

# Educational program and testing using images for the standardization of breast cancer screening by ultrasonography

Eriko Tohno · Hidehito Takahashi · Takayuki Tamada · Yasuhisa Fujimoto · Hidemitsu Yasuda · Noriaki Ohuchi

Received: 24 March 2010 / Accepted: 16 August 2010 / Published online: 6 October 2010  
© The Japanese Breast Cancer Society 2010

## Abstract

**Background** To improve the quality of breast cancer screening by ultrasonography, both effective training and evaluation of the performance of the examiners of breast ultrasound are essential.

**Methods** The Educational Committee of the Japan Association of Breast and Thyroid Sonology, an NPO, has established 2-day training programs on breast ultrasonography with tests at the end of the programs. The tests are performed using images to evaluate the ability of observers to detect and evaluate lesions on ultrasound. Ability to detect lesions was examined by using videos, and ability to evaluate lesions was examined by using still images. The

results of tests taken by 422 physicians and 415 technologists were analyzed.

**Results** In a comparison between physicians and technologists, the video specificity, the still image sensitivity, and the percentage of category agreement did not show any significant differences. The video sensitivity, the still image specificity, and the percentage of disease name agreement were significantly higher in technologists. Observers who had experienced <100 cases showed significantly poorer results in all subjects except for the video specificity in physicians and the still image specificity in technologists.

**Conclusions** Ultrasound technologists perform as well as physicians in recognizing and interpreting cancers on breast ultrasound, which supports their role in performing the initial screening examination.

---

E. Tohno (✉)  
Graduate School of Comprehensive Sciences,  
University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki, Japan  
e-mail: etohno@sakura.cc.tsukuba.ac.jp

H. Takahashi  
Epidemiology and Biostatistics, School of Medicine,  
University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki, Japan

T. Tamada  
School of Medicine, Cluster of Medical Sciences,  
University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki, Japan

Y. Fujimoto  
Tachibana Hospital, 4-3-18 Tachibana, Amagasaki,  
Hyogo, Japan

H. Yasuda  
Fourth Department of Surgery, International Medical Center  
of Japan, 1-21-1 Toyama, Shinjuku, Tokyo, Japan

N. Ohuchi  
Division of Surgical Oncology, Graduate School of Medicine,  
Tohoku University, 1-1 Seiryouchou Aoba-ku, Sendai,  
Miyagi, Japan

**Keywords** Breast · Screening · Ultrasonography · Education

## Introduction

Cancer screening by biennial mammography is recommended for women aged 40 years or above in Japan. Ultrasonography can detect breast cancers that escape mammography, and the supplemental contribution is particularly high in high-density breasts of young women. In Japan, a large-scale comparative study of whether the concomitant use of ultrasonography is effective for breast cancer screening in women in their 40s is in progress [1].

Ultrasonography is already used for breast cancer screening at many private health care organizations. Quality control is important for breast cancer screening by ultrasonography. While the equipment is an important

component of quality control, the ability of the examiner is even more important, because ultrasonographic detection and diagnosis of the lesions are performed in real-time. To improve the skill of the examiner for the future nationwide introduction of breast cancer screening by ultrasonography, the Educational Committee of the Japan Association of Breast and Thyroid Sonology (JABTS), an NPO, has organized 2-day training programs on breast ultrasonography for physicians and technologists with tests using images to evaluate the ability of the participants at the end of the programs. In this report, the test results were analyzed.

## Subjects and methods

Between April 2008 and March 2009, the Educational Committee of the JABTS sponsored training programs on breast ultrasonography 18 times (9 times for physicians, 9 times for technologists). The target population was physicians and technologists engaged, or expected to be engaged, in breast cancer screening by ultrasound and those at hospitals accepting secondary examinations of breast cancer screening. Each training program was performed over 2 days, with 49 physicians or 48 technologists participating at most. The subjects of this study were 422 physicians and 415 technologists. The specialties of the physicians and the numbers practicing each speciality were as follows: breast surgeons (196), surgeons of other or unspecified fields (118), gynecologists (37), radiologists (33), physicians working at screening institutions (21), and physicians of internal medicine (15); data were missing for 2. The technologists included medical technologists and radiographers; both are allowed to perform ultrasound in Japan. The experience levels were divided into four groups according to the numbers of ultrasound examinations performed during the past 5 years: <100, 100–499, 500–999, and  $\geq 1,000$ . The distribution of the physicians in the different experience levels was 66, 126, 74, and 156, respectively, and that of technologists was 105, 110, 53, and 147, respectively.

Table 1 shows a training program.

Tests using images were performed at the end of the training program using laptop-type personal computers. The observers answered using computers. The contents of the tests were as follows.

Fifty questions using videos (each 15–25 s long)

Based on the video images that mimicked ultrasonography and that did not pause at the lesion, participants were asked whether a lesion that requires secondary examination (category 3 or more advanced) was present or absent. The questions consisted of 25 videos including lesions of category 3 or greater and

**Table 1** Training program for physicians and technologists

Physicians	
Lectures for all participants	
	Basic knowledge of breast diseases (20 min) <sup>a</sup>
	Pathology of breast diseases (40 min)
	Ultrasonographic tissue characteristics of various breast diseases (40 min)
	Method of breast ultrasonography (30 min)
	Terminology of breast ultrasonography (mass image-forming lesions) (30 min)
	Terminology of breast ultrasonography (non-mass image-forming lesions) (30 min)
	Criteria for secondary examinations and method to describe findings on screening (30 min)
Group training (50 min, each)	
	Mass image-forming diseases, pt. 1
	Mass image-forming diseases, pt. 2
	Non-mass image-forming diseases
	Hands on, pt. 1 (scanning method, description of findings)
	Hands on, pt. 2 (interventional technique) <sup>a</sup>
	Skill of finding lesions
	Ultrasound images of various diseases
Tests	
	Tests using images (50 questions using videos, 50 questions using still images; 100 min)
Technologists	
Lectures for all participants	
	Basic knowledge of breast diseases (30 min) <sup>a</sup>
	Pathology of breast diseases (40 min)
	Ultrasonographic tissue characteristics of various breast diseases (40 min)
	Method of breast ultrasonography (30 min)
	Terminology of breast ultrasonography (mass image-forming lesions) (30 min)
	Terminology of breast ultrasonography (non-mass image-forming lesions) (30 min)
	Criteria for secondary examinations and method to describe findings on screening (30 min)
Group training (50 min)	
	Mass image-forming diseases, pt. 1
	Mass image-forming diseases, pt. 2
	Non-mass image-forming diseases
	Hands on, pt. 1 (scanning method, description of findings)
	Skill of finding lesions
	Ultrasound images of various diseases
Tests	
	Written test (25 questions in 40 min) <sup>a</sup>
	Tests using images (50 questions using videos, 50 questions using still images; 100 min)

<sup>a</sup> Differs between physicians and technologists

25 videos that included no lesions or included lesions of category 2.

Fifty questions using still images (each consisting of 1 or 2 orthogonal images)

Based on the still images, the category and disease considered most likely were selected. The questions included 18 sets of malignancies, 7 sets of benign lesions that need further examination, 17 sets of benign lesions that do not need recall (category 2), and 8 sets of normal breasts (including normal variations).

The videos were played using Windows Media Player. The observers were taught how to use the software during group training sessions. The answers (recommended categories and disease names) to questions presented using still images were determined based on a conference involving two doctors with more than 20 years of experience with breast ultrasound and one technologist with more than 10 years of experience. Table 2 defines the categories we use in Japan in comparison with BIRADS categories. Figure 1 and Table 3 show the criteria for category judgment. Table 4 shows the choices of disease names presented in the tests using still images. There were multiple possibly correct answers regarding both the category and disease name in still images, and the answer was regarded as correct if one of them was selected. The duration of tests using videos and those using still images combined was 100 min.

The following six items were evaluated for comparison:

1. Percentage of correct judgments of diseases on video images (video sensitivity)
2. Percentage of correct judgments of non-disease conditions on video images (video specificity)

**Table 2** Breast cancer categories and their definitions

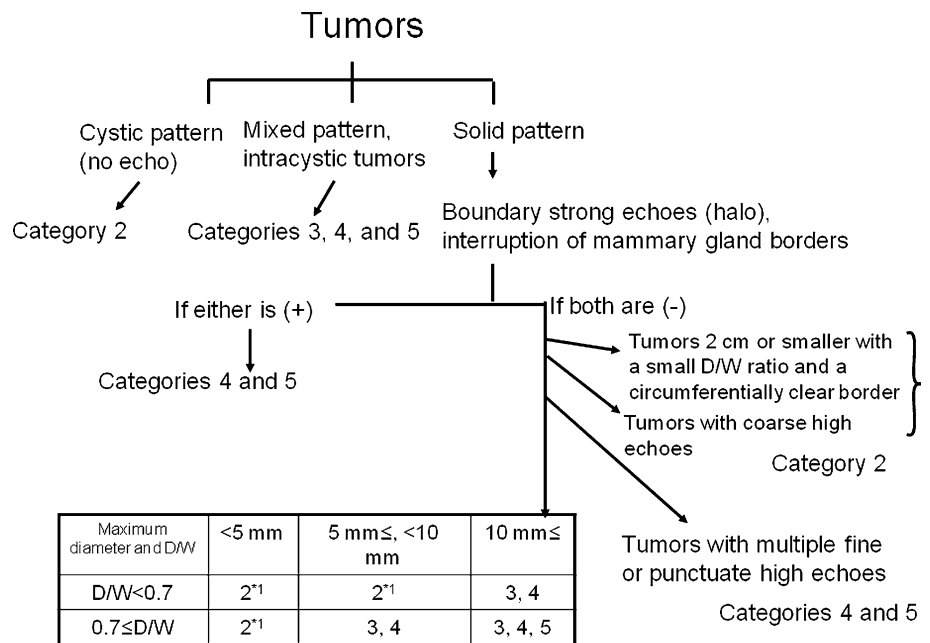
Category 1	No abnormality. Normal variations are included.
Category 2	Benign lesions not requiring secondary examinations on screening.
Category 3 (BIRADS category 3 and 4a)	Lesions that are probably benign but may also be malignant. They require further examinations including cytological or histological examinations or follow-up at hospitals.
Category 4 (BIRADS category 4b and 4c)	Lesions suspected to be malignant.
Category 5	Highly suggestive of malignancy—appropriate action should be taken.

The category classification in Japan is slightly different from that of BIRADS. Screening institutions usually do not follow-up the patient in Japan. If lesions that need either follow-up or further examinations are detected, the patients are referred to hospitals for secondary examinations

3. Percentage of correct judgments of diseases in still images (still image sensitivity)
4. Percentage of correct judgments of non-disease conditions in still images (still image specificity)
5. Percentage of category agreement for still images
6. Percentage of disease agreement for still images

Here, the video sensitivity and specificity are defined, respectively, as answering that a disease is “present” when viewing a video containing a category 3 or more advanced

**Fig. 1** Mass image-forming lesions



\*1: Tumors with irregular morphology may be classified as Category 3 or above.

**Table 3** Non-mass image-forming lesions

1. Unidirectional or segmental duct ectasia with internal echoes
(a) If the internal echoes sharply protrude → category 3
(b) If the internal echoes are broadly based → category 4 or 5
2. Localized or segmental hypoechoic (or isoechoic) areas with unclear borders → category 3, 4
If punctate high-echo spots suggestive of calcification are observed in the lesion, lesions of higher malignancy must be suspected → category 4 or 5
3. Clustered small cysts with localized or segmental distribution → category 3
4. Architectural distortion
(a) If its presence is questionable → category 3
(b) If its presence is definite → category 3, 4
(c) If its location coincides with the site of prior surgery → category 2

**Table 4** Choices of breast diseases

Normal or benign	Malignant
Normal (including normal variation)	Ductal carcinoma in situ
Fibroadenoma (including involuted fibroadenoma)	Scirrhous carcinoma <sup>a</sup>
Phyllodes tumor	Papillotubular carcinoma <sup>a</sup>
Intraductal papilloma	Solid-tubular carcinoma <sup>a</sup>
Cyst (including complicated cyst)	Invasive lobular carcinoma
Adenosis	Mucinous carcinoma
Sclerosing adenosis	Medullary carcinoma
Fibrosis	Squamous cell carcinoma
Duct ectasia	Inflammatory carcinoma
Fat necrosis (including oil cyst)	Malignant lymphoma
Panniculitis	
Mastitis	
Lipoma	
Hamartoma	
Atheroma of the skin	
Silicone granuloma	
Other benign mass	

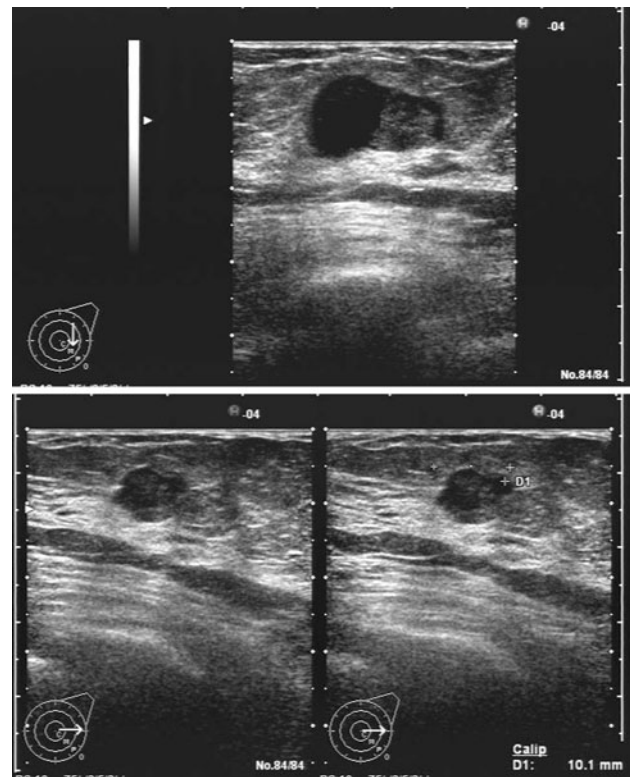
<sup>a</sup> Subgroups of infiltrating ductal carcinoma no special type (NOS) defined in the “General Rules for Clinical and Pathological Recording of Breast Cancer” by the Japanese Breast Cancer Society [7]

lesion and answering that a disease is “absent” when viewing a video not containing a category 3 or more advanced lesion. The still image sensitivity and specificity were defined, respectively, as the percentage of category 3 or more advanced lesions scored as 3 or higher and the percentage of category 2 or less advanced lesions scored as 2 or lower. Category and disease agreement for still images were defined, respectively, as the percentages of answers that agreed with the recommended categories in all cases

and the percentage of answers that agreed with the recommended disease names in all cases. Figure 2 shows an example of still image questions and answers and how the answers were evaluated.

Regarding each of these six items, the significance of the following differences was examined:

1. Difference between physicians and technologists
2. Differences among physicians in various fields of specialty such as breast surgery, other surgery, gynecology, radiology, working at screening institutions, and internal medicine; and between technologists working at a screening facility and those working at a hospital
3. Differences according to the number of patients the observers had examined by breast ultrasonography during the past 5 years (self-reported) (<100, 100–499, 500–999, and ≥1,000)



**Fig. 2** Example of still image tests. The recommended answers for this image set are category 3 or 4. Although this lesion was actually an intraductal papilloma, the differential diagnoses include intraductal papilloma or DCIS (non-invasive intracystic papillary carcinoma is classified as DCIS in Japan). The answers were evaluated as follows. Category 2, intraductal papilloma: sensitivity, false negative; category agreement, incorrect; disease agreement, correct. Category 3, intraductal papilloma: sensitivity, true positive; category agreement, correct; disease agreement, correct. Category 4, DCIS: sensitivity, true positive; category agreement, correct; disease agreement, correct. Category 5, DCIS: sensitivity, true positive; category agreement, incorrect; disease agreement, correct

**Table 5** Comparisons between physicians and technologists

	Physicians	Technologists	<i>p</i> value
Number	422	415	
Video sensitivity (%) (no of loss)	84.0	85.9 (1)	0.0370
Video specificity (%) (no of loss)	79.4	80.3 (1)	0.3469
Still image sensitivity (%)	95.9	95.8	0.7502
Still image specificity (%)	85.1	86.6	0.0260
Category agreement (%)	84.8	85.1	0.6157
Disease agreement (%)	78.4	81.1	<0.0001

#### 4. Differences according to age (20s, 30s, 40s, and 50s or above) of the observers

The differences between two groups such as physicians and technologists were examined by the *t* test, and those among three or more groups were examined by one-way ANOVA and Tukey's multiple comparison procedure. All statistical analyses were performed using SAS 9.13, with a significance level of 5%.

## Results

### Comparisons between physicians and technologists

The video sensitivity was 84.0% in physicians and 85.9% in technologists, being significantly higher in technologists

( $p = 0.037$ ) (Table 5). The still image specificity was 85.1% in physicians and 86.6% in technologists, being significantly higher in technologists ( $p = 0.026$ ). The percentage of disease name agreement was 78.4% in physicians and 81.1% in technologists, being significantly higher in technologists ( $p < 0.0001$ ).

### Comparisons according to physician fields of specialty

Among physicians, breast surgeons and radiologists showed better performance in many measures than doctors from other fields, especially gynecologists (Table 6). No significant difference was noted in the still image sensitivity.

No significant difference was observed between technologists working at a screening facility and those working at a hospital.

### Comparisons according to the number of patients the subjects examined by breast ultrasonography during the past 5 years (self-reported)

The video sensitivity improved with increases in the number of patients for both physicians and technologists (Table 7). In physicians, video sensitivity was 76.1% in those who had examined <100 patients and 85.5% in those who had examined 100 or more, with a significant difference ( $p < 0.0001$ ). In technologists, it was 79.8% in those who had examined <100 patients and 87.9% in those who had examined 100 or more, again showing a significant

**Table 6** Differences among physicians in various fields of specialty and between technicians working at a screening facility and those working at a hospital

	Physicians ( <i>N</i> = 420) <sup>a</sup>						<i>p</i> value (global)	Multiple comparison ( $p < 0.05$ )	Technologists ( <i>N</i> = 415)		<i>p</i> value
	Breast surgery 1	Other surgery 2	Gynecology 3	Radiology 4	Screening 5	Internist 6			Screening facility	Hospital	
Number	196	118	37	33	21	15			130	285	
Video sensitivity (%)	86.9	81.1	74.9	87.9	84.2	82.7	<0.0001	1-2, 1-3, 3-4	85.6	86.0	0.7499
Video specificity (%)	81.1	77.6	73.2	84.5	77.3	79.5	0.0027	1-3, 3-4	80.1	80.4	0.8394
Still image sensitivity (%)	96.6	95.3	95.0	96.1	95.4	94.1	0.1831		95.1	96.1	0.0690
Still image specificity (%)	87.8	83.2	77.6	86.4	84.2	83.5	<0.0001	1-2, 1-3, 2-3, 3-4	86.5	86.7	0.8803
Category agreement (%)	86.9	83.2	79.3	86.2	83.0	84.7	<0.0001	1-2, 1-3, 3-4	84.6	85.3	0.4144
Disease agreement (%)	80.9	76.3	72.6	81.7	74.4	75.7	<0.0001	1-2, 1-3, 1-5, 3-4	80.9	81.2	0.8257

<sup>a</sup> Speciality was missing in two doctors

**Table 7** Differences according to the number of patients the subjects had examined by breast ultrasonography during the past 5 years

	1 <100 patients	2 100–499 patients	3 500–999 patients	4 ≥1,000 patients	<i>p</i> value (global)	Multiple comparison ( <i>p</i> < 0.05)
Physicians ( <i>N</i> = 422)						
Number	66	126	74	156		
Video sensitivity (%)	76.1	82.4	84.3	88.5	0.0019	1-2, 1-3, 1-4, 2-4
Video specificity (%)	76.2	80.1	78.3	80.7	0.1162	
Still image sensitivity (%)	93.3	95.8	96.9	96.7	<0.0001	1-2, 1-3, 1-4
Still image specificity (%)	81.7	85.0	82.4	88.0	<0.0001	1-4, 2-4, 3-4
Category agreement (%)	81.0	84.2	84.0	87.3	<0.0001	1-2, 1-4, 2-4, 3-4
Disease agreement (%)	73.0	77.7	77.6	81.7	<0.0001	1-2, 1-3, 1-4, 2-4, 3-4
Technologists ( <i>N</i> = 415)						
Number	105	110	53	147		
Video sensitivity (%)	79.8	87.2	87.2	88.7	<0.0001	1-2, 1-3, 1-4
Video specificity (%)	73.1	81.1	82.6	84.1	<0.0001	1-2, 1-3, 1-4
Still image sensitivity (%)	95.0	95.3	96.5	96.5	0.0773	
Still image specificity (%)	84.1	86.1	87.6	88.4	0.0054	1-4
Category agreement (%)	82.4	84.5	86.5	87.0	<0.0001	1-3, 1-4
Disease agreement (%)	76.6	80.6	83.0	84.0	<0.0001	1-2, 1-3, 1-4, 2-4

difference ( $p < 0.0001$ ). The video specificity was 73.1% in technologists who had examined <100 patients and 82.8% in those who had examined 100 or more, with a significant difference ( $p < 0.0001$ ).

The still image sensitivity was 93.3% in physicians who had examined <100 patients and 96.4% for those who had examined 100 or more, with a significant difference ( $p < 0.0001$ ). The still image specificity was 81.7% in physicians who had examined <100 patients, being significantly lower than the 85.7% in those who had examined 100 or more ( $p = 0.0037$ ) and 88.0% for those who had examined 1,000 or more patients ( $p < 0.0001$ ). The value in those who had examined 1,000 or more patients was significantly higher than the 85.0% in those who had examined 100–499 ( $p = 0.0389$ ) and 82.4% for those who had examined 500–999 patients ( $p = 0.0002$ ). In technologists, the still image specificity was significantly higher at 88.4% in those who had examined 1,000 or more patients than the value in those who had examined <100 ( $p = 0.0033$ ).

The percentage of category agreement was highest in physicians who had examined 1,000 or more patients at 87.3% and was significantly higher than the 81.0% in those who had examined <100 ( $p < 0.0001$ ) and 84.2% in those who had examined 100–499 patients ( $p = 0.004$ ). In technologists, also, the percentage of category agreement was 82.4% in those who had examined <100 patients, being significantly lower than the 85.1% in those who had examined 100 or more patients ( $p < 0.0001$ ), 86.5% in those who had examined 500–999 patients ( $p = 0.0084$ ), and 87.0% in those who had examined 1,000 or more patients ( $p < 0.0001$ ).

The percentage of disease name agreement was 73.0% in physicians who had examined <100 patients, being significantly lower than the 79.4% in those who had examined 100 or more ( $p < 0.0001$ ). It was also lower than the 77.7% in those who had examined 100–499 patients ( $p = 0.0102$ ), 77.6% in those who had examined 500–999 patients ( $p = 0.0297$ ), and 81.7% in those who had examined 1,000 or more patients ( $p = 0.0001$ ). The results were also better in those who had examined 1,000 or more patients than in those who had examined 100–499 patients ( $p = 0.0029$ ).

In technologists, the percentage of disease name agreement was 76.6% in those who had examined <100 patients, being significantly lower than the 81.1% in those who had examined 100 or more ( $p < 0.0001$ ). It was also lower than the 80.6% in those who had examined 100–499 patients ( $p = 0.0079$ ), 83.0% in those who had examined 500–999 patients ( $p = 0.003$ ), and 84.0% in those who had examined 1,000 or more patients ( $p < 0.0001$ ). It was also higher in those who had examined 1,000 or more patients than in those who had examined 100–499 ( $p = 0.0201$ ).

#### Comparisons according to age

In physicians, the video specificity was 81.8% in those aged 50 years and above, being significantly lower than in those aged <50 years ( $p < 0.0001$ ) (Tables 8, 9). It was also significantly lower than the 82.4% in those in their 30s ( $p < 0.0001$ ) and 81.3% in those in their 40s ( $p < 0.0001$ ). In technologists, video sensitivity and video specificity were not significantly different for those aged 50 years and

**Table 8** Differences according to physician or technologist age

	1 20s	2 30s	3 40s	4 50s or above	<i>p</i> value (global)	Multiple comparison ( <i>p</i> < 0.05)
Physicians ( <i>N</i> = 419) <sup>a</sup>						
Number	30	118	149	122		
Video sensitivity (%)	81.2	85.3	86.4	80.3	<0.0001	2-4, 3-4
Video specificity (%)	81.9	82.4	81.3	73.2	<0.0001	1-4, 2-4, 3-4
Still image sensitivity (%)	94.4	96.9	96.7	94.4	<0.0001	2-4, 3-4
Still image specificity (%)	84.4	85.0	87.7	82.2	<0.0001	3-4
Category agreement (%)	84.1	85.2	87.0	82.0	<0.0001	2-4, 3-4
Disease agreement (%)	78.0	79.2	80.9	74.6	<0.0001	2-4, 3-4
Technologists ( <i>N</i> = 415)						
Number	104	151	129	31		
Video sensitivity (%)	83.3 (1)	87.0	87.6	81.9	0.0120	1-3
Video specificity (%)	79.8 (1)	81.3	80.9	74.8	0.1852	
Still image sensitivity (%)	95.8	96.2	95.8	93.5	0.1149	
Still image specificity (%)	87.0	87.7	86.0	83.2	0.1120	
Category agreement (%)	85.5	86.2	84.7	80.4	0.0023	1-4, 2-4, 3-4
Disease agreement (%)	80.8	82.4	80.7	77.5	0.0627	

<sup>a</sup> Age information was missing in three doctors

**Table 9** Comparison of differences between physicians and technologists according to age

	Physicians ( <i>N</i> = 419) <sup>a</sup>			Technologists ( <i>N</i> = 415)		
	Under 50	50s or above	<i>p</i> value (global)	Under 50	50s or above	<i>p</i> value (global)
Number	297	122		384	31	
Video sensitivity (%)	85.5	80.3	0.0034	86.2	81.9	0.1607
Video specificity (%)	81.8	73.2	<0.0001	80.8	74.8	0.1221
Still image sensitivity (%)	96.5	94.4	0.0020	96.0	93.5	0.1153
Still image specificity (%)	86.3	82.2	0.0003	86.9	83.2	0.1425
Category agreement (%)	86.0	82.0	<0.0001	85.5	80.4	0.0250
Disease agreement (%)	79.9	74.6	<0.0001	81.4	77.5	0.0656

<sup>a</sup> Age information was missing in three doctors

above compared with that in those aged <50 years. This is considered to have been partly due to the small number of subjects.

In physicians, the still image sensitivity was 94.4% in those aged 50 years or above, being significantly lower than the 96.5% in those aged <50 years ( $p = 0.0020$ ). It was also significantly lower than the 96.9% in those in their 30s ( $p = 0.0024$ ) and 96.7% in those in their 40s ( $p = 0.0037$ ). The still image specificity was highest at 87.7% in physicians in their 40s, being significantly higher than the 82.2% in those aged 50 years or above ( $p < 0.0001$ ). In technologists, no significant difference was noted in either still image sensitivity or still image specificity according to age.

The percentage of category agreement using still images was 82.0% in physicians aged 50 years or above, being significantly lower than the 84.8% in those aged <50 years

( $p = 0.0001$ ). It was also significantly lower than the 85.2% in those in their 30s ( $p = 0.0061$ ) and 87.0% in those in their 40s ( $p < 0.0001$ ). In technologists, the percentage of category agreement was 80.4% in those aged 50 years or above, also being significantly lower than the 85.5% in those aged <50 years ( $p < 0.0023$ ). It was also lower than the 85.5% in those in their 20s ( $p = 0.0078$ ), 86.2% in those in their 30s ( $p = 0.0012$ ), and 84.7% in those in their 40s ( $p = 0.0339$ ).

The percentage of disease name agreement using still images was 74.6% in physicians aged 50 years or above, being significantly lower than the 78.4% in those aged <50 years ( $p < 0.0001$ ). It was also lower than the 79.2% in those in their 30s ( $p = 0.002$ ) and 80.9% in those in their 40s ( $p < 0.0001$ ). In technologists, no significant difference was noted in the percentage of disease name agreement according to age.

## Discussion

In Japan, both the morbidity and mortality rates of breast cancer are increasing, and breast cancer became the most frequent type of cancer among Japanese women, overtaking stomach cancer in 1993 [2]. The incidence of breast cancer in Japanese women reaches a peak in the late 40s, unlike in Western countries. In Japan, screening for breast cancer has long been performed by palpation, but mammography was introduced after the results of randomized comparative studies in Western countries were reported. Today, breast cancer screening primarily by mammography has become widely available for women aged 40 years or above. However, the sensitivity of mammography screening for breast cancer in women in their 40s is 71.4%, which is lower than the figures in those in their 50s and 60s [3]. This is probably because many women in their 40s present high-density to extremely dense breasts on mammography. To increase the detection rate of breast cancer by screening, the use of MRI for high-risk groups is recommended in the United States [4]. In Japan, however, the concomitant use of ultrasonography, which can be performed at most hospitals, may be more convenient, and a large-scale comparative study is presently being conducted involving women in their 40s [1].

Since breast ultrasonographic diagnosis is performed in real-time, the ability of the examiner to detect lesions, evaluate them, and create appropriate records greatly affects the sensitivity and specificity of breast cancer screening. In Japan, it is stipulated that examination results must be evaluated by physicians but that breast ultrasonography can actually be performed by either a physician or a technologist. In the United States, the results of screening ultrasonography performed by physicians have been reported [5], but screening by technologists is considered to be advantageous from the viewpoints of the number of available examiners and cost. In this study, the video sensitivity was evaluated to assess the subjects' ability to detect breast cancer. Since it was significantly higher in technologists, screening by technologists is not considered to lead to the overlooking of lesions.

A low specificity is a problem with breast ultrasonography, although many lesions other than breast cancer are detected in this way. We previously performed tests using images before and after a 2-day training program and reported that the program was effective, because the sensitivity of screening using video and still images improved significantly after the program despite a slight but non-significant decrease in specificity [6]. In the program, the participants were trained to discriminate changes in the breast that should be eliminated from secondary examinations (clearly benign lesions or very small lesions that are likely to be benign, e.g., normal variations such as a

lactating breast and the interposition of fat, surgically enlarged breasts, cysts, and typical fibroadenoma). The results of this study were obtained from participants of a 2-day training program, who had acquired the same knowledge about the assessment of lesions. The results of this study indicate that technologists have an ability to read both still images and videos comparable to physicians and are sufficiently capable of conducting primary screening.

Among physicians of various specialties, gynecologists did poorly in the tests. This may be partly because they are slightly older than other physicians. The average ages of gynecologists and other specialties were 50 and 44 years, respectively. There is also the concept in breast cancer screening that secondary examinations should not be indicated for changes detected by breast ultrasonography unlikely to be cancer, which differs from the guidelines for uterine cancer screening. Gynecologists may not have been accustomed to the stricter elimination policy. There was no difference between technologists working at a screening facility and those working at a hospital. This probably suggests that technologists at a screening facility are also learning to categorize and identify diseases as well as acquiring skills for detecting lesions, but may also have been due to the difficulty in discriminating clearly between screening and hospital technologists, because some hospital technologists are working in screening departments.

In terms of the number of patients the subjects had examined, those who had experienced <100 cases did poorly in the tests. This suggests that learning in a training program is not sufficient and that clinical experience is important to improve the screening ability. This also shows that if an examiner has experience with fewer than 100 patients, the screening ultrasound should be done with the supervision of experienced examiners.

The results of the subjects aged 50 years or above were poor. Fatigue caused by having to watch the computer screen for a long time and the degree of familiarity with computer operation may have been related to this outcome, and considerations such as extending the testing time and allowing the taking of breaks may be necessary for this age group.

## Conclusion

We analyzed the results of tests using images after a breast ultrasonography training program. The test results were comparable between physicians and technologists and were poorer for participants with less clinical experience.

**Acknowledgments** The work was partially funded by Japan Strategic Anti-cancer Randomized Trial (J-START) (a project in the Third-term Comprehensive Strategy for Cancer Control).



## References

1. Ohuchi N, Suzuki A, Sakurai Y, Kawai M, Marikawa Y, Narimatsu H, et al. Current status and problems of breast cancer screening. *JMAJ*. 2009;52:45–9.
2. Cancer Statistics in Japan 2007. [http://ganjoho.ncc.go.jp/public/statistics/backnumber/2007\\_jp.html](http://ganjoho.ncc.go.jp/public/statistics/backnumber/2007_jp.html).
3. Saslow D, Boetes C, Burke W, Harms S, Leach MO, Lehman CD, et al. American Cancer Society guidelines for breast screening with MRI as an adjunct to mammography. *CA Cancer J Clin*. 2007;57:75–89.
4. Suzuki A, Kuriyama S, Kawai M, Amari M, Takeda M, Ishida T, et al. Age-specific interval breast cancers in Japan: estimation of the proper sensitivity of screening using a population-based cancer registry. *Cancer Sci*. 2008;99:2264–7.
5. Berg WA, Blume JD, Cormack JB, Mendelson EB, Lehrer D, Böhm-Vélez M, et al. Combined screening with ultrasound and mammography vs mammography alone in women at elevated risk of breast cancer. *JAMA*. 2008;14:2151–63.
6. Tohno E, Sawai K, Shimamoto K, Ueno E, Endou T, Tsunoda-Shimizu H, et al. Establishment of seminars to improve the diagnostic accuracy and effectiveness of breast ultrasound. *J Med Ultrasonics*. 2006;33:239–44.
7. Histological classification. *Breast Cancer*. 2005;12 Suppl:12–14.