Ultrasound in Med. & Biol., Vol. ■, No. ■, pp. 1–8, 2016

0301-5629/$ - see front matter

ELSEVIER

http://dx.doi.org/10.1016/j.ultrasmedbio.2016.06.017

● *Original Contribution*

INTER- AND INTRA-OBSERVER AGREEMENT IN ULTRASOUND BI-RADS CLASSIFICATION AND REAL-TIME ELASTOGRAPHY TSUKUBA SCORE ASSESSMENT OF BREAST LESIONS

Fabienne Schwab,* Katharina Redling,* Matthias Siebert,* Andy Schötzau,† Cora-Ann Schoenenberger,‡ and Rosanna Zanetti-Dällenbach*

*Department of Gynecology and Obstetrics, University Hospital of Basel, Basel, Switzerland; † Statistics, Department of Gynecology and Obstetrics, University Hospital of Basel, Basel, Switzerland; and ‡ Department of Chemistry, University of Basel, Basel, Switzerland

(*Received* 11 *January* 2016; *revised* 2 *May* 2016; *in final form* 8 *June* 2016)

Abstract—Our aim was to prospectively evaluate inter- and intra-observer agreement between Breast Imaging Reporting and Data System (BI-RADS) classifications and Tsukuba elasticity scores (TSs) of breast lesions. The study included 164 breast lesions (63 malignant, 101 benign). The BI-RADS classification and TS of each breast lesion was assessed by the examiner and twice by three reviewers at an interval of 2 months. Weighted κ values for inter-observer agreement ranged from moderate to substantial for BI-RADS classification (κ = 0.585–0.738) and was substantial for TS (κ = 0.608–0.779). Intra-observer agreement was almost perfect for ultrasound (US) BI-RADS (κ = 0.847–0.872) and TS (κ = 0.879–0.914). Overall, individual reviewers are highly self-consistent (almost perfect intra-observer agreement) with respect to BI-RADS classification and TS, whereas inter-observer agreement was moderate to substantial. Comprehensive training is essential for achieving high agreement and minimizing the impact of subjectivity. Our results indicate that breast US and real-time elastography can achieve high diagnostic performance. (E-mail: rosanna.zanetti@usb.ch)

Key Words: **Breast Imaging Reporting and Data System, BI-RADS, Tsukuba score, Breast ultrasound, Inter-observer agreement, Intra-observer agreement, Breast elastography.**

INTRODUCTION

In breast ultrasonography (US), lesions are classified on the basis of their sonographic features into categories 2–6 according to the Breast Imaging Reporting and Data System (BI-RADS) (American College of Radiology [ACR] 2013). Solid breast lesions are characterized according to shape (oval, round, irregular); orientation (parallel, not parallel); margins (circumscribed, not circumscribed); echo pattern (anechoic, hyper-echoic, complex cystic and solid, hypo-echoic, iso-echoic and heterogeneous); and posterior acoustic features. Benign masses are round or oval and wider than tall, with smooth, defined margins. In contrast, malignant masses tend to be irregular, with ill-defined to spiculated margins, and are taller than wide. In addition, BI-RADS classifies solid breast lesions according to their risk of malignancy. Cysts, solid lesions with unchanged or diminishing size and benign features in follow-up, as well as solid lesions with known histology, are classified as BI-RADS 2 or benign lesions. A solid breast lesion without any suspicious features is considered BI-RADS 3 or probably benign, whereas BI-RADS 4 indicates a suspicious finding and BI-RADS 5 is highly suggestive of malignancy. A known biopsy-proven malignancy is classified as BI-RADS 6. A normal sonographic breast examination without any lesions is considered BI-RADS 1. An incomplete examination or the need for additional imaging evaluation is described as BI-RADS 0.

A number of factors, however, influence BI-RADS classification. For example, considerable overlap in the US appearance between malignant and benign breast lesions introduces observer variability. More significantly, US is an operator-dependent examination because the

Address correspondence to: Rosanna Zanetti-Dällenbach, Department of Gynecology and Obstetrics, University Hospital of Basel, Spitalstrasse 21, 4031 Basel, Switzerland. E-mail: rosanna.zanetti@usb.ch

detection, description and interpretation of breast lesions are based solely on the examiner (Berg et al. 2006). Operator dependence has been investigated by determining weighted κ values of US BI-RADS final assessment between different examiners (see Table 1 and references therein). An early study by Skaane et al. (1997) reported moderate agreement between examiners in their interpretation of breast US (κ = 0.48). Subsequently, Baker et al. (1999) confirmed a moderate (κ = 0.51) inter-observer agreement. These authors also explored intra-observer agreement, which they found to be substantial (κ = 0.66).

In addition to the morphologic features seen in ultrasound, the elastic properties of soft tissues of the breast can be used as a parameter for diagnostic procedures. For example, real-time elastography (RTE) explores the difference in stiffness in benign and malignant breast lesions compared with the surrounding normal parenchyma at the macroscopic scale (Itoh et al. 2006; Sadigh et al. 2012). More recently, indentation-type atomic force microscopy has emerged as a novel diagnostic tool that allows mapping of the stiffness of unadulterated breast biopsies at nanoscale resolution (Plodinec et al. 2012).

In RTE, strain data obtained from tissue displacement produced by external compression with the probe is used to form a strain image (Itoh et al. 2006). Harder areas exhibit less tissue displacement, whereas softer areas exhibit more displacement. Strain distribution, which is inversely related to tissue stiffness, is visualized as a color-coded map that is superimposed on the B-mode image of conventional US. To standardize the interpretation of strain images, Itoh et al. (2006) developed the Tsukuba elasticity score (TS). Like the BI-RADS lexicon for breast US, TS defines parameters to standardize the interpretation of RTE. According to the Tsukuba elasticity score illustrated in Figure 1, TS1 and TS2 represent benign breast lesions, TS3 probably benign, and TS4 and TS5 malignant.

In clinics, RTE is not used as the sole examination procedure, but as a non-invasive adjunct in combination with breast US to further classify breast lesions (Wojcinski et al. 2010). However, the corresponding strain image is influenced by the compression technique of the individual examiner (Ciurea et al. 2011). Similar to US, not only data acquisition, but also interpretation of strain images, is operator dependent. Thus, TS is vulnerable to inter-observer variability. For example, Yoon at al. (2011) reported moderate (κ = 0.46) inter-observer agreement for TSs of static images. However, in real-time examination, inter-observer agreement was decreased to fair (κ = 0.28).

Overall, inter- and intra-observer variations in visual classification in breast US and RTE raise concern regarding the diagnostic performance of these procedures. Whereas a limited number of studies exploring inter- and intra-observer agreement in B-mode US exist, corresponding studies in RTE are rare. Here, we describe a prospective single-center study investigating inter- and intra-observer agreement of US BI-RADS classification and TSs in RTE.

METHODS

The single-center prospective study was approved by the institutional review board and conducted according to good clinical practice guidelines. Women with a solid breast lesion who were at least 18 y of age and scheduled for ultrasound-guided invasive breast biopsy at the outpatient breast clinic of the Women's Hospital

Table 1. Studies investigating inter- and intra-observer agreement

	κ Value			
	BI-RADS final classification		BI-RADS final classification	
Reference	Inter-observer	Intra-observer	Inter-observer	Intra-observer
Berg et al. 2006	0.52			
Lee et al. 2008	0.53	0.72–0.79		
Abdullah et al. 2009	0.30			
Calas et al. 2010		0.37–0.69		
Berg et al. 2012	0.53–0.59			
Elverici et al. 2013	0.35	0.64–0.83		
Thomas et al. 2006a			0.73	
Thomas et al. 2006b			0.86	
Yoon et al. 2011	0.37		0.28*/0.46	
Cho et al. 2011			0.587	
Cho et al. 2012			0.481	
Park et al. 2015	0.478		0.591	
Schafer et al. 2013	0.634	0.784	0.561	0.720
This study	0.585–0.738	0.847–0.872	0.608–0.779	0.879–0.914

* Real time.

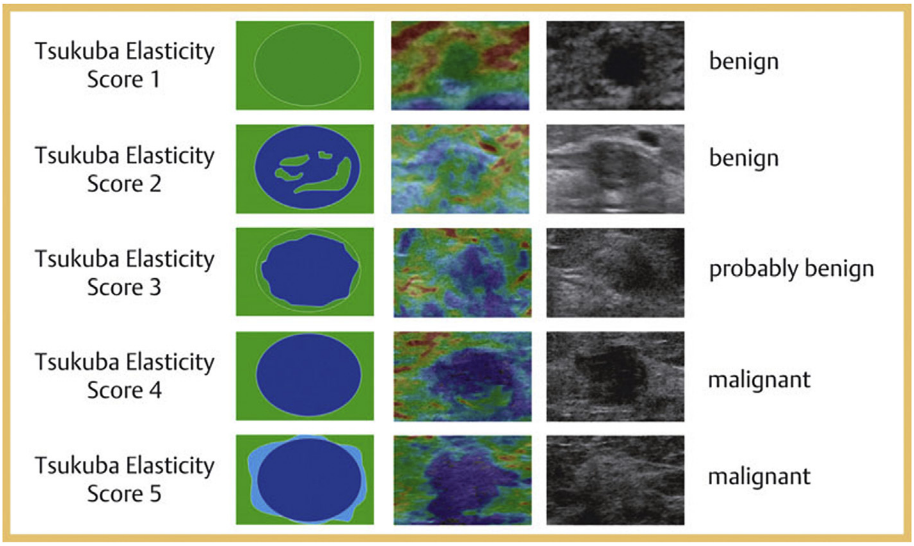

Fig.1. Tsukuba elasticity score. Breast elastography images are classified in five categories (Tsukuba elasticity scores 1–5) based on strain image patterns superimposed on B-mode images. Lesions scored 1 or 2 are considered benign, lesions scored 3 are considered to be probably benign and lesions scored 4 or 5 are considered malignant.

of the University Hospital, Basel, Switzerland, from August 2009 to December 2012 were recruited. After receiving full information about the study and providing written informed consent, patients underwent a breast workup including breast examination, bilateral whole-breast ultrasound and RTE of the lesion, mammography where indicated, and finally, ultrasound-guided breast biopsy. Histopathologic results of the ultrasound-guided breast biopsies were used as reference standards.

Sonographic and RTE data were acquired by examiners who routinely perform breast ultrasound examinations. Before the study began, they received 1 week training in RTE that included background theory as well as supervised elastography examinations. The primary examiners were aware of the clinical and, if available, mammographic findings at the time of examination. The ultrasound examination and RTE were performed with a 50-mm linear probe (EUP-L53 L) and a EUB-75 ooHV-LCD Hitachi scanner (Hitachi Medical Systems Europe Holding, Zug, Switzerland). For B-mode imaging, each lesion was scanned in transverse and sagittal planes. The dimensions of the lesions were recorded along three orthogonal planes, and the lesion volume was calculated. Based on the B-mode images, lesions were categorized according to BI-RADS (ACR 2003).

Real-time elastography of the lesion was performed immediately after recording the conventional B-mode image. The strain image was superimposed on the B-mode scan and displayed with color mapping from *red* (soft lesion with maximum strain) to *green* (intermediate strain) and *blue* (hard lesion with minimum strain). The area of interest was selected such that it spanned from subcutaneous fatty tissue to the superficial portion of the pectoralis muscle, avoiding the rib, and contained the breast lesion surrounded by sufficient normal tissue. Subsequently, a manual freehand cyclic compression technique was employed as described by Itoh et al. (2006). To avoid excessive or insufficient pressure on the tissue and to augment reproducibility, a pressure between 3 and 4 according to the pressure amplitude displayed by the Hitachi software (1 = lowest and 6 = highest pressure) was applied. Elastographic images were recorded and classified according to a 5-point scoring system known as the Tsukuba score (Fig. 1) (Itoh et al. 2006).

For each lesion, one representative B-mode image and one to five elastograms were recorded. Individual elastograms were acquired in series under comparable imaging conditions by the same examiner and saved for later review.

In addition to the assessment by the initial examiner, BI-RADS classification (ACR, 2003) of static B-mode images and Tsukuba scoring (Itoh et al. 2006) of static RTE images were repeated by three independent, experienced reviewers who were blinded to the clinical findings, as well as mammographic and histopathologic results. After at least 2 months, all images were re-assessed in random order by all three reviewers.

Data analysis

Patient characteristics and data derived from ultrasound examination, including BI-RADS classification and elasticity score, were retrieved from electronic medical records (View Point, Bildverarbeitung, Wessling, Germany) and entered in an Excel worksheet. Before statistical analysis, all data were anonymized.

To compare study parameters descriptively between study groups, the mean, standard deviation (SD), median, minimum (min) and maximum (max) were reported as appropriate. Additionally, p values of t-tests, Mann–Whitney U-tests and Fischer's exact tests were displayed as appropriate.

Inter- and intra-rater agreement was estimated using weighted κ values with corresponding 95% confidence intervals (CIs) (Cohen 1968). κ values were interpreted as poor ($\kappa < 0$), small/slight (κ = 0.0–0.20), fair (κ = 0.21–0.40), moderate (κ = 0.41–0.60), substantial (κ = 0.61–0.80) and almost perfect (κ = 0.81–1.00) agreement between examiners.

A p value <0.05 was considered to indicate significance. All evaluations were done using the statistical package R, Version 3.1.1 (R Development Core Team 2014).

RESULTS

Analysis of inter- and intra-observer agreement in US BI-RADS classification and TS included 613 elastograms of 164 breast lesions in 156 patients. Patients ranged from 18 to 89 y of age with a mean of 50.3 y. Patients with benign lesions were significantly younger than patients with malignant lesions. Core needle biopsy was performed on all 164 breast lesions. Histologic evaluation revealed 38.4% (n = 63) malignant and 61.6% (101) benign lesions. Among the benign lesions, 40.6% (n = 41) were fibroadenomas, 34.7% (n = 35) fibrocystic changes, 15.8% (n = 16) ductal hyperplasias and 8.9% (n = 9) other benign lesions. Among the malignancies, 77.8% (n = 49) were invasive ductal cancers, 12.7% (n = 8) invasive lobular cancers and 9.5% (n = 6) other types of breast cancer. The mean maximal diameter of the lesions was 17.7 mm (4.6–67.9 mm), and the mean lesion volume was 2.55 mL (0.05–30.1 mL). Benign and malignant lesions did not significantly differ in size (Table 2).

Initial BI-RADS classification by the examiner of all 164 lesions yielded 51.8% (n = 85) in BI-RADS 3, 26.8% (n = 44) in BI-RADS 4 and 21.3% (n = 35) in BI-RADS 5. The cancer rates for each BI-RADS classification are listed in Table 2. Benign lesions were significantly more often classified as BI-RADS 3 than malignant lesions, whereas malignant lesions were significantly more often classified as BI-RADS 4 or 5 than benign lesions ($p < 0.001$).

The elastograms were scored according to the Tsukuba elasticity score developed by Itoh and Ueno (Itoh et al. 2006). The 5-point scoring system is based on visual assessment of the degree and distribution of strain in the breast lesion and surrounding tissue (Fig. 1). TS 1 and 2 lesions are considered benign, TS 3 lesions probably benign and TS 4 and 5 lesions malignant. One to five elastograms of each lesion (n = 613) were recorded and scored. In the real-time assessment by the examiner, 9.5% (n = 58) of elastograms were assessed as TS 1, 60.8% (n = 373) as TS 2, 10.9% (n = 67) as TS 3, 4.6% (n = 28) as TS 4 and 14.2% (n = 87) as TS 5. The cancer rates for different TSs are listed in Table 2. Among the elastograms of benign lesions, 71.2%

Table 2. Patient and lesion characteristics

		Pathohistologic diagnosis		
Patient/lesion characteristics	All breast lesions (n = 164)	Benign	Malignant	p value
Age (y)				
Mean ± SD	50.3 ± 17.4	43.1 ± 14.5	61.2 ± 15.8	<0.001
Range, min–max	18–89	18–84	19–89	
Percentage (No.) of breast lesions	164	61.6% (101)	38.4% (63)	
Lesion volume (mL)				
Mean ± SD	2.55 ± 4.66	2.06 ± 4.28	3.35 ± 5.15	0.101
Range, min–max	0.05–30.1	0.06–30.1	0.05–26.8	
Mean maximum lesion ø (mm)				
Mean ± SD	17.7 ± 10.5	16.6 ± 9.08	19.1 ± 12.4	0.202
Range, min–max	4.6–67.9	5.2–57.0	4.6–67.9	
BI-RADS category by examiner, % (No.)				
3	85	96.5% (82)	3.5% (3)	<0.001
4	44	40.9% (18)	59.1% (26)	
5	35	2.9% (1)	97.1% (34)	
Tsukuba score by examiner, % (No.)	613	61.2% (375)	38.8% (238)	<0.001
1	58	96.5% (56)	3.5% (2)	
2	373	68.1% (254)	31.9% (119)	
3	67	65.7% (44)	34.3% (23)	
4	28	57.1% (16)	42.9% (12)	
5	87	5.7% (5)	94.3% (82)	

Table 3. Weighted κ values of inter-observer and intra-observer agreement in BI-RADS classification and Tsukuba score assessment

	BI-RADS			Tsukuba score		
	Reviewer 1	Reviewer 2	Reviewer 3	Reviewer 1	Reviewer 2	Reviewer 3
Inter-observer						
Examiner	0.585 (0.485–0.686)	0.738 (0.664–0.811)	0.622 (0.540–0.705)	0.693 (0.649–0.738)	0.608 (0.557–0.658)	0.689 (0.643–0.735)
Reviewer 1		0.678 (0.601–0.754)	0.559 (0.474–0.645)		0.753 (0.714–0.793)	0.779 (0.745–0.813)
Reviewer 2			0.701 (0.618–0.783)			0.750 (0.715–0.786)
Intra-observer						
Reviewer 1	0.847 (0.768–0.927)			0.891 (0.870–0.913)		
Reviewer 2		0.852 (0.802–0.902)			0.879 (0.857–0.901)	
Reviewer 3			0.872 (0.813–0.930)			0.914 (0.895–0.933)

(n = 354) had a TS between 1 and 3, whereas 81.7% (n = 94) of images of malignant lesions had a TS of 4 or 5. The assignment of benign lesions to TS 1–3 versus 4 and 5 and that of malignant lesions to 4 and 5 versus 1–3 are statistically significant ($p < 0.001$).

To assess inter-observer reproducibility of US BI-RADS classification and TSs, static B-mode images and elastograms were independently evaluated in random order by three reviewers who were blinded to all clinical data, in addition to the primary examiner. Weighted κ values for inter-observer agreement are summarized in Table 3. For inter-observer agreement of US BI-RADS classification between the examiner and each of the three reviewers, κ values between 0.585 and 0.738 were calculated, which correspond to a moderate to substantial agreement. κ values between reviewers ranged from 0.559 to 0.701, also representing moderate to substantial inter-observer agreement. The κ values for inter-observer agreement of TSs ranged from 0.608 to 0.693 (examiner vs. reviewer) and from 0.750 to 0.779 (inter-reviewer). Theses κ values represent substantial inter-observer agreement.

To determine the intra-observer reliability of US BI-RADS classification and elasticity scoring for each reviewer, the κ values were calculated based on two sequential reviews of B-mode images and elastograms that were carried out in random order with an interval of 2 months between reviews (Table 3). The κ values for US BI-RADS ranged from 0.847 to 0.872, and those for TS, from 0.879 to 0.914, reflecting almost perfect intra-observer agreement for both image evaluations.

DISCUSSION

Breast US is a well-established tool in the diagnostic imaging of lesions despite its operator dependency. In RTE, different compression techniques of individual examiners add to the variability of elastograms. However, for clinical application, a minimal reproducibility of the lesion assessment is required. Therefore, we comprehensively evaluated inter- and intra-observer agreement on US BI-RADS classification and on TSs of 156 breast lesions. As discussed below, there exist a number of studies on the inter- and/or intra-observer agreement of either US BI-RADS classification or Tsukuba score assessment (see Table 1). In contrast, our study at the same time analyzes inter- and inter-observer agreement for US BI-RADS and TSs, which allows comparison of the reliability of conventional US and RTE.

The earliest κ values published for US before implementation of US BI-RADS classification ($\kappa = 0.48$) indicate moderate inter-observer agreement (Skaane et al. 1997). For US BI-RADS classification, inter-observer agreement ranging from fair to substantial has been reported (Table 1).

Our data for US BI-RADS inter-observer agreement between the examiner and each of the three reviewers (κ = 0.585–0.738) and among the reviewers (κ = 0.559–0.701) indicated moderate to substantial agreement and are in line with the published data.

With the exception of Elverici et al. (2013), who reported κ values indicating substantial to almost perfect agreement, the majority of studies on intra-observer agreement for US BI-RADS reveal substantial agreement (Table 1). With κ values between 847 and 0.872, our data indicate almost perfect intra-observer agreement on US BI-RADS classification for all reviewers. In contrast, Calas et al. (2010) found that among eight reviewers, the intra-observer agreement of the most experienced reviewer was only fair (κ = 0.37) (Calas et al. 2010).

In agreement with other studies (Table 1), our data indicate better intra-observer than inter-observer agreement on US BI-RADS classification. It appears that reviewers are consistent with respect to lesion description and interpretation of B-mode images, which results in an almost perfect intra-observer agreement on BI-RADS classification. This standardization differs among individuals, and reduced inter-observer compared with intra-observer agreement is observed.

As seen in other studies (Table 1), our data revealed higher κ values for TSs than for US BI-RADS. Moderate to almost perfect inter-observer agreement for TSs has been reported (Table 1). Our data for TS inter-observer agreement between the examiner and each of the three reviewers (κ = 0.608–0.693) and among the three reviewers (κ = 0.750–0.779) indicated substantial agreement.

To our knowledge, intra-observer agreement for TSs has so far only been studied by Schaefer et al. (2011). Similar to our study, Schaefer et al. reported better intra-observer than inter-observer agreement for TSs. However, they reported substantial agreement (κ = 0.72), whereas our data indicated almost perfect agreement (κ = 0.879–0.914). More significantly, our results differ from those of Schaefer et al. in that we obtained higher κ values for inter-observer agreement in TSs compared with BI-RADS.

It has previously been reported that a 1-h didactic session in US BI-RADS classification improved inter-observer agreement (Berg et al. 2012). These data emphasize the importance of training in the interpretation and characterization of breast lesions. Consistent with this notion, the 1-week training in RTE at the onset of our study (see Appendix) might explain the substantial inter-observer agreement and almost perfect agreement for intra-observer agreement in TSs. Our results on inter-observer and inter-observer agreement for TSs, together with published data, indicate that the TS has high diagnostic performance.

One limitation of our study was that all patients were scheduled for breast biopsy, and therefore, no BI-RADS 2 lesions were included. In addition, retrospective review of B-mode and RTE images was based on static images.

CONCLUSIONS

Our data indicated moderate to substantial inter-observer agreement for US BI-RADS classification, substantial inter-observer agreement for TSs and almost perfect intra-observer agreement for US BI-RADS and TS assessment. Additional training and periodic performance evaluations may help to further improve observer agreement. Ideally, new study designs should incorporate real-time examination.

Acknowledgments—We thank Dominique Amy for RTE training sessions.

REFERENCES

Abdullah N, Mesurolle B, El-Khoury M, Kao E. Breast imaging reporting and data system lexicon for us: Interobserver agreement for assessment of breast masses. Radiology 2009;252:665–672.

American College of Radiology. ACR BI-RADS® Atlas ultrasound. In: Breast Imaging Reporting and Data System. 4th edition. Reston, VA: Author 2003.

American College of Radiology. ACR BI-RADS® Atlas. In: Breast Imaging Reporting and Data System. 5th edition. Reston, VA: Author 2013.

Baker JA, Kornguth PJ, Soo MS, Walsh R, Mengoni P. Sonography of solid breast lesions: Observer variability of lesion description and assessment. AJR Am J Rroentgenol 1999;172:1621–1625.

Berg WA, Blume JD, Cormack JB, Mendelson EB. Operator dependence of physician-performed whole-breast US: Lesion detection and characterization. Radiology 2006;241:355–365.

Berg WA, Blume JD, Cormack JB, Mendelson EB. Training the ACRIN 6666 investigators and effects of feedback on breast ultrasound interpretive performance and agreement in BI-RADS ultrasound feature analysis. AJR Am J Roentgenol 2012;199:224–235.

Calas MJ, Almeida RM, Gutfilen B, Pereira WC. Intraobserver interpretation of breast ultrasonography following the BI-RADS classification. Eur J Radiol 2010;74:525–528.

Cho N, Jang M, Lyou CY, Park JS, Choi HY, Moon WK. Distinguishing benign from malignant masses at breast US: Combined US elastography and color doppler US—Influence on radiologist accuracy. Radiology 2012;262:80–90.

Cho N, Moon WK, Chang JM, Yi A, Koo HR, Park JS, Park IA. Sonoelastographic lesion stiffness: Preoperative predictor of the presence of an invasive focus in nonpalpable DCIS diagnosed at US-guided needle biopsy. Eur Radiol 2011;21:1618–1627.

Ciurea AI, Bolboaca SD, Ciortea CA, Botar-Jid C, Dudea SM. The influence of technical factors on sonoelastographic assessment of solid breast nodules. Ultraschall Med 2011;32(Suppl 1):S27–S34.

Cohen J. Weighted kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. Psychol Bull 1968;70: 213–220.

Elverici E, Zengin B, Nurdan Barca A, Didem Yilmaz P, Alimli A, Araz L. Interobserver and intraobserver agreement of sonographic birads lexicon in the assessment of breast masses. Iran J Radiol 2013;10:122–127.

Itoh A, Ueno E, Tohno E, Kamma H, Takahashi H, Shiina T, Yamakawa M, Matsumura T. Breast disease: Clinical application of us elastography for diagnosis. Radiology 2006;239:341–350.

Lee HJ, Kim EK, Kim MJ, Youk JH, Lee JY, Kang DR, Oh KK. Observer variability of Breast Imaging Reporting and Data System (BI-RADS) for breast ultrasound. Eur J Radiol 2008;65:293–298.

Park CS, Kim SH, Jung NY, Choi JJ, Kang BJ, Jung HS. Interobserver variability of ultrasound elastography and the ultrasound BI-RADS lexicon of breast lesions. Breast Cancer 2015;22:153–160.

Plodinec M, Loparic M, Monnier CA, Obermann EC, Zanetti-Dallenbach R, Oertle P, Hyotyla JT, Aebi U, Bentires-Alj M, Lim RY, Schoenenberger CA. The nanomechanical signature of breast cancer. Nat Nanotechnol 2012;7:757–765.

R Development Core Team. R: A language and environment for statistical computing. Vienna: Foundation for Statistical Computing; 2014.

Sadigh G, Carlos RC, Neal CH, Dwamena BA. Ultrasonographic differentiation of malignant from benign breast lesions: A meta-analytic comparison of elasticity and birads scoring. Breast Cancer Res Treat 2012;133:23–35.

Schaefer FK, Heer I, Schaefer PJ, Mundhenke C, Osterholz S, Order BM, Hofheinz N, Hedderich J, Heller M, Jonat W, Schreer I. Breast ultrasound elastography—Results of 193 breast lesions in a prospective study with histopathologic correlation. Eur J Radiol 2011;77:450–456.

Schafer FK, Hooley RJ, Ohlinger R, Hahne U, Madjar H, Svensson WE, Balu-Maestro C, Juhan V, Athanasiou A, Mundinger A, Order B, Locatelli M, Cosgrove D, Wolf OJ, Henry JP, Moutfi M, Gay JM, Cohen-Bacrie C. Shearwave™ elastography BE1 multinational breast study: Additional SWE™ features support potential to downgrade BI-RADS®-3 lesions. Ultraschall Med 2013;34:254–259.

Skaane P, Engedal K, Skjennald A. Interobserver variation in the interpretation of breast imaging. Comparison of mammography, ultrasonography, and both combined in the interpretation of palpable noncalcified breast masses. Acta Radiol 1997;38:497–502.

Thomas A, Fischer T, Frey H, Ohlinger R, Grunwald S, Blohmer JU, Winzer KJ, Weber S, Kristiansen G, Ebert B, Kummel S. Real-time elastography–an advanced method of ultrasound: First results in 108 patients with breast lesions. Ultrasound Obstet Gynecol 2006a;28:335–340.

Thomas A, Kummel S, Fritzsche F, Warm M, Ebert B, Hamm B, Fischer T. Real-time sonoelastography performed in addition to B-mode ultrasound and mammography: Improved differentiation of breast lesions? Acad Radiol 2006b;13:1496–1504.

Wojcinski S, Farrokh A, Weber S, Thomas A, Fischer T, Slowinski T, Schmidt W, Degenhardt F. Multicenter study of ultrasound real-time tissue elastography in 779 cases for the assessment of breast lesions: Improved diagnostic performance by combining the BI-RADS®-US classification system with sonoelastography. Ultraschall Med 2010;31:484–491.

Yoon JH, Kim MH, Kim EK, Moon HJ, Kwak JY, Kim MJ. Interobserver variability of ultrasound elastography: How it affects the diagnosis of breast lesions. AJR Am J Roentgenol 2011; 196:730–736.

APPENDIX

All examiners underwent both theoretical and practical training in RTE. The theoretical introduction included two half-day training sessions by a Hitachi representative who explained and demonstrated RTE instrumentation. In addition, a radiologist experienced in RTE provided background information on correct application and interpretation of RTE. Practical training consisted of two half-day training sessions in which examiners carried out RTE in patients under the guidance and supervision of the expert radiologist.