Breast Cancer (2014) 21:542–549
DOI 10.1007/s12282-012-0434-2

ORIGINAL ARTICLE

Effect of screening mammography on cumulative survival of Japanese women aged 40–69 years with breast cancer

Masaaki Kawai · Akihiko Suzuki · Yoshikazu Nishino · Koji Ohnuki · Takanori Ishida · Masakazu Amari · Daisuke Shibuya · Noriaki Ohuchi

Received: 9 October 2012 / Accepted: 29 November 2012 / Published online: 14 December 2012

Abstract
Background The effectiveness of screening mammography (MMG) has mainly been demonstrated by studies in western countries. This study was conducted to evaluate cumulative survival and the risk of breast cancer death among Japanese women aged 40–69 years with screening-detected and interval breast cancer divided into three groups: MMG with clinical breast examination (CBE), CBE alone, and self-detection.
Methods By matching a list of 126,537 women (358,242 person-screenings) who participated in the Miyagi Cancer Society Screening program between 1 April 1995 and 31 December 2002 with the Miyagi Prefectural Cancer Registry, 429 MMG with CBE, 522 CBE, and 3,047 self-detected cases were included in this study. Follow-up was performed until the date of death or 31 December 2007. Survival was estimated by the Kaplan–Meier method. The Cox proportional hazards model was used to estimate hazard ratios (HR) and 95 % confidence intervals (CI) for breast cancer death.
Results Five-year survival for women in the MMG with CBE, CBE, and self-detection groups was 96.8, 92.7, and 86.6 %, respectively. The HR (95 % CI) for breast cancer death was 2.38 (0.72–7.94) among CBE-screened and 4.44 (1.42–13.89) among self-detected cases for women aged 40–49 years, but was 3.00 (1.63–5.50) among CBE-screened and 4.51 (2.69–7.56) among self-detected cases for women aged 50–69 years relative to cases screened by use of MMG with CBE.
Conclusions In terms of the survival and risk of breast cancer death, MMG with CBE may be more effective than MMG alone or self-detection for Japanese women aged 40–69 years.

Keywords Breast cancer · Mammography · Screening · Survival

M. Kawai · A. Suzuki · T. Ishida · M. Amari · N. Ohuchi (✉)
Department of Surgical Oncology, Tohoku University Graduate School of Medicine, 1-1 Seiryo-machi, Aoba-ku, Sendai, Miyagi 980-8574, Japan
e-mail: noriaki-ohuchi@med.tohoku.ac.jp

Y. Nishino
Division of Epidemiology, Miyagi Cancer Center Research Institute, 47-1 Nodayama, Medeshima-Shiode, Natori, Miyagi 981-1293, Japan

K. Ohnuki
Division of Breast Surgery, Iwate Prefectural Central Hospital, 1-4-1 Ueda, Morioka, Iwate 020-0066, Japan

D. Shibuya
Miyagi Cancer Society, 5-7-30 Kamisugi, Aoba-ku, Sendai, Miyagi 980-0011, Japan

Introduction

Breast cancer is one of the most common cancers worldwide. Among Japanese women, the age-standardized incidence of breast cancer has now risen to first place among all cancers, and it is increasing rapidly [1]. Furthermore, the age-specific incidence of breast cancer among Japanese women aged 45–49 years and mortality due to breast cancer among Japanese women aged 35–64 years are the highest for any type of cancer [1]. Therefore, screening mammography (MMG) is regarded as an important public health priority.

Randomized controlled trials (RCTs) conducted in western countries have clarified the effectiveness of MMG screening for women aged 40–69 years, and especially for

those aged 50–69 years [2]. Breast cancer screening by clinical breast examination (CBE) was introduced for Japan for women aged 30 years and over in 1987 in the absence of any evidence of its effectiveness [3]. Studies in Japan to evaluate the efficacy of screening using MMG with CBE compared with CBE alone revealed the former was superior to the latter in terms of sensitivity, specificity, and success of detection in women aged over 50 years [4, 5]. Based on the results of those studies, screening using MMG with CBE was endorsed in 2000 for women aged over 50 years, and in 2004 for those aged over 40 years. However, this initiative was based mainly on data obtained from RCTs of MMG screening in western countries [2]. The efficacy of screening using MMG with CBE for Japanese women was further examined by cost-effectiveness analysis based on actual screening data for those aged 40–49 years [6] and by a validation study of accurate false-negativity data for MMG with CBE screening [7]. Furthermore, our previous study revealed that the survival of women with MMG-detected breast cancer was superior to that of women with CBE-detected or self-detected breast cancer, especially for those aged 50–69 years, although the effectiveness of the screening program for women aged 40–49 years was not assessed at that time [8]. In relation to the effectiveness of the screening program, our previous study [8] may have included inherent bias, because it did not consider the presence of interval breast cancer [9], which may grow rapidly and have a poor outcome [10]. Therefore, to properly assess the effectiveness of MMG screening there is still a need to evaluate the survival and risks of breast cancer death among Japanese women aged 40 years and over with screening-detected and interval cancer [9].

For this purpose, this retrospective cohort study was conducted to clarify the efficacy of screening using MMG with CBE by investigating cumulative survival and the risk of breast cancer death among Japanese women aged 40–69 years with screen-detected and interval cancer by dividing them into groups according to the screening methods used (MMG with CBE, CBE alone, or self-detection) and stratifying the subjects according to age. Improvements in the survival of women with breast cancer and the risk of breast cancer death for MMG with CBE screening in comparison to CBE screening alone and self-detection were evaluated with reference to the Miyagi Prefectural Cancer Registry [11].

Materials and methods

The Miyagi Cancer Society has performed breast cancer screening for women in Miyagi prefecture since 1989 [4, 5]. In brief, women aged 50 years and over living in Miyagi prefecture underwent annual single-view MMG with CBE in 32 registered communities; initially CBE only was provided in another 27 communities for breast cancer screening (Miyagi trial). Women aged 40 years and over underwent annual single-view MMG with CBE or CBE for breast cancer screening in 1995 and biennial single-view MMG with CBE or CBE for breast cancer screening between 1996 and 2004 [6]. The process of transition from CBE to MMG with CBE depended on the decision of each community and was gradual. Screening MMG was performed with CBE, and the mammograms were reviewed for each subject by two physicians at the Cancer Detection Center of the Miyagi Cancer Society. CBE is defined as inspection and palpation of breasts and regional lymph nodes by the attending physician at the screening. Women with any abnormal findings detected by MMG with CBE, or by CBE alone, were referred to community hospitals or followed up at the Cancer Detection Center of the Miyagi Cancer Society [4, 5]. All results of diagnostic examinations were reported by the hospitals that had performed the diagnostic MMG and/or ultrasonography (biopsy and/or surgical operation if necessary). Screening-detected cancer was defined as a case diagnosed pathologically within 6 months after a positive screening test (detected case) [7]. Interval cancers were defined as cases that were diagnosed as non-malignant at the primary screening but then clinically diagnosed as breast cancer during the interval until the next screening was conducted [7].

The end-point of this analysis was the cumulative survival of women with screening-detected and interval breast cancer (for women who underwent MMG with CBE, or CBE alone) and the survival of women with self-detected breast cancer, defined as topography code C50.0–C50.9 of the International Classification of Disease for Oncology, second edition (ICD-O-2) [12]. In the Miyagi Prefectural Cancer Registry, the relevant patients were abstracted from the medical records of the hospitals by a physician or trained medical records reviewer, except for patients reported directly to the registry by an institution. The clinical staging system was that of the Research Group for Population-Based Cancer Registration in Japan, among the methods used for detection. Lesions were classified into five stages (in situ, localized, lymph node metastasis, regional invasion, or distant metastasis) on the basis of information about tumor extension and metastasis to lymph nodes and distant sites [13]. This clinical staging system was available for breast cancer from 1 April 1995. Between that date and 31 December 2004, 6,134 cases of primary breast cancer were registered. The percentage of cases registered by death certificates only (DCO) for breast cancer was 2.82 % (178/6,134 primary breast cancers). DCO cases were excluded from the analysis.

Matching of records from the screening program database with the Miyagi Prefectural Cancer Registry was conducted with the aid of registry officials, using name, address, and date of birth to identify individuals. By matching the cancer registry with the Miyagi Cancer Society Screening program for a total of 126,537 subjects (358,242 person-screenings) from 1 April 1995 to 31 December 2002, 662 screening-detected cases and 289 interval breast cancer cases in patients aged 40–69 years were found and included in this analysis. Among the remaining 5,005 cases, 450 were excluded because they were entered in the Miyagi Prefectural Cancer Registry as having been detected by other screening programs, and a further 1,508 cases were excluded because age at diagnosis was under 40 or over 70 years. The remaining 3,047 cases, registered as having been detected by other methods, or those for which the details were unknown, for women aged 40–69 years, were regarded as having been self-detected. Thus, a final total of 3,998 cases were included in this analysis.

The numbers of women in the 40–49, 50–59, and 60–69-year age groups were 1,545, 1,270, and 1,183, respectively. The screening methods (MMG with CBE, and CBE alone) used for each cancer patient were confirmed from the breast cancer database of the Miyagi Cancer Society Screening program. Self-detection is defined as a patient finding a lesion by herself, the lesion being later diagnosed as breast cancer. Finally, we separated the subjects into three groups (429 screened by MMG with CBE, 522 screened by CBE alone, and 3,047 with self-detected lesions).

Follow-up was conducted for each of the subjects from the date of diagnosis of breast cancer until the date of death or the end of follow-up (31 December 2007), whichever occurred first. Patients who died from causes other than breast cancer were treated as censored cases. Patients for whom no information on death was available were regarded as alive at the end of the follow-up period. On the basis of these data, the association between type of screening method used and patient outcome was analyzed. Kaplan–Meier survival analysis was performed for each screening group. Differences between survival in the two groups were assessed statistically by use of the log-rank test. The Cox proportional hazards regression model was used to estimate the hazard ratios (HR) and 95 % confidence intervals (CI) for relative mortality risk in comparison with the MMG with CBE screened group [14]. All statistical analysis was performed by use of SAS version 9.3 (SAS, Cary, NC, USA). All reported p values were two-sided, and differences were considered statistically significant at $p < 0.05$.

The study protocol was approved by the institutional review board of Tohoku University Graduate School of Medicine, the Miyagi Cancer Society, and the committee of the Miyagi Prefectural Cancer Registry. The study was conducted in accordance with the principles of the Declaration of Helsinki.

Results

Four-hundred and twenty-nine cancers (10.7 %) were detected by MMG with CBE, 522 (13.1 %) by CBE alone, and 3,047 (76.2 %) were self-detected. Among the cancers, 1,545 (38.6 %), 1,270 (31.8 %), and 1,183 (29.6 %) occurred in women aged 40–49, 50–59, and 60–69 years, respectively. The proportion of interval cancer was higher in women screened by CBE alone (40.0 %; 209/522) than in those screened by use of MMG with CBE (18.6 %; 80/429) (Table 1).

Among cancers detected by MMG with CBE, 85 (19.8 %) were in situ, 248 (57.8 %) were localized, 68 (15.9 %) were lymph node metastases, 6 (1.4 %) were regional invasion, 2 (0.5 %) were distant metastases, and the stages of 20 patients (4.7 %) were unknown. Among cancers detected by CBE alone, 64 (12.3 %) were in situ, 273 (52.3 %) were localized, 68 (24.3 %) were lymph node metastases, 9 (1.7 %) were regional invasion, 12 (2.3 %) were distant metastases, and the stages of 37 patients (7.1 %) were unknown. Among 3,047 (76.2 %) self-detected cancers, 157 (5.2 %) were in situ, 1,324 (43.5 %)

Table 1 Age distribution of the study subjects according to modality

Modality	Age group (years), *N*			Total		Median age (years)	SD
	40–49	50–59	60–69	*N*	%		
MMG with CBE	78	174	177	429	10.7	58.4	7.7
MMG detected	55	139	155	349		59.1	7.6
MMG interval	23	35	22	80		54.4	7.5
CBE	273	126	78	522	13.1	49.4	8.2
CBE detected	165	78	70	313		49.2	8.1
CBE interval	108	48	53	209		49.7	8.4
Self-detection	1,194	970	883	3,047	76.2	52.8	8.4
Total	1,545	1,270	1,183	3,998	100.0	53.0	8.4

MMG mammography, *CBE* clinical breast examination

Table 2 Cancer stages of the study subjects according to modality

Modality	Stage												Total	
	In situ		Localized		Lymph node metastasis		Regional invasion		Distant metastasis		Unknown			
	N	%	*N*	%	*N*	%	*N*	%	*N*	%	*N*	%	*N*	%
MMG with CBE	85	19.8	248	57.8	68	15.9	6	1.4	2	0.5	20	4.7	429	10.7
CBE	64	12.3	273	52.3	127	24.3	9	1.7	12	2.3	37	7.1	522	13.1
Self-detection	157	5.2	1,324	43.5	851	27.9	197	6.5	213	7.0	305	10.0	3,047	76.2
Total	306	7.7	1,845	46.1	1,046	26.2	212	5.3	227	5.7	362	9.1	3,998	100.0

MMG mammography, *CBE* clinical breast examination

Table 3 Cause of death of the study subjects according to modality

Modality	Status								Total	
	Alive		Breast cancer death		Other cancer death		Other causes			
	N	%	*N*	%	*N*	%	*N*	%	*N*	%
MMG with CBE	393	91.6	18	4.2	12	2.8	6	1.4	429	10.7
CBE	449	86.0	57	10.9	13	2.5	3	0.6	522	13.1
Self-detection	2,366	77.7	568	18.6	55	1.8	58	1.9	3,047	76.2
Total	3,208	80.2	643	16.1	80	2.0	67	1.7	3,998	100.0

MMG mammography, *CBE* clinical breast examination

were localized, 851 (27.9 %) were lymph node metastases, 197 (6.5 %) were regional invasion, 213 (7.0 %) were distant metastases, and the stages of 305 patients (10.0 %) were unknown (Table 2).

Among the patients whose cancers had been detected by MMG with CBE, 393 (91.6 %) were alive, 18 (4.2 %) died from breast cancer, 12 (2.8 %) died from other cancers, and 6 (1.4 %) died from other causes. Among patients whose cancers had been detected by CBE alone, 449 (86.0 %) were alive, 57 (10.9 %) died from breast cancer, 13 (2.5 %) died from other cancers, and 3 (0.6 %) died from other causes. Among the 3,047 patients whose cancers had been self-detected, 2,366 (77.7 %) were alive, 568 (18.6 %) died from breast cancer, 55 (1.8 %) died from other cancers, and 58 (1.9 %) died from other causes (Table 3).

The mean observation time for patients who had been screened using MMG with CBE was slightly shorter than that for patients who had been screened using CBE alone, or for patients with self-detected cancer. Five-year survival of breast cancer patients who had been screened by use of MMG with CBE, by CBE alone, and by self-detection was 96.8, 92.7, and 86.6 %, respectively. The corresponding 8-year survival was 94.9, 88.7, and 82.1 %, respectively (Table 4).

Table 4 Survival of the study subjects according to modality

Modality	Survival		Mean observation time (month)	SD
	5-year (%)	8-year (%)		
MMG with CBE	96.8	94.9	86.1	29.5
CBE	92.7	88.7	94.8	36.6
Self-detection	86.6	82.1	96.9	39.8
Total	88.5	84.3	95.5	38.5

MMG mammography, *CBE* clinical breast examination

Outcome and survival analysis according to detection method

Statistically significant differences in outcome and survival were observed between the patients screened by use of MMG with CBE and those screened by use of CBE alone ($p = 0.0008$), and those with self-detected cancers ($p < 0.0001$). The difference between the CBE screening and self-detection groups was significant ($p < 0.0001$) (Fig. 1).

Mortality risk analysis according to detection method

Mortality risk of breast cancer was determined by age-adjusted risk analysis. Mortality risk among patients screened by CBE alone was 2.59-fold (95 % CI 1.52–4.41, $p = 0.0005$) and that among patients with self-detected cancers was 4.37-fold (95 % CI 2.73–6.99, $p < 0.0001$) higher than that among patients screened by MMG with CBE. The subjects were stratified into two age groups (40–49 years and 50–69 years) for statistical analysis of

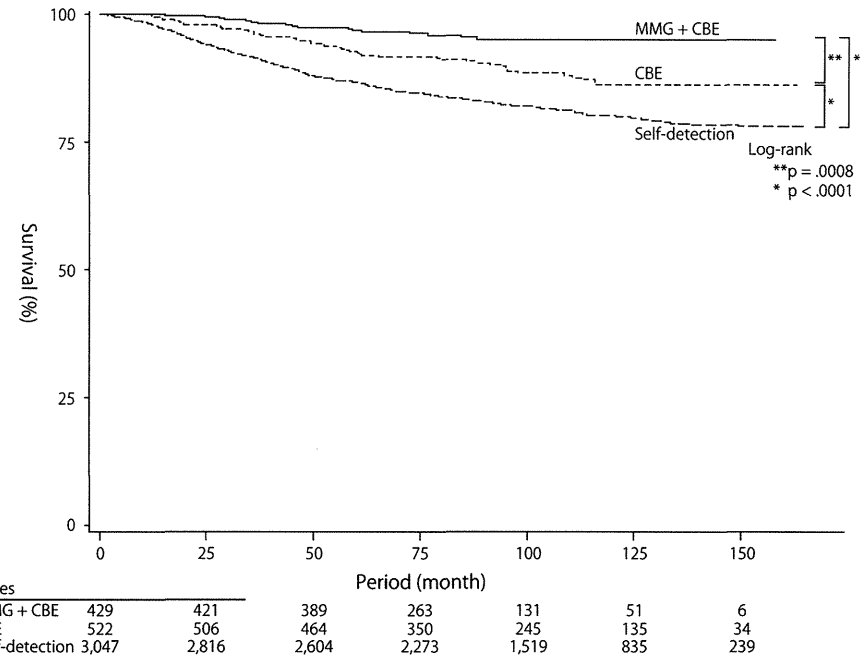

Fig. 1 Kaplan–Meier survival curves for screening mammography (MMG) with clinical breast examination (CBE) (429 patients), CBE alone (522 patients), and self-detection (3,047 patients). Statistically significant differences were observed between the MMG with CBE group and the self-detection group ($p < 0.0001$), the CBE alone and self-detection groups ($p < 0.0001$), and the MMG with CBE and CBE alone groups ($p = 0.0008$)

Table 5 Hazard ratio (HR) and 95 % confidence interval (CI) for the mortality risk of each group compared by screening mammography

Age group	Modality	Cases	Person-years	Breast cancer death	HR	95 % CI	*p*
All	MMG with CBE	429	3,077.1	18	1.00 (reference)[a]	–	–
	CBE	522	4,124.0	57	2.59	1.52–4.41	0.0005
	Self-detection	3,047	24,646.1	568	4.37	2.73–6.99	<.0001
40–49	MMG with CBE	78	686.6	3	1.00 (reference)	–	–
	CBE	273	2,145.5	23	2.38	0.72–7.94	0.16
	Self-detection	1,194	10,217.5	197	4.44	1.42–13.89	0.01
50–69	MMG with CBE	351	2,390.5	15	1.00 (reference)	–	–
	CBE	249	1,978.5	34	3.00	1.63–5.50	0.0004
	Self-detection	1,853	14,398.6	371	4.51	2.69–7.56	<.0001

HR hazard ratio, *CI* confidence interval, *MMG* mammography, *CBE* clinical breast examination

[a] Adjusted by age

mortality according to the detection method used. In the 40 to 49-year age group, mortality risk for CBE alone was 2.38-fold (95 % CI 0.72–7.94, $p = 0.16$) and that for self-detection was 4.44-fold (95 % CI 1.42–13.89, $p = 0.01$) higher than that for MMG with CBE. In the 50 to 69-year age group, however, the mortality risk for CBE alone was 3.00-fold (95 % CI 1.63–5.50, $p = 0.0004$) and that for self-detection was 4.51-fold (95 % CI 2.69–7.56, $p < .0001$) higher than that for MMG with CBE (Table 5).

Discussion

Several trials have been conducted to evaluate the effectiveness of MMG screening in western countries [2], on the basis of the relative risk of breast cancer death for women aged 40–69 years, and especially for those aged 50–69 years. Our previous study revealed that the survival of women aged over 50 years whose breast cancers had been detected by MMG was superior to that of women whose cancers had been detected by CBE alone or by self-examination; although the effectiveness of MMG for detecting breast cancer in women aged 40–49 years could not be evaluated by age-stratified analysis [8], there may have been some bias for screening-detected breast cancers, which may grow slowly and have a better prognosis [9]. In this retrospective cohort study of Japanese women aged 40–69 years, we evaluated whether the efficacy of screening using MMG with CBE was superior to that using CBE alone or to self-detection by investigating the

cumulative survival of women with screening-detected and interval cancer by reference to the Miyagi Prefectural Cancer Registry, one of the oldest and most reliable population-based cancer registries in Japan [11]. As in western countries, RCTs are required for evaluation of the effectiveness of MMG [15], but in Japan this is not realistic because many women have been included in the MMG program through studies to evaluate the efficacy of MMG [4–8] and because of endorsement by the Ministry of Health, Labour and Welfare in 2000. Therefore, this retrospective cohort study is one of the best recent attempts to clarify whether MMG is able to reduce breast cancer mortality in Japan.

Here we found that 5-year survival was over 90 % for women whose cancers had been detected by MMG with CBE, and by CBE alone. Survival of women who had been screened by use of MMG with CBE was significantly better than that of women who had been screened by use of CBE alone, or whose cancers had been self-detected. One possible reason for this better survival was the lower proportion of interval cancers in the group screened by use of MMG with CBE (18.6 %; 80/429) than in the group screened by use of CBE alone (40.0 %; 209/522). A previous study demonstrated that interval cancers tended to be more advanced, less well differentiated, and included a significantly higher proportion of triple-negative cancers, thus resulting a poorer outcome than for screening-detected and self-detected cancers [10]. In this study, the proportion of early, in situ, or localized, breast cancers was higher in the MMG with CBE group (77.6 %; 333/429) than in the CBE alone (64.6 %; 337/522) and self-detection (48.6 %; 1,481/3,047) groups. On the other hand, the proportion of advanced breast cancers, which are thought to be directly related to breast cancer death, was lower in women screened by use of MMG with CBE than in other groups. The proportion of breast cancer deaths was also lower in the MMG with CBE group (4.2 %) than in the CBE alone (10.9 %) and self-detection (18.6 %) groups. In the MMG with CBE group, age tended to be higher and mean observation time shorter than that in the CBE alone and self-detection groups; therefore, there was a possibility that the MMG with CBE group had a higher proportion of deaths from other causes. The proportion of deaths from other cancers and other causes in the MMG with CBE group (4.2 %; 18/429) was similar to that in the CBE alone (3.1 %; 16/522) and self-detection (3.7 %; 113/3,047) groups. Our observation period might have been sufficient because the three Kaplan–Meier curves were parallel at the end of follow-up; therefore, other causes of death might not have distorted the results. Ten-year survival might be almost the same as 8-year survival, because the three survival curves were almost horizontal after 96 months. The different clinical stages and pathological features of cancers in the MMG with CBE group might have resulted in the lower proportion of breast cancer deaths, which may be the presupposition of the declining mortality.

In the 40 to 49-year age group, the risk of breast cancer death among women screened by use of CBE alone was 2.38-fold higher than that among women screened by use of MMG with CBE, but this was not statistically significant (95 % CI 0.72–7.94, $p = 0.16$). The mortality risk in the self-detection group was significantly higher than that in the MMG with CBE group. In the 50 to 69-year age group, the mortality risk for women screened using CBE alone or self-detection was significantly higher than that in the MMG with CBE group. A meta-analysis of major MMG trials in western countries [2] found that MMG was effective in all age groups, but especially for woman aged 50 years and over. Our findings are consistent with those in that study, indicating that screening using MMG with CBE is more effective than self-detection for reduction of breast cancer mortality among women aged 40–69 years, and especially those aged 50–69 years.

The efficacy of MMG may be lower for women aged 40–49 years, for whom breast density is higher [7]. One approach for overcoming this weakness of MMG was evaluated in the "Japan Strategic Anti-cancer Randomized Trial (J-START)", which evaluated the effectiveness of MMG with ultrasound for breast cancer screening in comparison with mammography alone for women aged 40–49 years [16].

This study had both limitations and strengths. First, it was vulnerable to a variety of bias because of comparison of survival. Breast cancer screening presumably reduces mortality by detecting breast cancer and thus enabling a patient to be treated appropriately at an earlier stage. Differences in mortality risk between groups screened by use of MMG with CBE, CBE alone, and self-detection were presumably caused by the effect of MMG with CBE in reducing mortality and bias, for example self-selection bias (healthy screenee bias). However, the proportion of deaths from other cancers and other causes in the MMG with CBE group was almost the same as that in the CBE alone group and in the self-detection group, even though age tended to be higher in the MMG with CBE group. Therefore, any such bias might have been too small to distort the results. Second, we speculated whether lead time bias could cause these differences in survival [9]. However, the Kaplan–Meier survival curves for the three groups screened using these methods did not cross over, despite of our long observation period; it is, therefore, assumed that the effect of this bias on survival would have been too small to have affected the results.

One of the strengths of our study was that it was able to evaluate the effectiveness of MMG with CBE for women aged 40–49 years by analysis of mortality risk in different

age groups. Our previous study was unable to evaluate the effectiveness of MMG for women aged 40–49 years, for whom breast cancer incidence and mortality would be high in Japan [1], by analysis of mortality risk by age groups [8], because there were no deaths among patients whose cancers had been detected by MMG with CBE and the number of patients in this age group was small. Second, the quality of CBE and reading of MMG were controlled. The screening program was performed by registered surgeons who were approved by the committee for breast cancer screening of the Miyagi Cancer Society as having sufficient experience in general surgery, including the treatment of breast cancer. Statistically significant differences in survival were observed between self-detection and CBE alone, and between self-detection and MMG (Fig. 1). In comparison with self-detection, survival was significantly better for women whose cancers had been detected by CBE alone or by MMG with CBE. The difference in survival between the self-detection and CBE groups was larger than that between the CBE alone and MMG with CBE groups. This implies that the quality of CBE conducted by registered physicians was well controlled. It can be said that MMG with CBE is better than quality controlled CBE alone, although CBE is, of course, better than self-detection. Third, the Miyagi Prefectural Cancer Registry is one of the earliest and most accurate population-based cancer registries in Japan [11]. Therefore, the quality of the data is regarded as sufficiently reliable.

In conclusion, this analysis, conducted by reference to the population-based cancer registry in Miyagi, Japan and which included screening-detected and interval cancers, revealed that by screening using MMG with CBE it is possible to reduce breast cancer mortality from the perspective of survival and risk of breast cancer death among women aged 40–69 years in Japan. To reduce future breast cancer mortality in Japan, national screening by use of MMG with CBE should be increased. Currently, the prevalence of screening using MMG with CBE is 32.1 % (2005) in Miyagi Prefecture [17], as opposed to 60.8 % (2003) in the United States and 69.5 % (2005) in the United Kingdom [18]. This means there is still a higher proportion of self-detected cases in Japan. The only sure indicator of the effectiveness of MMG screening will be a decline in breast cancer mortality; before this can occur the problem of low screening must be addressed. Invitation to MMG screening for each eligible woman might effectively increase the amount of screening, as reported elsewhere [19].

Acknowledgments This work was supported by KAKENHI, including Grant-in-Aid for Scientific Research (C) (24590781), Grant-in-Aid for Young Scientists (A) (24689032), and the 3rd Term Comprehensive Control Research for Cancer (H23-Sanjigan-shitei-002) from the Ministry of Health, Labour and Welfare, Japan.

Conflict of interest The authors have no conflict of interest.

References

1. Matsuda T, Marugame T, Kamo K, Katanoda K, Ajiki W, Sobue T. Cancer incidence and incidence rates in Japan in 2006: based on data from 15 population-based cancer registries in the monitoring of cancer incidence in Japan (MCIJ) project. Jpn J Clin Oncol. 2012;42:139–47.
2. Nelson HD, Tyne K, Naik A, Bougatsos C, Chan BK, Humphrey L. Screening for breast cancer: an update for the US Preventive Services Task Force. Ann Intern Med. 2009;151:727–37. W237–42.
3. Kanemura S, Tsuji I, Ohuchi N, Takei H, Yokoe T, Koibuchi Y, et al. A case control study on the effectiveness of breast cancer screening by clinical breast examination in Japan. Jpn J Cancer Res. 1999;90:607–13.
4. Ohuchi N, Yoshida K, Kimura M, Ouchi A, Kamioki S, Shiiba K, et al. Improved detection rate of early breast cancer in mass screening combined with mammography. Jpn J Cancer Res. 1993;84:807–12.
5. Ohuchi N, Yoshida K, Kimura M, Ouchi A, Shiiba K, Ohnuki K, et al. Comparison of false negative rates among breast cancer screening modalities with or without mammography: Miyagi trial. Jpn J Cancer Res. 1995;86:501–6.
6. Ohnuki K, Kuriyama S, Shoji N, Nishino Y, Tsuji I, Ohuchi N. Cost-effectiveness analysis of screening modalities for breast cancer in Japan with special reference to women aged 40–49 years. Cancer Sci. 2006;97:1242–7.
7. 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.
8. Kawai M, Kuriyama S, Suzuki A, Nishino Y, Ishida T, Ohnuki K, et al. Effect of screening mammography on breast cancer survival in comparison to other detection methods: a retrospective cohort study. Cancer Sci. 2009;100:1479–84.
9. Raffle AE, Gray JAM. Screening: evidence and practice. Oxford: Oxford University Press; 2007.
10. Domingo L, Blanch J, Servitja S, Corominas JM, Murta-Nascimento C, Rueda A, et al. Aggressiveness features and outcomes of true interval cancers: comparison between screen-detected and symptom-detected cancers. Eur J Cancer Prev. 2012. doi: 10.1097/CEJ.0b013e328354d324.
11. Nishino Y, Tsuji I, Moro K. Cancer incidence in Miyagi, Japan (1998–2002). In: Curado MP, Edwards B, Shin HR, Storm H, Ferlay J, Heanue M, et al., editors. Cancer incidence in five continents, vol IX, IARC Scientific Publications No 160. Lyon: IARC; 2007.
12. World Health Organization. International Classification of Diseases for Oncology, 2nd edn (ICD-O-2). Lyon: World Health Organization; 1990.
13. Sobue T, Tsukuma H, Okamoto N, Ajiki W, editors. Handbook on population-based cancer registration in Japan fifth revision. Tokyo: Japanese Association of Cancer Registries; 2007 (in Japanese).
14. Cox DR. Regression models and life-tables (with discussion). J R Stat Soc. 1972;34:187–220.
15. Canadian Task Force on the Periodic Health Examination. The periodic health examination. Can Med Assoc J. 1979;121:1193–254.
16. Ohuchi N, Ishida T, Kawai M, Narikawa Y, Yamamoto S, Sobue T. Randomized controlled trial on effectiveness of ultrasonography

screening for breast cancer in women aged 40–49 (J-START): research design. Jpn J Clin Oncol. 2011;41:275–7.
17. The Ministry of Health, Labour and Welfare. Health statistics in Japan, 2007. Tokyo: Health Statistics Office, Statistics and Information Department, p. 35.
18. Armesto SG, Lapetra ML, Wei L, Kelley E, the Members of the HCQI Expert Group. Health Care Quality Indicators Project 2006 Data Collection Update Report. 29th ed. Paris: OECD; 2007. p. 35.
19. Ferroni E, Camilloni L, Jimenez B, Furnari G, Borgia P, Guasticchi G, et al. How to increase uptake in oncologic screening: a systematic review of studies comparing population-based screening programs and spontaneous access. Prev Med. 2012. doi:10.1016/j.ypmed.2012.10.007.

Breast Cancer (2014) 21:325–333
DOI 10.1007/s12282-012-0387-5

ORIGINAL ARTICLE

Analysis of clinically relevant values of Ki-67 labeling index in Japanese breast cancer patients

Kentaro Tamaki · Takanori Ishida · Nobumitsu Tamaki · Yoshihiko Kamada · Kanou Uehara · Minoru Miyashita · Masakazu Amari · Akiko Tadano-Sato · Yayoi Takahashi · Mika Watanabe · Keely McNamara · Noriaki Ohuchi · Hironobu Sasano

Received: 16 February 2012 / Accepted: 13 June 2012 / Published online: 11 July 2012

Abstract
Background It has become important to standardize the methods of Ki-67 evaluation in breast cancer patients, especially those used in the interpretation and scoring of immunoreactivity. Therefore, in this study, we examined the Ki-67 immunoreactivity of breast cancer surgical specimens processed and stained in the same manner in one single Japanese institution by counting nuclear immunoreactivity in the same fashion.
Methods We examined 408 Japanese breast cancers with invasive ductal carcinoma and studied the correlation between Ki-67 labeling index and ER/HER2 status and histological grade of breast cancer. We also analyzed overall survival (OS) and disease-free survival (DFS) of these patients according to individual Ki-67 labeling index.
Results There were statistically significant differences of Ki-67 labeling index between ER positive/HER2 negative and ER positive/HER2 positive, ER negative/HER2 positive or ER negative/HER2 negative, and ER positive/HER2 positive and ER negative/HER2 negative groups (all $P < 0.001$). There were also statistically significant differences of Ki-67 labeling index among each histological grade ($P < 0.001$, respectively). As for multivariate analyses, Ki-67 labeling index was strongly associated with OS (HR 39.12, $P = 0.031$) and DFS (HR 10.85, $P = 0.011$) in ER positive and HER2 negative breast cancer patients. In addition, a statistically significant difference was noted between classical luminal A group and "20 % luminal A" in DFS ($P = 0.039$) but not between classical luminal A group and "25 % luminal A" ($P = 0.105$).
Conclusions A significant positive correlation was detected between Ki-67 labeling index and ER/HER2 status and histological grades of the cases examined in our study. The suggested optimal cutoff point of Ki-67 labeling index is between 20 and 25 % in ER positive and HER2 negative breast cancer patients.

Keywords Ki-67 · Breast cancer · Cutoff point · Estrogen receptor · HER2 · Histological grade

K. Tamaki (✉) · N. Tamaki · Y. Kamada · K. Uehara
Department of Breast Surgery, Nahanishi Clinic, 2-1-9 Akamine, Naha, Okinawa 901-0154, Japan
e-mail: nahanisikenta@yahoo.co.jp; k-tamaki@naha-nishi-clinic.or.jp

K. Tamaki · T. Ishida · M. Miyashita · M. Amari · A. Tadano-Sato · N. Ohuchi
Department of Surgical Oncology, Tohoku University Graduate School of Medicine, Sendai, Japan

K. Tamaki · M. Miyashita · Y. Takahashi · M. Watanabe · K. McNamara · H. Sasano
Department of Pathology, Tohoku University Hospital, Sendai, Japan

Introduction

Tumor proliferation fraction has become an established predictive marker for clinical outcome of breast cancer patients [1–3]. Uncontrolled cell proliferation has also been considered a hallmark of malignancy and can be assessed by various laboratory methods, including counting mitotic figures under light microscopy, flow or image cytometric evaluation of the fraction of the cells in S phase, and immunohistochemistry of various nuclear antigens associated with cell proliferation [3–5]. The proliferation antigen Ki-67 is localized in nuclei of the cells at all phase of the cell cycle except for those at G0 phase and, in particular, the Ki-67 labeling index (percentage of cells with Ki-67

positive nuclear immunoreactivity) is considered to represent the status of tumor proliferation [1–3, 6, 7].

The statistically significant correlation between the Ki-67 labeling index of carcinoma cells and clinical outcome has been reported in human breast cancer patients [8–10]. Trihia et al. reported that a relatively higher Ki-67 labeling index within the carcinoma was significantly associated with adverse clinical outcome regardless of the subtypes of breast cancer [9, 10]. These results indicate that the Ki-67 labeling index in breast carcinoma cells may confer a higher risk of relapse and subsequently a worse overall survival in those with early breast cancer [8–10].

While results obtained using the Ki-67 labeling index of carcinoma cells resemble those obtained by the Oncotype Dx assay in ER positive and lymph node negative breast cancer patients (largely because the results of the Oncotype Dx assay are based on the status of cell proliferation genes) [11], additional information can be gained from assessing the Ki-67 labeling index within the carcinoma cells. The information obtained from such an assessment is not limited to predictions of prognosis or clinical outcome but also includes prediction of relative responsiveness or resistance to chemotherapy or endocrine therapy in adjuvant settings and the treatment efficacy in tissue specimens obtained before, during, and after neoadjuvant therapy, particularly neoadjuvant endocrine therapy [3]. Because of this additional predictive value, results of the Ki-67 labeling index in carcinoma cells have been incorporated into surgical pathology reports of breast cancer patients in an increasing number of diagnostic pathology laboratories in many countries [3].

However, as in any study utilizing immunohistochemical staining to evaluate clinical samples, it is cardinal and pivotal to standardize the method of Ki-67 measurement, including pre-analytical, analytical, interpretation, and scoring assessment [3], because otherwise results are far from reproducible and applicable in routine clinical settings. This may be particularly true of the methodology used in the stratification of early breast cancer patients into high and low proliferation groups. This stratification is markedly important in clinical settings and many attempts have been made to define the optimal cutoff value [12–14]; however, the reported value suggested to optimally distinguish these two groups of patients has been strikingly variable, from 1 to 28.6 %, thereby markedly limiting its clinical utility [3]. The 12th St. Gallen International Breast Cancer Conference 2011 recommended that patients with ER positive and HER2 negative breast cancer with a Ki-67 labeling index of 14 % or more may be recommended to receive adjuvant chemotherapy in addition to endocrine therapy [12]. The use of this cutoff point must, however, be approached with some caution as Nishimura et al. [13] recently demonstrated that the optimal cutoff of Ki-67 was 25 % in Japanese early breast cancer patients. In addition, the International Ki-67 in Breast Cancer Working Group also proposed that the direct application of specific cutoffs for decision making must be considered unreliable unless analyses were conducted in a highly experienced laboratory with its own reference data [3].

Careful and critical review of the previously reported studies of Ki-67 in human breast cancer revealed that the great majority of Ki-67 labeling index studies have not necessarily been performed under stringent conditions as described above, especially under those recommended by the International Ki-67 in Breast Cancer Working Group. Therefore, in this study, we evaluated the Ki-67 labeling index in breast cancer surgical pathology specimens processed in the same manner in a single institute, Tohoku University Hospital, Sendai, Japan and by the same observers using the same evaluation criteria. We then evaluated the correlation between the Ki-67 labeling index and ER/HER2 status and histological grade in Japanese cases of invasive ductal carcinoma. We then attempted to determine the clinical relevant cutoff value or the percentage of Ki-67 positive invasive breast carcinoma cells that could differentiate eventual clinical outcome of ER positive breast cancer cases.

Materials and methods

Carcinomas

We examined 408 Japanese patients with invasive ductal carcinomas of the breast, all of whom had undergone surgery at Tohoku University Hospital, Sendai and Nahanishi Clinic Okinawa. The study protocol was approved by the Ethics Committee at Tohoku University Graduate School of Medicine. The median age of the patients was 56 years (range 25–89 years). Estrogen receptor (ER) and human epidermal growth factor receptor 2 (HER2) status were reevaluated and summarized as follows: ER positive and HER2 negative, ER positive and HER2 positive, ER negative and HER2 positive, and ER negative and HER2 negative. These specimens had been first cut into 5-mm slices after carefully inking the margins, fixed in 10 % formalin for 46–48 h at room temperature, and embedded in paraffin wax.

Immunohistochemistry

Immunohistochemical analyses were all performed by a single experienced histotechnician at the Department of Pathology, Tohoku University Hospital using the same protocol. All the blocks were freshly cut into 4-μm sections, placed on glue-coated glass slides (Matsunami Glass

Ind., Ltd, Osaka, Japan), and left at room temperature for 3–5 days. Sections were then deparaffinized in xylene, and hydrated with graded alcohols and distilled water at room temperature. Endogenous peroxidase activity was blocked with freshly prepared 3 % hydrogen peroxidase for 10 min at room temperature. Antigen retrieval was performed in an autoclave (Tomy SX-500 high pressure steam sterilizer, Tomy Seiko Co., Ltd., Tokyo, Japan) using citrate buffer for Ki-67 heated at 121 °C for 5 min. Sections were subsequently incubated for 30 min at room temperature in a blocking solution of 10 % rabbit serum (Nichirei Biosciences, Tokyo, Japan) for Ki-67, and then immunostained for 16 h at 4 °C with the primary antibody. The primary antibody of Ki-67 was MIB-1 mouse monoclonal antibody (code M7240; Dako, Copenhagen, Denmark) diluted at 1:300. Secondary antibody reaction for Ki-67 immunohistochemistry was performed using biotinylated rabbit anti-mouse antibody (Nichirei Bioscience) at a dilution of 1:100 for 30 min at room temperature and peroxidase-conjugated avidin (Nichirei Bioscience) was used according to the manufacture's instruction. Reacted sections were visualized using 3,3′-diaminobenzidine-tetrachloride (DAB)/30 % H_2O_2 in 0.05 mol/l Tris buffer (pH 7.6) and counterstained with hematoxylin for nuclear staining. We used the avidin–streptavidin immunoperoxidase method using the clone 6F11 antibody (Ventana, Tucson, AZ, USA) in an automated immunostainer (Benchmark System; Ventana) for immunohistochemistry of ER. A standardized immunohistochemistry kit (Hercep-Test for Immunoenzymatic Staining; Dako) was used for HER2 staining as previously reported [15, 16].

Histopathological analysis

Histopathological evaluations were based on the World Health Organization (WHO) histological classification of tumors of breast and *Rosen's Breast Pathology* [17, 18]. Histological grades were assessed according to the criteria of Elston and Ellis [17, 18]. The Ki-67 immunoreactivity was evaluated independently by two of the authors by first identifying the areas of the most densely stained areas in the whole tissue sections by scanning at low power fields and then counting 1000 carcinoma cells in these areas [3]. We used an Olympus BX50 (Olympus, Tokyo, Japan) and ×20 objectives for the analysis. Figure 1 represents characteristic immunohistochemical findings of Ki-67 positive and negative carcinoma cells (Fig. 1). The presence of ER was determined by distinctive nuclear immunoreactivity and was graded from 0 to 8 using the Allred score, with positivity of the cases defined as a score of 3 [19]. With regard to HER2 evaluation, membranous staining was graded as 0–1+, 2+, and 3+ [20]. The cases scored as 2+ were subjected to FISH to calculate the gene copy ratio of

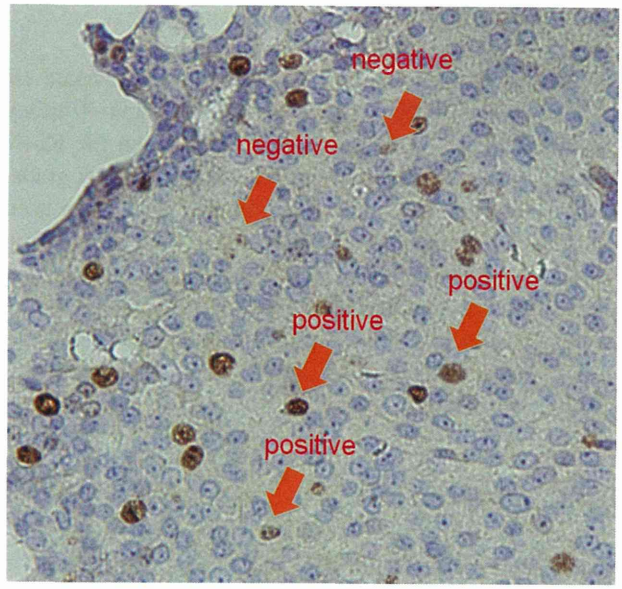

Fig. 1 Representative immunohistochemical findings of Ki-67 positive and negative carcinomas. The specimens were fixed in neutral buffered 10 % formalin and sections stained for Ki-67 with MIB1 antibody (*brown stain*) and counterstained with Mayer's hematoxylin (*blue stain*) (color figure online)

HER2 to CEP17 (PathVysion HER2 DNA Probe kit; Abbott, Chicago, IL, USA), as previously reported [15, 21]. HER2 positive cases were defined as a HER2/CEP17 signal ratio (FISH score) greater than 2.2 [20].

On the basis of the values obtained in the manner above, we examined the correlation between the Ki-67 labeling index and ER/HER2 status and histological grade. We also analyzed overall survival (OS) and disease-free survival (DFS) stratified according to the Ki-67 labeling index, in order to examine the utility of various cutoff points of Ki-67 in predicting clinical outcome within various ER+ breast cancer subgroups (luminal A, luminal B). In order to do this we tentatively assigned luminal A cases as follows: "classical luminal A" as the ER positive and HER2 negative group [22]; "14 % luminal A", based upon the proposal made at the St. Gallen 2011 consensus meeting [12], with a Ki-67 labeling index of less than 14 %; "20 % cutoff luminal A" with a Ki-67 labeling index of less than 20 %; "25 % cutoff luminal A" with a Ki-67 labeling index of less than 25 %; and "30 % cutoff luminal A" with a Ki-67 labeling index of less than 30 % [14, 23]. As for luminal B, we defined "classical luminal B" as ER positive and HER2 positive [24]; "14 % luminal B", proposed at St. Gallen 2011 [12], with a Ki-67 labeling index of more than 14 %; "20 % cutoff luminal B" with a Ki-67 labeling index of more than 20 %; "25 % cutoff luminal B" with a Ki-67 labeling index of more than 25 %; and "30 % cutoff luminal B" with a Ki-67 labeling index of more than 30 % [14, 22].

Statistical analyses

Statistically analyses were performed using StatMate IV for Windows (ATMS, Tokyo, Japan). The Mann–Whitney test was used to assess the correlation between the Ki-67 labeling index and ER/HER2 status and histological grade. The Cox proportional hazards regression model was used for multivariate analyses to evaluate each factor including the Ki-67 labeling index, TNM stages, ER expression, HER2 status, and adjuvant therapy of the patients. The analyses of OS or DFS curves were performed using the Kaplan–Meier method. The results were considered significant at $P < 0.05$.

Results

Correlation between Ki-67 labeling index and ER and HER2 status

Figure 2 summarizes the Ki-67 labeling index results according to ER and HER2 status of the cases examined. The Ki-67 labeling index in carcinoma cells was 11 % (median) and 17.9 % (average) in ER positive/HER2 negative, 40 % (median) and 36.4 % (average) in ER positive/HER2 positive, 40 % (median) and 46.8 % (average) in ER negative/HER2 positive, and 60 % (median) and 56.3 % (average) in ER negative/HER2 negative groups. There were statistically significant differences of the Ki-67 labeling index between ER positive/HER2 negative and ER positive/HER2 positive, ER negative/HER2 positive or ER negative/HER2 negative, and ER positive/HER2 positive and ER negative/HER2 negative groups (all $P < 0.001$).

Correlation between Ki-67 labeling index and histological grades

Figure 3 summarizes the Ki-67 labeling results index in each histological grade of the cases examined. The Ki-67 labeling index was 6 % (median) and 8.5 % (average) in grade 1, 19 % (median) and 24.0 % (average) in grade 2, and 60 % (median) and 55.8 % (average) in grade 3. The Ki-67 labeling index was significantly different between histological grades ($P < 0.001$, respectively).

OS of luminal A and B groups according to Ki-67 labeling index

Table 1 shows the distribution of patients according to the subtypes classical luminal, 14 % luminal, 20 % luminal, 25 % luminal, and 30 % luminal. The 5-year OS rates of patients in luminal A groups were 0.949 in classical luminal A, 1.000 in "14 % luminal A", 1.000 in "20 % luminal A", 1.000 in "25 % luminal A", and 1.000 in "30 % luminal A". There were no statistically significant differences of OS rates among these groups. The 5-year OS rates of luminal B were 1.000 in classical luminal B, 0.875 in "14 % luminal B", 0.853 in "20 % luminal B", 0.822 in "25 % luminal B", and 0.812 in "30 % luminal B". No statistically significant differences were detected among these groups.

DFS of luminal A and B groups according to the Ki-67 labeling index

Figure 4 summarizes the DFS rates of the patients according to each subgroup determined by the Ki-67 labeling index of individual cases. The 5-year DFS rates of patients in luminal A groups were 0.956 in classical luminal A, 1.000 in "14 % luminal A", 0.993 in "20 % luminal A", 0.989 in "25 % luminal A", and 0.983 in "30 % luminal A". There were statistically significant differences between classical luminal A and "14 % luminal A" or "20 % luminal A" ($P = 0.010$ and $P = 0.039$, respectively). A similar tendency was also noted between classical luminal A and "25 % luminal A" or "30 % luminal A" ($P = 0.105$ and 0.159, respectively) but the difference did not reach statistical significance. The 5-year DFS rates of patients in luminal B groups were 0.885 in classical luminal B, 0.880 in "14 % luminal B", 0.871 in "20 % luminal B", 0.840 in "25 % luminal B" and 0.835 in "30 % luminal B". There were no statistically significant differences among these groups above.

Multivariate analyses of OS and DFS according to Ki-67 labeling index

Among the factors examined, including the Ki-67 labeling index, tumor size, nodal status, stage, and adjuvant chemotherapy status, the Ki-67 labeling index was markedly associated with OS (HR 39.12, $P = 0.031$) and DFS (HR 10.85, $P = 0.011$) in ER positive and HER2 negative breast cancer patients. However, the Ki-67 labeling index was not statistically associated with OS (HR 9.28, $P = 0.198$) and DFS (HR 5.76, $P = 0.420$) in all cases including ER positive/HER2 positive, ER negative/HER2 negative, and ER negative/HER2 positive breast cancer patients.

Determination of Ki-67 labeling index cutoff values of carcinoma cells according to the clinical outcome of ER positive breast cancer cases

We evaluated the statistical significance of cutoff values of the Ki-67 labeling index in carcinoma cells segregated by 5 %. There were no statistically significant differences in OS of the patients. A statistically significant difference was

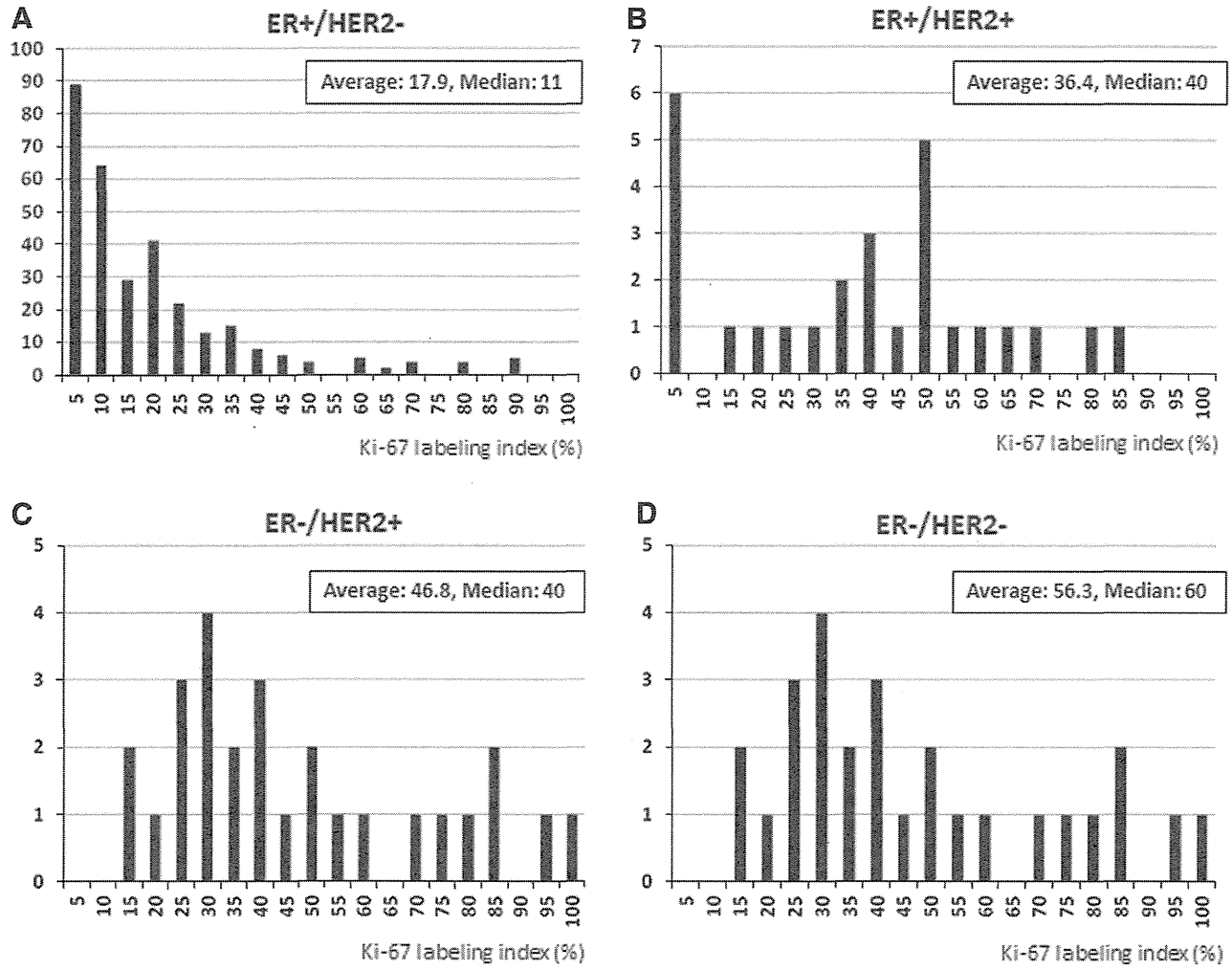

Fig. 2 Correlation between Ki-67 labeling index and ER or HER2 status. The distribution of Ki-67 labeling index in **a** ER positive and HER2 negative cases, **b** ER positive and HER2 positive cases, **c** ER negative and HER2 positive cases, **d** ER negative and HER2 negative cases

noted between classical luminal A group and "20 % luminal A" in DFS ($P = 0.039$) but not between classical luminal A group and "25 % luminal A" ($P = 0.105$). Therefore, the optimal cutoff point of the Ki-67 labeling index was suggested to be between 20 and 25 %.

Discussion

Ki-67 has been established as a well-known biomarker of cell proliferation in many human malignancies including breast cancer. The Ki-67 labeling index has been utilized to obtain both prognosis and prediction of the sensitivity to systemic therapy of breast cancer patients [2, 10, 21]. Some examples of this are the statistically significant correlation between a high Ki-67 labeling index of carcinoma cells and increased risk of cancer relapse and death in breast cancer patients [10] and the utility of mid-course evaluation of Ki-67 labeling index, even after 2 weeks of endocrine therapy, in predicting the subsequent response to endocrine therapy in ER positive breast cancer patients [23]. In addition the group of breast cancer patients associated with a high Ki-67 labeling index studied in the Breast International Group trial (BIG) 1-98 was associated with a potential clinical benefit in selecting letrozole over tamoxifen in post-menopausal patients [2]. Despite these important aspects of Ki-67 immunohistochemistry, the necessary standardized guidelines have not been developed [12, 25].

The International Ki-67 in Breast Cancer Working Group recently recommended the fixation of the specimens with neutral buffered formalin for 4–48 h or more and the counting of at least 500 invasive carcinoma cells using MIB-1 mouse monoclonal antibody [3]. In our present study, all the specimens examined had been processed in the same manner and according to the guidelines above and the Ki-67 labeling index was also evaluated accordingly.

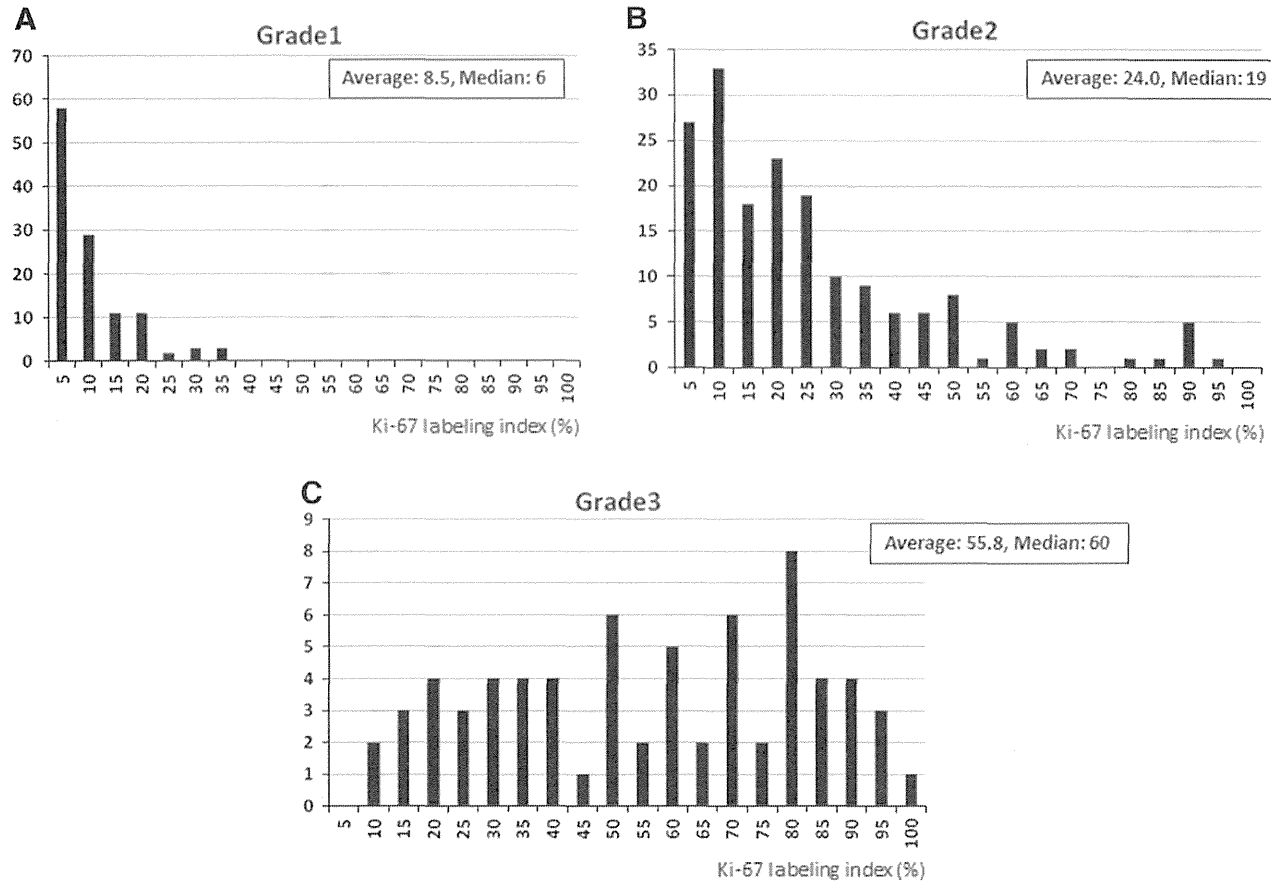

Fig. 3 Correlation between Ki-67 labeling index and histological grade of the patients. The distribution of Ki-67 labeling index in **a** grade 1, **b** grade 2, **c** grade 3 groups

Table 1 Distribution of patients according to the subtypes classical luminal, 14 % luminal, 20 % luminal, 25 % luminal, and 30 % luminal

	n	Ki-67 (median %)	Ki-67 (average %)
Classical lum A	289	11	17.9
14 % lum A	160	5	6.0
20 % lum A	186	6	7.5
25 % lum A	215	8	9.2
30 % lum A	225	9	10.1
Classical lum B	23	40	36.4
14 % lum B	152	27	33.2
20 % lum B	126	31	36.7
25 % lum B	97	35	41.1
30 % lum B	87	40	42.9

Previous studies conducted by Nishimura et al. [26–28] on Japanese breast cancer patients demonstrated that the Ki-67 value as significantly higher in triple negative cases. However, the Ki-67 labeling index was also statistically lower in ER positive/HER2 negative cases [26–28]. We therefore examined the correlation between the Ki-67 labeling index and hormone receptor, HER2 status, or histological grade using surgical pathology specimens processed in the same manner and immunostained in the same fashion by one single experienced histotechnician in one single institution.

The results of our present study demonstrated that the ER positive and HER2 negative group was associated with a significantly lower Ki-67 labeling index of carcinoma cells than in other subtypes examined. The cases with a high Ki-67 labeling index in the ER positive and HER2 negative group have been considered as potential candidates for receiving chemotherapy in addition to endocrine therapy as in the patients with a high histological grade [12–14]. In our present study, there was also a statistically significant correlation between the Ki-67 labeling index and histological grades of individual cases. Collectively our findings suggest that it may be better to review the slides when there is a significant discrepancy between the results of Ki-67 labeling index and histological grade in invasive ductal carcinoma cases. The results of our present study also demonstrated that subtyping of the tumors using immunohistochemical surrogate markers such as ER,

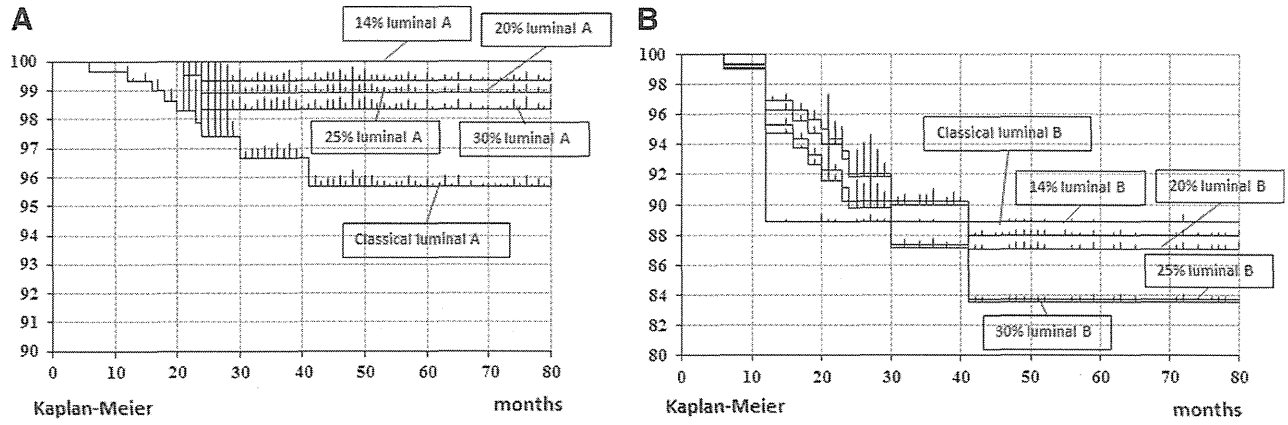

Fig. 4 DFS according to Ki-67 labeling index of the patients. **a** Luminal A: *classical luminal A* ER positive and HER2 negative; *14 % luminal A* Ki-67 labeling index less than 14 %; *20 % luminal A* Ki-67 labeling index less than 20 %; *25 % luminal A* Ki-67 labeling index less than 25 %; *30 % luminal A* Ki-67 labeling index less than 30 %. **b** Luminal B: *classical luminal B* ER positive and HER2 positive; *14 % luminal B* Ki-67 labeling index more than 14 %; *20 % luminal B* Ki-67 labeling index more than 20 %; *25 % luminal B* Ki-67 labeling index more than 25 %; *30 % luminal B* Ki-67 labeling index more than 30 %

HER2, and Ki-67, if using appropriately processed surgical pathology specimens and well-controlled immunohistochemical procedures, could at least contribute to identifying high-risk Japanese breast cancer patients within the hormone receptor positive subgroup of breast cancers. Nishimura et al. [26] also indicated that ER/PgR, HER2, and Ki-67 are all important biological markers for predicting prognosis and making effective treatment decisions in Japanese breast cancer patients by using only these biomarkers. The combination of these markers has been proposed at least in defining luminal A and B types of breast cancer without necessarily performing gene profiling studies with some exceptions [12, 29]. Luminal B type breast cancer represents a clinically important subgroup generally associated with adverse clinical outcome regardless of systemic adjuvant therapy [19]. It was recently recommended at the St. Gallens consensus meeting that chemotherapy was indicated for the majority of these patient defined as ER positive and with a Ki-67 labeling index of more than 14 % [12]. However, it is also true that the optimal cutoff points of the Ki-67 labeling index in these cases have been reported as 10–25 % [3, 12]. For instance, no pathological responders were reported in the cases with more than 25 % Ki-67 in neoadjuvant chemotherapy of Japanese breast cancer patients [13]. These discrepancies or variations of proposed values of Ki-67 labeling may be all due to differences of methodologies involved in obtaining the Ki-67 labeling index including pre-analytical factors such as fixation of the specimens and/or ethnical or racial backgrounds of the patients and further investigations are required for clarification.

The direct application of a specific cutoff for clinical decision making may be considered unreliable unless analyses are conducted in a highly experienced laboratory with its own reference data [3]. The International Ki-67 in Breast Cancer Working Group demonstrated that no consensus has been reached regarding the ideal cutoff point of the Ki-67 labeling index. The results of our present study demonstrated that there were statistically significant differences of DFS between classical luminal A and luminal with a 14 or 20 % cutoff of Ki-67. In addition, we examined the cutoff values of the Ki-67 labeling index segregated by 5 %. A statistically significant difference was noted between classical luminal A group and "20 % luminal A" in DFS but not between classical luminal A group and "25 % luminal A". Therefore, we propose an optimal cutoff point of the Ki-67 labeling index of between 20 and 25 %. These results were similar to that of a previous study from Japan mentioned above [13]. Therefore, ER positive and HER2 negative Japanese breast cancer patients with a Ki-67 labeling index of 20–25 % are associated with more aggressive biological course than those not and additional chemotherapy may be of further help or benefit to these patients.

It was recently proposed that the prognostic information provided by ER, PgR, HER2, and Ki-67 immunostaining performed in a rigorously controlled fashion was considered at least equivalent to that provided by 21 gene signature analysis and highlights the relevance of these readily available routine histopathological parameters in the clinical management of early ER positive breast cancer [30]. In addition, we demonstrated using multivariate analysis that the Ki-67 labeling index was one of the most important prognostic factors for the ER positive and HER2 negative group in this study. Therefore, it has become important to standardize the type of fixation, time to fixation, appropriate primary antibody, and methods of immunostaining and interpretation, especially in countries like Japan where

the expensive gene signature tests are and will be out of reach for the great majority of breast cancer patients. We also noted the statistically significant correlation between the Ki-67 labeling index and ER/HER2 status and histological grade of individual patients performed in a single institution. It is true that our present study was retrospective, the number of the patients is relatively small, and the patients were all Japanese but the results still provided sufficient evidence to support the value of the Ki-67 labeling index in the clinical management of breast cancer patients. Further investigations employing larger numbers of patients with longer periods of clinical follow-up may be required for determining the most clinically relevant cutoff points of the Ki-67 labeling index in breast cancer patients, especially those in the early stage in order to confer the maximal clinical benefits upon individual breast cancer patients.

Acknowledgments The authors appreciate the continuous excellent technical assistance of the staff in the Department of Pathology, Tohoku University Hospital, Sendai, Japan, especially their uninterrupted laboratory service toward the care of breast cancer patients despite enormous and unprecedented damages inflicted upon glass slides and instruments such as tissue processors, cryostat instruments, and automatic stainers, and harsh working conditions such as continuous aftershocks, total blackout, and interruption of running water in our laboratory as a result of the 3/11 earthquake. This work was supported in part by a Grant-in Aid from the "Kurokawa Cancer Research Foundation".

Conflict of interest The authors have no conflict of interest.

References

1. Clahsen PC, van de Velde CJ, Duval C, Pallud C, Mandard AM, Delobelle-Deroide A, et al. The unit of mitotic index, oestrogen receptor and Ki-67 measurements in the creation of novel prognostic indices for node-negative breast cancer. Eur J Surg Oncol. 1999;25:356–63.
2. Viale G, Giobbie-Hurder A, Regan MM, Coates AS, Mastropasqua MG, Dell'Orto P, et al. Prognostic and predictive value of centrally reviewed Ki-67 labeling index in postmenopausal women with endocrine-responsive breast cancer: results from Breast International Group Trial 1–98 comparing adjuvant tamoxifen with letrozole. J Clin Oncol. 2008;26:5569–75.
3. Dowsett M, Nielsen TO, A'Hern R, Bartlett J, Coombes RC, Cuzick J, et al. Assessment of Ki67 in breast cancer: recommendations from the International Ki67 in Breast Cancer Working Group. J Natl Cancer Inst. 2011;103:1–9.
4. Dressler LG, Seamer L, Owens MA, Clark GM, McGuire WL. Evaluation of a modeling system for S-phase estimation in breast cancer by flow cytometry. Cancer Res. 1987;47:5294–302.
5. Tovey SM, Witton CJ, Bartlett JM, Stanton PD, Reeves JR, Cooke TG. Outcome and human epidermal growth factor receptor (HER) 1–4 status in invasive breast carcinomas with proliferation indices evaluated by bromodoxyuridine labeling. Breast Cancer Res. 2004;6:246–51.
6. Gerdes J, Schwab U, Lemke H, Stein H. Production of a mouse monoclonal antibody reactive with a human nuclear antigen associated with cell proliferation. Int J Cancer. 1983;3:13–20.
7. Lehr HA, Hansen DA, Kussick S, Li M, Hwang H, Krummenauer F, et al. Assessment of proliferative activity in breast cancer: MIB-1 immunohistochemistry versus mitotic figure count. Hum Pathol. 1999;30:1314–20.
8. Tamaki K, Moriya T, Sato Y, Ishida T, Maruo Y, Yoshinaga K, et al. Vasohibn-1 in human breast carcinoma: a potential negative feedback regulator of angiogenesis. Cancer Sci. 2009;100:88–94.
9. Trihia H, Murray S, Price K, Gelber RD, Golouh R, Goldhirsch A, et al. Ki-67 expression in breast carcinoma: its association with grading system, clinical parameters, and other prognostic factors-a surrogate marker? Cancer. 2003;97:1321–31.
10. de Azambuja E, Cardoso F, de Castro G, Colozza M Jr, Mano MS, Durbecq V, et al. Ki-67 as prognostic marker in early breast cancer: a meta-analysis of published studies involving 12,155 patients. Br J Cancer. 2007;96:1504–13.
11. Paik S, Shak S, Tang G, Kim C, Baker J, Cronin M, et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. N Engl J Med. 2004;351:2817–26.
12. Goldhirsch A, Wood WC, Coates AS, Gelber RD, Thurlimann B. Panel members. Strategies for subtypes-dealing with the diversity of breast cancer: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2011. Ann Oncol. 2011;22:1736–47.
13. Nishimura R, Osako T, Okumura Y, Hayashi M, Arima N. Clinical significance of Ki-67 in neoadjuvant chemotherapy for primary breast cancer as a predictor for chemosensitivity and for prognosis. Breast Cancer. 2010;17:269–75.
14. Goldhirsch A, Ingle JN, Gelber RD, Coates AS, Thürlimann B. Panel members. Thresholds for therapies: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2009. Ann Oncol. 2009;20:1319–29.
15. Tamaki K, Ishida T, Miyashita M, Amari M, Mori N, Ohuchi N, et al. Multidetector row helical computed tomography for invasive ductal carcinoma of breast: the correlation between radiological findings and the corresponding biological characteristic of the patients. Cancer Sci. 2012;103:67–72.
16. Tamaki K, Sasano H, Ishida T, Ishida K, Miyashita M, Takeda M, et al. Comparison of core needle biopsy (CNB) and surgical specimens for accurate preoperative evaluation of ER, PgR and HER2 status of breast cancer patients. Cancer Sci. 2010;101:2074–9.
17. Tavassoli FA, Devilee P. World Health Organization classification of tumors. Tumor of the breast and female gential organs. Lyon: IARC; 2003.
18. Rosen PP. Rosen's breast pathology. 3rd ed. Philadelphia: Lippincott Williams & Wilkins; 2009.
19. Allred DC, Harvey JM, Berardo M, Clark GM. Prognostic and predictive factors in breast cancer by immunohistochemical analysis. Mod Pathol. 1998;11:155–68.
20. Wolff AC, Hammond ME, Schwartz JN, Hagerty KL, Allred DC, Cote RJ, et al. American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer. J Clin Oncol. 2007;25:118–45.
21. Miyashita M, Ishida T, Ishida K, Tamaki K, Amari M, Watanabe M, et al. Histopathological subclassification of triple negative breast cancer using prognostic scoring system: five variables as candidates. Virchows Arch. 2011;458:65–72.
22. Jalava P, Kuopio T, Juntti-Patinen L, Kotkansalo T, Kronqvist P, Collan Y. Ki67 immunohistochemistry: a valuable marker in prognostication but with a risk of misclassification: proliferation subgroups formed based on Ki67 immunoreactivity and standardized mitotic index. Histopathology. 2006;48:674–82.
23. Dowsett M, Smith IE, Ebbs SR, Dixon JM, Skene A, A'Hern R, Salter J. Prognostic value of Ki67 expression after short term presurgical endocrine therapy for primary breast cancer. J Natl Cancer Inst. 2007;99:167–70.

24. Spitale A, Mazzola P, Soldini D, Mazzucchelli L, Bordoni A. Breast cancer classification according to immunohistochemical markers: clinicopathologic features and short-term survival analysis in a population-based study from the South of Switzerland. Ann Oncol. 2009;20:628–35.
25. Cheang MCU, Chia SK, Voduc D, Gao D, Leung S, Snider J. Ki67 index, Her2 status, and prognosis of patients with luminal B breast cancer. J Natl Cancer Inst. 2009;101:736–50.
26. Nishimura R, Osako T, Okumura Y, Tashima R, Toyozumi Y, Arima N. Changes in the ER, PgR, HER2, p53 and Ki-67 biological markers between primary and recurrent breast cancer: discordance rates and prognosis. World J Surg Oncol. 2011;9:131.
27. Nishimura R, Okumura Y, Arima N. Trastuzumab monotherapy versus combination therapy for treating recurrence breast cancer: time to progression and survival. Breast Cancer. 2008;15:57–64.
28. Nishimura R, Arima N. Is triple negative a prognostic factor in breast cancer? Breast Cancer. 2008;15:303–8.
29. Bhargava R, Striebel J, Beriwal S, Flickinger JC, Onisko A, Ahrendt G. Prevalence, morphologic features and proliferation indices of breast carcinoma molecular classes using immunohistochemical surrogate markers. Int J Clin Exp Pathol. 2009;2:444–55.
30. Cuzick J, Dowsett M, Wale C. Prognostic value of a combined ER, PgR, Ki67, HER2 immunohistochemical (IHC4) score and comparison with the GHI recurrence score-results from TransATAC. Cancer Res. 2009;69:503s.

Breast Cancer (2014) 21:532–541
DOI 10.1007/s12282-012-0423-5

ORIGINAL ARTICLE

Cost-effectiveness analysis for breast cancer screening: double reading versus single + CAD reading

Miho Sato · Masaaki Kawai · Yoshikazu Nishino · Daisuke Shibuya · Noriaki Ohuchi · Tadashi Ishibashi

Received: 3 August 2012 / Accepted: 15 October 2012 / Published online: 27 October 2012

Abstract
Background Computer-aided detection (CAD) increases breast cancer detection, but its cost-effectiveness is unknown for breast cancer screening in Japan. We aimed to determine whether screening mammography diagnosed by one physician using CAD is cost-effective when compared with the standard double reading by two physicians.
Methods We established our model with a decision tree and Markov model concept based on feasible screening and clinical pathways, combined with prognosis of the health state transition of breast cancer. Cost-effectiveness analysis between double reading by two readers and single reading with CAD by one reader was performed from a social perspective in terms of the expected cost, life expectancy and incremental cost-effectiveness ratio (ICER). The hypothetical population comprised 50-year-old female breast cancer screening examinees. Only direct medical costs related to breast cancer screening and treatment were considered. One simulation cycle was 2 years, and the annual discount rate was 3 %. Sensitivity analysis was performed to evaluate the robustness of the model and input data.
Results Single reading with CAD increased expected costs by 2,704 yen and extended life expectancy by 0.0087 years compared with double reading. The ICER was 310,805 yen per life year gained, which is below the threshold. Sensitivity analysis showed that the sensitivity and specificity of CAD and the number of breast cancer screening examinees greatly affected the results.
Conclusions Single reading using CAD in mammography screening is more cost-effective than double reading, although the results are highly sensitive to the sensitivity and specificity of CAD and the numbers of examinees.

Keywords Computer-aided diagnosis · Cost effectiveness · Breast cancer · Decision trees · Diagnostic techniques and procedures

M. Sato · T. Ishibashi (✉)
Division of Clinical Imaging, Tohoku University Graduate School of Medicine, 2-1 Seiryo-machi, Aoba-ku, Sendai, Miyagi 980-8575, Japan
e-mail: tisibasi@mail.tains.tohoku.ac.jp

M. Kawai · N. Ohuchi
Department of Surgical Oncology, Tohoku University Graduate School of Medicine, 1-1 Seiryo-machi, Aoba-ku, Sendai, Miyagi 980-8574, Japan

M. Kawai
Division of Community Health, Tohoku University Graduate School of Medicine, 2-1 Seiryo-machi, Aoba-ku, Sendai, Miyagi 980-8575, Japan

Y. Nishino
Miyagi Cancer Center Research Institute, 47-1 Aza-Nodayama, Medeshimasiote, Natori, Miyagi 981-1293, Japan

D. Shibuya
Miyagi Cancer Society, 5-7-30 Kamisugi, Aoba-ku, Sendai, Miyagi 980-0011, Japan

Introduction

Breast cancer screening is a beneficial tool not only from the perspective of preventive medicine, but also for health economics [1–4]. Biennial screening mammography with double reading and palpation is recommended for women aged over 40 in Japan, aiming at effective screening with precise diagnosis of breast cancer. However, some communities have difficulties recruiting two individual doctors certified by the Central Committee on

Quality of Mammographic Screening for their double reading needs. This issue is especially important in areas where onsite consensus reading is done by two certified doctors with extensive experience. Moreover, the Japanese breast cancer screening rate is still lower than in other developed countries, and the Ministry of Health, Labor and Welfare Japan is trying to raise this to 50 % (from 30.6 % in 2010) [5]. This reader shortage problem for double reading will become serious in the near future whether we use independent or consensus double reading. To improve the screening rate and efficiency given the limited numbers of readers, we expect that computer-aided detection (CAD) could be a solution as a second reader.

CAD uses an algorithm to indicate suspicious regions in medical images, allowing abnormal areas to be detected objectively. However, its effectiveness in breast cancer screening has been controversial because it detects calcifications at a higher rate, but at a lower rate for focal asymmetric density (FAD) or architectural distortion [6–11]. The accuracy of CAD assessment on mammography has been studied and discussed in Japan [6, 7, 9, 12]. However, a socioeconomic analysis has yet to be performed because of fewer practical examples of CAD installation for mammography screening, and its health-economic impact remains unclear. CAD cannot be sold as software in Japan, but rather as a medical device because of the Japanese Pharmaceutical Affairs Act, and the cost usually includes at least the cost of CAD itself, high-resolution monitors, storage servers and maintenance fees. It is an expensive investment without cost-effectiveness evidence, even though CAD could possibly solve the reader shortage and save the second reading fee.

This situation has left CAD marketing in Japan 10 years behind other developed countries, and a cost-effectiveness analysis has now become necessary to facilitate effective CAD usage in Japan [13]. Without actual clinical research examples, long-term estimation in a hypothetical setting gives an indication of decision making, and a modeling analysis with decision trees and a Markov model could suggest whether future investment is effective [3, 14]. Decision trees are often used to analyze alternatives in outcomes research and are also a useful tool in clinical decision-making, while the Markov model evolves a hypothetical cohort over a number of time cycles [15–17]. In particular, disease carries a long-term prognosis, and repetitive screening such as for breast cancer is appropriate for Markov models.

The aim of this study is to assess whether one physician reading using CAD is cost-effective compared with the standard double reading in mammography breast cancer screening by a modeling method.

Materials and methods

Modeling method

We established a breast cancer screening model using a decision tree and a Markov model to perform decision analysis with long-term cost-effectiveness analysis (Figs. 1, 2). The TreeAge Pro 2009 Healthcare module (TreeAge Software Inc., Williamstown, MA, USA) was used for model construction and simulation. The structure of our decision tree started with two mammography reading methods: onsite consensus double reading by two physicians certified in mammography reading (double reading), and single reading by one certified physician using CAD (single + CAD). The structure of both methods was the same because the clinical pathways and health condition changes in breast cancer were considered similarly in both, and the decision nodes follow a Markov model.

We defined six health states: women with non-breast cancer (BC) history, early stage by screening (SC) follow-up, advanced stage by SC follow-up, early stage by outpatient visit (OP) follow-up, advanced stage by OP follow-up and death (Fig. 3). Women with non-BC history is the state in which the entire hypothetical population starts. The detailed tree structure for screening and treatment in "women with non-BC history" is shown in Fig. 2, where the branches represent pathways of procedures and treatments that lead to the next health state. The first node divides the population into breast cancer positive (BC+) or breast cancer negative (BC−) subject to screening results. BC+ refers to detailed examination including fine-needle aspiration cytology and core needle biopsy. We defined stages 0–I of breast cancer as early stage breast cancer (early stage) and stages II–IV as advanced stage breast cancer (advanced stage) in terms of the TNM classification, and thus those diagnosed as BC+ branched out into the early and advanced stages. BC+ patients were treated in the optimal way and continued to follow-up or death based on screening cancer data. Conversely, we determined that false negatives were found by screening and biopsy at outpatient visits as an interval cancer and shifted them to follow-up or death following interval cancer data. False positives by screening and biopsy were denied at the first outpatient visit and continued to "women with non-BC history" to undergo biennial screening unless they died as a result of non-breast cancer causes.

Early and advanced stage by SC follow-up are those who were detected with breast cancer at the screening and who continued follow-up through outpatient visits. Early and advanced stages by OP follow-up are the early and advanced stage breast cancer patients found at outpatient visits who continued their follow-up on Tree 1. One cycle of simulation is 2 years, and the annual discount rate is 3 % on both cost and effectiveness. Cost-effectiveness analysis was discussed from a social perspective.

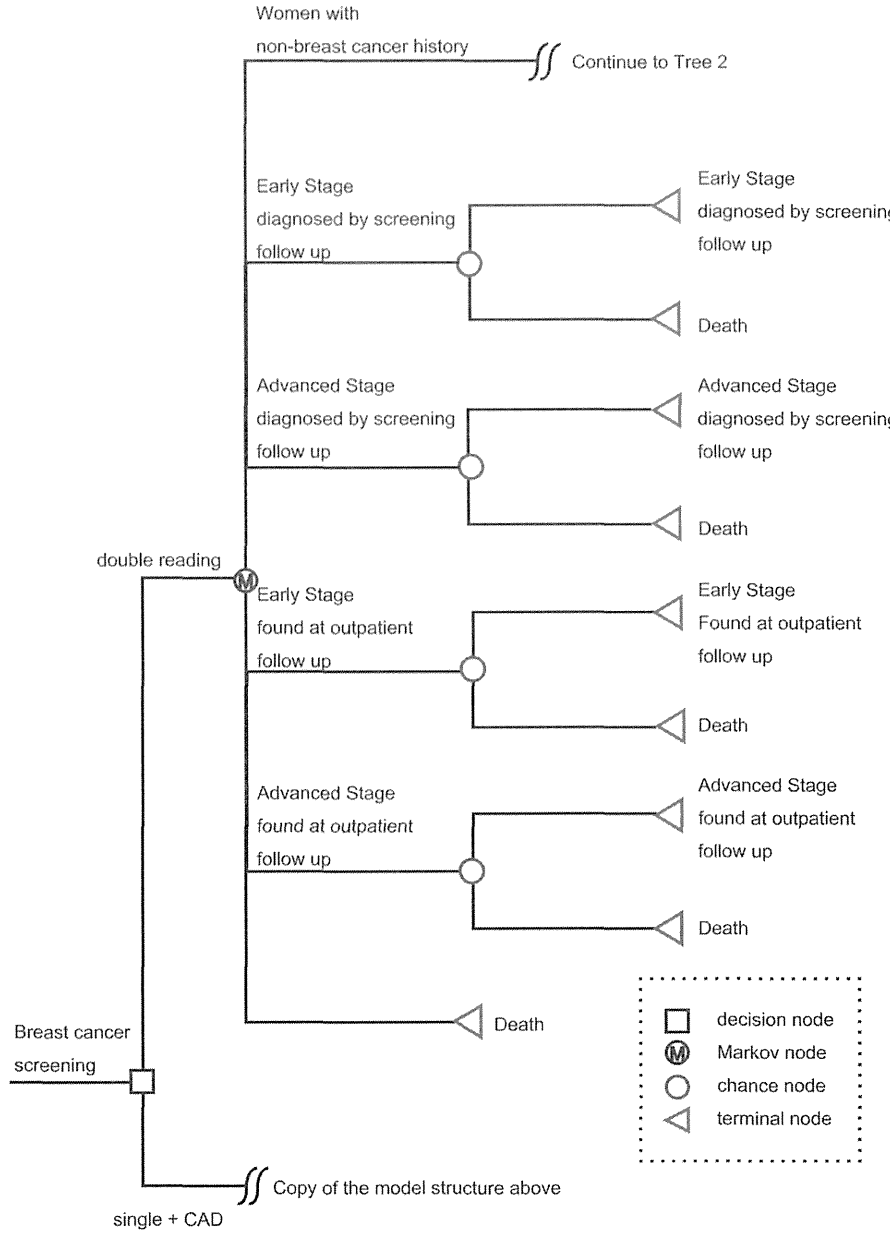

Fig. 1 The analysis model, Tree 1. The hypothetical population comprises 50-year-old women who have undergone breast cancer screening. All start from the health state of "woman with non-breast cancer history," and the details of transitions are shown in Tree 2. The single + CAD method has the same structure as double reading and is omitted from this figure. Terminal nodes at the end of branches jump to the given Markov node, although death is an absorb state

Input data

Base case

The hypothetical population comprised women aged 50 who have undergone biennial breast cancer screening with mammography and palpation. The facility setting was a screening center undertaking population-based screening with an average of about 16,000 examinees annually. Referring physicians pay a reading fee per examinee. The CAD system hypothetically installed in the base case facility includes the basic reading workstation setting with two 5-M monitors, a 3-T server and full-maintenance, and the facility is assumed to have a digital mammography setting. Most of the input data were from Miyagi Cancer Society and Miyagi Prefectural Cancer Registry (MPCR). The other data came from national statistics, the literature and receipt data from our hospital. The institutional ethics committees of Miyagi Cancer Society, MPCR, and our school approved the study protocol.

Probabilities

All probabilities in the base case are summarized in Table 1. Breast cancer prevalence, sensitivity and specificity of screening mammography are from Suzuki et al., who used the data from Miyagi Cancer Society and MPCR [18]. For the biopsy data, we established a database with biopsy results