O'KEEFFE, MBChB, FRCR and ²M G WALLIS, MBChB, FRCR

¹Department of Radiology, Cambridge Breast Unit and ²National Institute for Health Research Cambridge Biomedical Research Centre, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK

Objective: The UK 5-point breast imaging scoring system, recently formalised by the Royal College of Radiologists Breast Group, does not specify the likelihood of malignancy in each category. The breast imaging and reporting data system (BI-RADS) is widely used throughout North America and much of Europe. The main purpose of this study is to quantify the cancer likelihood of each of the UK 5-point categories and map them to comparable BI-RADS categories to facilitate comparison with North American and European literature and publication of UK research abroad.

Methods: During the 8 year study period, mammogram and ultrasound results were UK scored and the percentage of cancer outcomes within each group calculated. These were then compared with the percentage incidence of the BI-RADS categories.

Results: Of 23 741 separate assessment episodes, 15 288 mammograms and 10 642 ultrasound examinations were evaluated. There was a direct correlation between UK scoring and BI-RADS for categories 1 and 5. UK Score 2 lipomas and simple cysts correlated with BI-RADS 2, with the remaining UK Score 2 lesions (mostly fibroadenomas) assigned to BI-RADS 3. BI-RADS 4 incorporates a wide range of cancer risk (2–95%) with subdivisions a, b and c indicating increasing, but unspecified, likelihood of malignancy. UK Score 3 correlated with BI-RADS 4 a/b and UK Score 4 corresponded with BI-RADS 4c.

Conclusion: This study quantifies the cancer likelihood of the UK scoring and maps them to parallel BI-RADS categories, with equivalent cancer risks. This facilitates the ability to share UK research data and clinical practice on an international scale.

Received 4 August 2010
Revised 7 January 2011
Accepted 20 January 2011

DOI: 10.1259/bjr/48490964

© 2011 The British Institute of
Radiology

Until recently, UK breast units used variations of a standard 1–5 scoring system for normal, benign, uncertain, suspicious and malignant appearances, respectively. This led to disparity in reporting and made sharing information between centres difficult. In an effort to address this, the standard UK scoring system was recently formalised in a publication on behalf of the Royal College of Radiologists (RCR) Breast Group (RCRBG) [1].

The breast imaging reporting and data system (BI-RADS) has been successfully used in North America and most of Europe since its publication by the American College of Radiologists (ACR). It was first used for mammography reporting and subsequently ultrasound and MRI [2]. It encompasses a lexicon of descriptors based on ACR recommendations and the ultrasound publications of Stavros et al [3] ranging from BI-RADS 1 to 5, where BI-RADS 1 is normal, BI-RADS 2 is benign, BI-RADS 3 has a likelihood of malignancy of <2%, BI-RADS 4 is between 2% and 95% and BI-RADS 5 >95% likelihood

of a malignancy (Table 1). The literature has suggested subdivision of the large BI-RADS 4 category into 4a, 4b and 4c as a sliding scale of increasing suspicion from 3 (probably benign) to 5 (malignant) [4]. The BI-RADS system has enabled standardisation and uniformity of reporting, which in turn facilitates accurate comparison of performance, research data and audit. Unlike the BI-RADS classification, the UK system does not specify the percentage likelihood of each category.

The primary purposes of this study are to quantify the percentage risk of malignancy for each of the five RCRBG categories and determine whether a UK RCRBG scored mammogram or ultrasound can be assigned a corresponding BI-RADS category in terms of the percentage likelihood of a malignant outcome. Successful mapping of the UK RCRBG scoring system to BI-RADS could facilitate more accurate comparison of data from the UK with North America and Europe, particularly in terms of research and publication.

Finally, we have undertaken a correlation between the most common sonographic features of each UK scored group and the descriptors of corresponding BI-RADS categories to determine whether this could assist in more accurate BI-RADS category assignment.

Address correspondence to: Mrs Kathryn Taylor, Department of Radiology, Cambridge Breast Unit, Box 97, Cambridge University Hospitals NHS Foundation Trust, Cambridge CB2 0QQ, UK. E-mail: kathryn.taylor@addenbrookes.nhs.uk

Table 1. Breast imaging and reporting data system (BI-RADS) categories for mammography and ultrasound with associated feature descriptors and percentage likelihood of malignancy

BI-RADS category	Mammogram description	Ultrasound description	Percentage likelihood of malignancy (%)
1	Normal	Normal	0
2	Benign: no suspicious findings. Also involuting, calcified fibroadenomas, multiple secretory calcifications, fat-containing lesions, such as oil cysts, lipomas, galactocoeles and mixed-density hamartomas	Benign: no suspicious findings Also simple cyst, uniform hyperechoic lesions.	0
3	Probably benign: neither category 2, nor 4 or 5. Non-palpable, circumscribed masses punctate calcifications. Focal asymmetry resolved on spot compression views	Probably benign: neither category 2, nor 4 or 5. Typical fibroadenomas	<2
4 ^a	Suspicious: Not category 3 or 5. Clustered amorphous calcifications. Atypical fibroadenomas	Suspicious: one or more suspicious features, not category 5. Biopsy minor – round shape. Microlobulated, indistinct or angular margin. Not parallel orientation. Duct extension. Complex echogenicity. Posterior shadowing	2–95
5	Highly suspicious: spiculated, irregular high-density masses, segmental or linear fine calcifications, an irregular spiculated mass with associated pleomorphic calcifications	Highly suspicious: two or more major suspicious findings. Major – irregular shape. Spiculated margin Microcalcifications	>95

^aCategory 4 has been subdivided into a, b, and c to indicate an increasing likelihood of malignancy from category 3 to 5. No formal percentage cancer rates have been assigned to these subdivisions.

BI-RADS 4 is described as having one or more minor features and BI-RADS 5 is described as having 2 or more major suspicious features.

Method and materials

All patients referred by their general practitioner (GP) to a symptomatic breast clinic over an 8 year period (1 January 2001 to 31 December 2008) were retrospectively identified from an in-house dedicated breast database. During their visit, patients underwent triple assessment in line with local protocols adapted from national guidelines [5, 6]. Patients over 35 years of age underwent bilateral mammography on either a Siemens Mammomat 3000 Nova analogue machine (Siemens Healthcare, Erlangen, Germany) a GE Senographe DS digital machine or a GE Senographe Essential machine (General Electric Medical Systems UK, Little Chalfont, UK). Those with a focal, palpable abnormality also had targeted ultrasound using a 12–16 MHz matrix linear array transducer on a Toshiba Aplio ultrasound platform (Toshiba Medical Systems, Tochigi, Japan). These were performed by either 1 of 5 consultant breast radiologists or a year 5/6 specialist breast radiology registrar. Any patients having multiple referrals during the study period were included as separate assessment episodes. Mammogram and ultrasound results were classified according to the standard UK 5-point classification scale, which was in place during the study period prior to the agreed classification system formalised by the RCRBG in 2009. All information was entered prospectively into the joint clinical information system (JCIS) at the time of assessment. JCIS is an *n*-tier web-based clinical information system supported by a structured query language (SQL) server database. The system was built in-house in partnership with the breast clinical team using i5™ Web application (Dataline Software, Brighton, UK) and Microsoft

technologies including Visual Basic (Microsoft Corp, Redmond, WA). For the purposes of analysis, patients were divided into two groups: those with breast cancer (based upon a malignant core biopsy or malignant surgical histology result) and those without. Results were scored and grouped in accordance with the standard UK 5-point scale and the percentage of cancer outcomes within each group calculated. These were then compared with the percentage incidence of the BI-RADS categories.

Finally, the sonographic features most commonly present within each ultrasound scoring group 1–5 were documented in terms of mass boundary, mass echogenicity and posterior acoustic appearances. Only examinations with all features documented were included in the study and the percentage cancer incidence within each group was calculated.

Results

During the 8 year study period 19 130 patients attended the symptomatic breast clinic following GP referral. A proportion of these were referred more than once and a total of 23 741 separate assessment episodes were included for analysis. Of these, 1635 (6.9%) resulted in a cancer diagnosis based upon a malignant core biopsy or surgical histology result and the remaining 22 106 (93.1%) were found to be cancer free following triple assessment.

8391 patients did not undergo mammography and a further 62 women did not have their mammogram results recorded. These were excluded leaving a total of 15 288 mammograms for evaluation. The percentage of patients with a Category 1 report subsequently diagnosed with

Table 2. UK Royal College of Radiologists Breast Group (RCRBG) scorings for mammography during the study period with associated number of cancers and percentage likelihood of cancer for each group. Each group has then been linked to the breast imaging and reporting data system (BI-RADS) category with the same score for comparison of cancer likelihood

Mammography					
UK (RCR) category	UK RCRBG scorings			Equivalent BI-RADS	
	Number of mammograms	Number of cancers	Likelihood of malignancy (%)	Likelihood of malignancy (%)	BI-RADS category
1. Normal	10744	190	1.8	0	1–2
2. Benign	2843	37	1.3	<2	3
3. Indeterminate /probably benign findings	657	268	40.8	2–95	4a–4b
4. Findings suspicious of malignancy	423	400	94.6	2–95	4c
5. Findings highly suspicious of malignancy	621	607	97.8	>95	5
Total	15288	1502			

breast cancer was 1.8%, with a Category 2, 1.3%; Category 3, 40.8%; Category 4, 94.6%; and Category 5, 97.8%. The number of mammograms allocated to each of the five standard UK categories and the percentage of cancers diagnosed is shown in Table 2 and compared with the equivalent BI-RADS category.

8131 patients did not undergo ultrasound examination and a further 58 women did not have ultrasound results recorded. These were excluded leaving a total of 10 462 ultrasound examinations for evaluation. The percentage of patients with a Category 1 ultrasound report subsequently diagnosed with breast cancer was 0.4%, Category 2, 1.8%; Category 3, 17.7%; Category 4, 88.2%; and Category 5, 97.1%. The number of ultrasounds allocated to each of the five standard UK categories and the percentage of cancers diagnosed are shown in Table 3 and compared with the equivalent BI-RADS category. Of the patients with a Category 1 or 2 mammogram report and a Category 1 or 2 ultrasound report subsequently diagnosed with cancer, 28% presented with single duct, bloody nipple discharge and either M1/U1 imaging or M1/U2 imaging, (ultrasound report of retroareolar duct ectasia). Cancer was subsequently diagnosed at duct surgery. 18% presented with a clinically benign palpable lump and either M1/U2 or M2/U2 imaging with features of a lesion thought to probably represent a fibroadenoma; cancer was diagnosed at core biopsy. 27% presented with a clinically suspicious lump and M1/U1 or M2/U2 imaging (ultrasound report of probable dense breast tissue); cancer was diagnosed at core biopsy. 9% had core biopsy because of patient anxiety or family history of breast cancer. 4%

presented with abnormal axillary lymph nodes and in the remaining 14% the method of diagnosis was unclear from the JCIS entry.

For mammograms, there was direct correlation between UK scoring and BI-RADS for categories 1 and 5. UK Score 2 lipomas and simple cysts correlated with BI-RADS 2, with the remaining UK Score 2 lesions (mostly fibroadenomas) being assigned to BI-RADS 3. BI-RADS 4 (subdivisions a, b and c) covers a wide range of cancer risk (2–95%) with increasing likelihood of malignancy. UK Score 3 correlated with BI-RADS 4 a/b and UK Score 4 corresponded with BI-RADS 4c. Ultrasound data are similar to mammography with the exception of a lower percentage cancer incidence in the standard UK indeterminate (Score 3) group.

The ultrasound descriptors used in this institution in order of frequency and by final UK 5-point scoring are shown in Table 4. The percentage likelihood of malignancy associated with these ultrasound features, according to their mass boundary, echogenicity and posterior acoustic appearances are also included. The ACR lexicon suggesting features for inclusion in each BI-RADS category for ultrasound is shown in Table 1.

Our results show lesion boundary and posterior acoustic features to be the most important predictors of malignancy and their importance is also evident in the BI-RADS ultrasound descriptors. A combination of both a well-defined boundary and posterior enhancement gives the lowest likelihood of malignancy (2.1%, equivalent to BI-RADS 3) and an ill-defined boundary and posterior attenuation gives the highest likelihood of malignancy

Table 3. UK Royal College of Radiologists Breast Group (RCRBG) scorings for ultrasound during the study period with associated number of cancers and percentage likelihood of cancer for each group. Each group has then been linked to the breast imaging and reporting data system (BI-RADS) category with the same score for comparison of cancer likelihood

Ultrasound					
UK RCRBG category	UK RCRBG scorings			Equivalent BI-RADS	
	Number of ultrasounds	Number Of cancers	Likelihood of malignancy (%)	Likelihood of malignancy (%)	BI-RADS category
1. Normal	6337	25	0.4	0	1–2
2. Benign	1853	33	1.8	<2	3
3. Indeterminate /probably benign findings	870	154	17.7	2–95	4a/4b
4. Findings suspicious of malignancy	492	434	88.2	2–95	4c
5. Findings highly suspicious of malignancy	910	884	97.1	>95	5
Total	10462	1530			

Table 4. Cambridge breast unit ultrasound descriptors in order of frequency, by final UK Royal College of Radiologists 5-point scoring and percentage likelihood of cancer (PLC) associated with each set of ultrasound descriptors. For example, the most frequent features of a cancer were an ill-defined, heterogeneous, hypoechoic lesion with posterior acoustic attenuation. This occurred in 58.6% cancers with an associated 87.4% likelihood of malignancy. Similarly, the second most common features of a cancer were an ill-defined, heterogeneous, hypoechoic lesion with isoechoic posterior echogenicity. This occurred in 19.6% cancers with an associated 70.4% likelihood of malignancy

UK (RCR) 1–5 score	1st most common			2nd most common			3rd most common			4th most common			Other	Total
	Description	%	PLC	Description	%	PLC	Description	%	PLC	Description	%	PLC	%	
1	Normal			Normal			Normal			Normal				
2	Well-defined homogeneous, hypoechoic, enhancement	36.6	2.1	Well-defined homogeneous, hypoechoic, isoechoic	27.5	7.5	Well-defined heterogeneous, hypoechoic, isoechoic	8.9	23.9	Well-defined homogeneous, hyperechoic, isoechoic	4.9	0	21.9	100.0
3	Ill-defined heterogeneous, hypoechoic, isoechoic	16.7	70.4	Well-defined homogeneous, hypoechoic, isoechoic	13.3	7.5	Ill-defined heterogeneous, hypoechoic, attenuation	12.2	87.4	Well-defined heterogeneous, hypoechoic, isoechoic	12.0	23.9	45.6	100.0
4	Ill-defined heterogeneous, hypoechoic, attenuation	39.8	87.4	Ill-defined, heterogeneous, hypoechoic, isoechoic	20.0	70.4	Ill-defined, homogeneous, hypoechoic, isoechoic	7.2	45.3	Well-defined heterogeneous, hypoechoic, isoechoic	6.7	23.9	26.0	100.0
5	Ill-defined heterogeneous hypoechoic, attenuation	58.6	87.4	Ill-defined heterogeneous, hypoechoic, isoechoic	19.6	70.4	Ill-defined, homogeneous, hypoechoic, attenuation	5.9	78.9	Ill-defined heterogeneous, hypoechoic, enhancement	3.0	72.0	12.7	100.0

(87.4%, equivalent to BI-RADS 4c/5), while a hypoechoic lesion can be indicative of either benignity or malignancy. Posterior acoustic enhancement is usually a predictor of benignity except when associated with an ill-defined heterogeneous lesion. The malignancy risk is then still relatively high at 60% (equivalent to BI-RADS 4b/4c).

Discussion

Although the 5-point scoring system has been used in UK breast centres for many years, the classification was only formalised in 2009 by the RCRBG in an attempt to standardise reporting using a similar rationale to the ACR development of BI-RADS. Although we would strongly support the formalisation of the UK reporting system the classification does not specify the probable cancer risk of each category, which limits potential data comparison. Although we acknowledge that our data are from a single unit, this paper does provide initial information on cancer risk for each of the UK's five reporting categories.

The BI-RADS categorisation system was introduced to standardise reporting for mammography, and later ultrasound. Its use is now established practice in North America and much of Europe and there is plenty of literature endorsement that enables direct comparison of performance, research and audit [7–9]. A literature search performed in the three radiological journals with the highest impact factor ratings (*Radiology*, *American Journal of Roentgenology* and *European Radiology*) from January 2007 to January 2010 illustrates the extensive usage of the BI-RADS system in the published literature. Performing a search of the term “BI-RADS” in the abstract or title resulted in 54 articles, but only 7 relevant articles were obtained searching for the terms “breast” and “scoring system” or “point-scale”. Of these seven articles, three had established scoring systems simply for the purpose of the research undertaken and only four papers referred to a scoring system other than BI-RADS used in clinical practice (three from Norway, one from Ireland) [10–13].

The mapping of our UK data to the BI-RADS classification paves the way for more comparable research between the UK, North America and Europe and facilitates data sharing. The most significant remaining difference between the two systems is the lack of percentage cancer likelihood associated with each BI-RADS category which is not part of the UK RCRBG scoring system. Although this study uses data produced prior to RCRBG formalisation, comparisons are nevertheless possible between standard UK scoring and BI-RADS categories.

The lower BI-RADS Categories 1 and 2 are associated with a very low likelihood of malignancy. Although our results for standard UK Score 1 shows a percentage likelihood of cancer that is quantifiable in comparison to the 0% assigned to BI-RADS Categories 1 and 2, this represents a realistic view of the clinical setting where a negative imaging result does not always exclude a cancer diagnosis. We think it is therefore reasonable to assign standard UK Score 1 lesions to BI-RADS 1 and for BI-RADS 2 (benign) to encompass some of our standard UK Score 2 findings where there are no suspicious features, such as lipomas, simple cysts and unequivocally benign lesions, such as fibroadenomas with “popcorn” calcification.

The rest of the standard UK Score 2 lesions can be assigned to BI-RADS Category 3 with a comparable percentage likelihood of malignancy. These lesions have a low likelihood of malignancy and comply with neither BI-RADS Category 2 nor 4 or 5. Lesions with the typical features of fibroadenomas best typify this category. This study demonstrates that the ultrasound appearances of such a lesion (a well-defined lesion that is homogeneously hypoechoic with posterior acoustic enhancement) were the most common RCRBG Score 2 features and had a percentage likelihood of malignancy of 2.1% (Table 4) *i.e.* comparable with BI-RADS 3. Table 4 shows some of these Category 2 lesions having a higher percentage cancer likelihood than the 2% advocated for BI-RADS 3. This possibly represents erroneous lesion scoring by some operators, which will be reduced with implementation of the recommended classification system formulated by the RCRBG.

While, outside the UK, presumed fibroadenomas are frequently managed by repeat ultrasound surveillance, the protocol in the UK is to manage them on an age basis [1]. Frequently, women less than 25 years old (and in this centre <30 years old) are discharged without core biopsy and those aged 25 years and over are routinely biopsied. This represents an advantageous flexibility of the UK RCRBG scoring system over BI-RADS and we would not advocate a change in this approach. Accurate comparison of the UK RCRBG scoring system with the BI-RADS classification system is possible without changing current UK management practices.

The standard UK Score 3 and 4 groups have comparable percentage cancer likelihood with the BI-RADS 4 category which encompasses a wide range of cancer risk on a sliding scale of increased suspicion from 2–95% and subdivision into a, b and c has been suggested [4]. Although not part of the formal BI-RADS classification, the subdivisions can be useful in assessing the likelihood of malignancy and increasing the accuracy of planning patient management prior to needle biopsy result. The UK RCRBG Score 3 can be equated with BI-RADS 4a/b. BI-RADS 4c correspond with the UK RCRBG scoring 4 where there is a high level of suspicion but insufficient to warrant a definitely malignant scoring. There are no published figures for cancer incidence within each subcategory but inclusion criteria are suggested (Table 1). The standard UK Score 5 group can be assigned to the BI-RADS 5 category and encompass the highly suspicious lesions with two or more suspicious findings.

With no standard UK descriptors for ultrasound features and their associated likelihood of malignancy, it is currently difficult to establish uniformity of reporting and the sharing of data. However, we have made an attempt to calculate the malignant risk of certain features (Table 4) to assist in the correlation between current UK RCRBG scoring and the BI-RADS glossary of descriptors. We have shown there is similarity in percentage cancer risk between the sonographic features associated with our UK scored groups and BI-RADS categories, and although accurate use of the latter requires interoperator correlation and training, the BI-RADS lexicon promotes a more structured assessment of feature combinations and resultant estimate of associated malignancy risk [4, 14].

Conclusion

We have calculated the potential malignant risk associated with the recently published UK RCRBG five category reporting system. This has enabled us to successfully map the five UK categories to BI-RADS classification with equivalent cancer risk. Although these results relate to a single institution and cancer likelihood may differ between units, they nevertheless provide a benchmark for other units to compare their own findings. The ability to accurately assign UK RCRBG score lesions to a comparable BI-RADS category gives us the ability to share comparable research data and uniform practice in the international arena.

Acknowledgments

The authors were supported with research funding from the National Institute for Health Research (NIHR) Cambridge Biomedical Research Centre (BMRC).

The JCIS team are funded by Cancer Research UK and NIHR Cambridge BMRC. We would particularly like to thank Jackie Foreman the Cambridge Breast Cancer Research Unit database manager for extracting the relevant patient information.

References

1. Maxwell AJ, Ridley NT, Rubin G, Wallis MG, Gilbert FJ, Michell MJ. The Royal College Of Radiologists Breast Group breast imaging classification. *Clin Rad* 2009;64: 624–7.
2. Breast imaging reporting and data system (BI-RADS) atlas, 4th edn. Reston, VA: American College Of Radiology; 2003.
3. Stavros T, Thickman D, Rapp C, Dennis M, Parker S, Sisney G. Solid breast nodules: Use of sonography to distinguish between benign and malignant lesions. *Radiology* 1995; 196:123–34.
4. Lazarus E, Mainiero M, Schepps B, Koelliker S, Livingston L. BI-RADS lexicon for US and mammography: interobserver variability and positive predictive value. *Radiology* 2006;239: 385–91.
5. Guidance on screening and symptomatic breast imaging, 4th edn. London, UK: The Royal College Of Radiologists 2003.
6. Guidelines for the management of symptomatic breast disease ESJO 31: S1–S21. London, UK: The Association Of Breast Surgery (BASO) 2005.
7. Zonderland H, Pope TL, Nieborg AJ. The positive predictive value of the breast imaging reporting and data system (BI-RADS) as a method of quality assessment in breast imaging in a hospital population. *Eur Radiol* 2004;14:1743–50.
8. Costantini M, Belli P, Lombardi R, Franceschini G, Mule A, Bonomo L. Characterization of solid breast masses: Use of the sonographic breast imaging and reporting data system lexicon. *J Ultrasound Med* 2006;25:649–59.
9. Kim EK, Ko KH, Oh KK, Kwak JY, You JK, Kim MJ, et al. Clinical application of the BI-RADS final assessment to breast sonography in conjunction with mammography. *AJR Am J Roentgenol* 2008;190:1209–15.
10. Hofvind S, Geller BM, Rosenberg RD, Skaane P. Screen detected breast cancers: discordant independent double reading in a population based screening program. *Radiology* 2009;253:652–660.
11. Skaane P, Hofvind S, Skjennald A. Randomised trial of screen-film versus full-field digital mammography with soft-copy reading in population-based screening program: Follow up and final results of Oslo II study. *Radiology* 2007;244:708–17.
12. Skaane P, Kshirsagar A, Stapleton S, Young K, Castellino RA. Effect of computer-aided detection on independent double reading of paired screen-film and full-field digital screening mammograms. *AJR Am J Roentgenol* 2007;188: 377–84.
13. Hambly N, McNicholas M, Phelan N, Hargaden G, O'Doherty A, Flanagan F. Comparison of digital mammography and screen film mammography in breast cancer screening: A review in the Irish breast screening program. *AJR Am J Roentgenol* 2009;193:1010–18.
14. 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 Am J Roentgenol* 2000;174:1769–77.