Comparison of mammography, sonography, MRI and clinical examination in patients with locally advanced or inflammatory breast cancer who underwent neoadjuvant chemotherapy

[1]H J SHIN, MD, [1]H H KIM, MD, [2]J H AHN, MD, [2]S-B KIM, MD, [2]K H JUNG, MD, [3]G GONG, MD, [4]B H SON, MD and [4]S H AHN, MD

Departments of [1]Radiology and Research Institute of Radiology, [2]Oncology, [3]Pathology and [4]Surgery, Asan Medical Centre, University of Ulsan, College of Medicine, 388-1 Pungnap-2-dong, Songpa-Gu, Seoul, 138-736, Korea

Objectives: The purpose of this study was to determine the relative accuracies of mammography, sonography, MRI and clinical examination in predicting residual tumour size and pathological response after neoadjuvant chemotherapy for locally advanced or inflammatory breast cancer. Each prediction method was compared with the gold standard of surgical pathology.
Methods: 43 patients (age range, 25–62 years; mean age, 42.7 years) with locally advanced or inflammatory breast cancer who had been treated by neoadjuvant chemotherapy were enrolled prospectively. We compared the predicted residual tumour size and the predicted response on imaging and clinical examination with residual tumour size and response on pathology. Statistical analysis was performed using weighted kappa statistics and intraclass correlation coefficients (ICC).
Results: The ICC values between predicted tumour size and pathologically determined tumour size were 0.65 for clinical examination, 0.69 for mammography, 0.78 for sonography and 0.97 for MRI. Agreement between the response predictions at mid-treatment and the responses measured by pathology had kappa values of 0.28 for clinical examination, 0.32 for mammography, 0.46 for sonography and 0.68 for MRI. Agreement between the final response predictions and the responses measured by pathology had kappa values of 0.43 for clinical examination, 0.44 for mammography, 0.50 for sonography and 0.82 for MRI.
Conclusion: Predictions of response and residual tumour size made on MRI were better correlated with the assessments of response and residual tumour size made upon pathology than were predictions made on the basis of clinical examination, mammography or sonography. Thus, the evaluation of predicted response using MRI could provide a relatively sensitive early assessment of chemotherapy efficacy.

The advantages of neoadjuvant chemotherapy are multiple and it has been used widely during the past few years [1]. Its primary role is to induce tumour shrinkage and permit breast-conserving surgery, primarily in patients with advanced breast cancer [2–4]. Neoadjuvant chemotherapy allows earlier treatment of micrometastatic disease and the study of biological markers that might predict tumour response [5]. The effectiveness of chemotherapeutic agents in treating both primary breast cancer and potential metastatic disease may be enhanced by the presence of tumour neovascularity. If chemotherapy is given before surgery, while tumour vascularity remains intact, the chemotherapeutic agents may be better able to reach the tumour and thus be more effective.

Neoadjuvant chemotherapy of locally advanced breast cancer (LABC) has also been shown to improve the resectability rate, offering disease-free and overall survival rates that are at least equivalent to those offered by surgery alone [6, 7]. Pathological complete response (pCR) is clinically significant because it is associated with improved long-term prognosis and decreased risk of recurrence [6, 8]. Decisions regarding the continuation of current regimens and the appropriate type and timing of surgery depend on the radiological and clinical assessment of residual tumour size during neoadjuvant chemotherapy [9, 10]. Until now, many studies have shown that physical examinations, mammography and sonography provide suboptimal evaluations of lesion extent that do not allow accurate assessments of pathological response or residual tumour size [5, 11–13]. In the case of LABC, physical examination, mammography or sonography may be suitable for detecting the larger lesions of non-responders, but they have limited sensitivity for responders with smaller residual lesions [14, 15]. For mammography, calcifications may persist or even increase in patients who respond to neoadjuvant chemotherapy [14, 16, 17].

Many previous studies have shown that MRI is the most reliable technique for evaluating residual disease after neoadjuvant chemotherapy, although initial reports described frequent false-negatives with smaller-volume disease [18–27]. Recent studies have increased the

Address correspondence to: Dr Hak Hee Kim, Department of Radiology and Research Institute of Radiology, College of Medicine, University of Ulsan, 388-1 Pungnap-2-dong, Songpa-Gu, Seoul 138-736, South Korea. E-mail: hhkim@amc.seoul.kr

sensitivity of MRI, with increased resolution, reduced slice thickness and lower enhancement thresholds being used to minimise the underestimation of residual disease [15, 22–27]. It is still difficult, however, to distinguish residual scarring, necrosis and fibrosis from viable residual malignancy and to predict accurate response after neoadjuvant chemotherapy, especially in responders. Few published studies have described work with patients with inflammatory breast cancer who underwent neoadjuvant chemotherapy because the incidence of this disease is very low [28, 29]. The purpose of our study was to determine the relative accuracies of mammography, sonography, MRI and clinical examination in predicting residual tumour size and pathological response after neoadjuvant chemotherapy for locally advanced and inflammatory breast cancer. We compared each prediction method with the gold standard of surgical pathology.

Methods and materials

Patient selection and neoadjuvant chemotherapy

Between May 2005 and March 2009, we prospectively enrolled 43 women with locally advanced ($n=19$) or inflammatory breast cancer ($n=24$). Patients who were eligible for this prospective study included patients with locally advanced or inflammatory breast cancer and at least one tumour-positive axillary node who were scheduled to receive neoadjuvant chemotherapy; patients with histologically confirmed breast cancer, including inflammatory breast cancer, that was unsuitable for breast-conserving surgery; and patients with breast cancer and either skin or chest wall involvement who were scheduled for neoadjuvant chemotherapy. The mean patient age was 42.7 years (range, 25–62 years).

The patients received neoadjuvant chemotherapy in an ongoing single-institution randomised Phase II trial. The institutional review board approved the study protocol, and written informed consent was obtained from all patients. In this trial, patients were randomised to receive (1) adriamycin 60 mg m^{-2}, (2) cyclophosphamide 600 mg m^{-2} and docetaxel 75 mg m^{-2} or (3) capecitabine 1000 mg m^{-2}, vinorelbine 25 mg m^{-2} and docetaxel 75 mg m^{-2}. Chemotherapy was administered every 3 weeks with a total of 8 cycles. All 43 patients were evaluated by clinical examination, mammography, sonography and MRI prior to the first course of chemotherapy (baseline assessment), after the fourth course of chemotherapy (mid-treatment assessment) and after the last course of chemotherapy (final assessment). After the last course of chemotherapy, all study patients underwent mastectomy or breast-conserving surgery according to the standard protocols at our institution, which take into account factors such as tumour size, distance from the nipple, multifocality, multicentricity and the patient's desires. All patients underwent ipsilateral axillary node dissection.

Clinical examination, mammography and sonography

All patients underwent clinical examination by one of three breast oncologists. Clinical response assessment was based on change in the tumour's longest diameter, as estimated by palpation at physical examination. Standard bilateral mammograms, with additional views as necessary, were obtained using a Senographe DMR or Senographe DS scanner (GE Medical Systems, Milwaukee, WI). Whole-breast sonography was performed by one of six breast radiologists using broadband linear array transducers (7–12 MHz) on an HDI-5000 or IU-22 unit (Philips Medical Systems, Bothell, WA).

MRI technique and interpretation

All 43 patients were scanned on a 1.5 T scanner (Magnetom Avanto, Siemens Medical Solutions, Erlangen, Germany) with a bilateral breast array coil (Siemens Medical Solutions). The standard MRI protocol included, first, an axial two-dimensional T_2 weighted short tau inversion recovery (STIR) turbo spin-echo pulse sequence (repetition time/echo time/time interval (TR/TE/TI), 6700/74/150 ms, field of view (FOV) 300 × 300 mm, matrix= 448 × 448, slice thickness 5 mm); second, pre- and post-contrast-enhanced fat-saturated axial three-dimensional T_1 weighted fast low angle shot volume-interpolated breathhold examination (FLASH VIBE) pulse sequences (TR/TE, 5.2/2.4 ms, FOV 340 × 340 mm, matrix=384 × 384, slice thickness 0.9 mm); and third, an axial three-dimensional delayed contrast-enhanced turbo spin-echo pulse sequence (TR/TE, 767/12 ms; FOV 350 × 350 mm; matrix=768 × 768; slice thickness, 5 mm) for the evaluation of the supraclavicular and axillary lymph nodes. The six dynamic sequences were performed before and during injection of contrast medium. The contrast medium (0.2 ml kg^{-1} body weight, Magnevist; Schering, Berlin, Germany) was injected before the first dynamic acquisition using a MR compatible power injector (Spectris; Medrad, Pittsburgh, PA) with a flow of 1 ml s^{-1} followed by a 20 ml saline flush. Post-processing manipulation included the production of standard subtraction, reverse subtraction and maximum-intensity-projection images.

Assessment of tumour response

Tumour size was measured by physical examination and imaging studies before the patient began their first chemotherapy treatment, after the fourth course (midtreatment) and after the last course (final). One of three breast oncologists recorded the size of the tumour in two dimensions on physical examination. The tumour size on mammography, sonography and MRI was measured in three dimensions by two breast radiologists, who reached consensus. Disease extent on MRI was assessed on the basis of the size of the lesion and using morphological and kinetic criteria. The distinction between malignant and benign findings was based on a combination of morphological and kinetic features. The criterion for the presence of residual tumour after chemotherapy was the visual observation of residual enhancement at the location of the initial tumour. Time-to-signal intensity curves obtained either in the area of greatest or most homogeneous enhancement or in areas where visible residual enhancement was detected were used to help differentiate residual tumour from adjacent

breast parenchyma. The pathological response to treatment and the pathological residual tumour size were assessed by direct examination of the excised tumour specimens by a breast pathologist. If the patient had a re-excision, the sum of the tumour sizes in each specimen was used as the total amount of residual tumour. Pathological CR (pCR) was defined as the absence of invasive cancer. Near pCR was defined as the presence of only a very small residual invasive cancer of less than 0.3 cm in diameter or of a small number of scattered tumour cells.

To evaluate the response of neoadjuvant chemotherapy, the response evaluation criteria in solid tumours (RECIST) criteria were used for solid tumours in order to simplify the response evaluation procedure [30]. The RECIST classification is based on uni-dimensional measurement of the largest tumour diameter rather than on the previously used bi-dimensional measurement. According to the RECIST guidelines, complete response is defined as no measurable disease, and partial response (PR) is defined as a decrease of at least 30% in tumour diameter. Progressive disease (PD) is defined as an increase of at least 20% in the longest tumour diameter. Stable disease (SD) is defined as a tumour that does not fulfil the criteria for complete or partial response or progressive disease. Each patient's tumour response was classified as CR, PR, SD or PD according to the RECIST guidelines [30] on the basis of tumour measurements made on clinical examination and during the imaging studies.

The baseline tumour size was compared individually with each of the post-chemotherapeutic tumour sizes measured on clinical examination and in each of the imaging studies. For each of the four methods of predicting tumour response to neoadjuvant chemotherapy, we compared the predicted responses with the pathological responses, both of which were determined on the basis of the corresponding baseline measurements. We also determined the sensitivity and specificity of the four methods with respect to the detection of pCR and near pCR.

Two breast radiologists with 5 and 15 year's experience reviewed the imaging studies. Each radiologist was blind to the results from the other radiologist at the initial review. When there was a discrepancy, the two radiologists reviewed these cases together and reached a consensus. The radiologists were blinded to the clinical response and pathology data and recorded the largest diameter according to the RECIST guidelines [30]. We defined tumour size on imaging studies as equal to that determined by pathology if the measurements of longest diameter obtained by the two methods were within 0.5 cm of each other. The tumour size as determined by imaging was defined as an underestimate if it was 0.5 cm or more below that seen on pathology and as an overestimate if 0.5 cm or more greater than that determined on pathology.

We assessed changes in the largest diameter on physical examination, mammography, sonography and MRI between baseline and mid-treatment studies and between the mid-treatment and final studies. We also evaluated the lesion type on mammography, sonography and MRI. On mammography and sonography, we classified the lesion types into three categories: dominant masses, diffuse infiltrative lesions and mass or diffuse infiltrative lesions with calcifications. On MRI, we classified the lesion types into two categories: dominant masses and diffuse or multinodular lesions.

Statistical analysis

Weighted kappa statistics were used to compare the agreement between responses predicted on the basis of clinical examination and imaging studies and the pathological responses using SAS version 9.1 (SAS Institute, Inc., Cary, NC). A kappa value of less than 0.20 was considered as slight agreement, of between 0.21 and 0.40 as fair agreement, of between 0.41 and 0.60 as moderate agreement, of between 0.61 and 0.80 as substantial agreement and of more than 0.81 as almost perfect agreement [31]. Intraclass correlation coefficient analysis was used to evaluate the agreement between the residual tumour size predicted by clinical examination and imaging studies and that measured by pathology, using SPSS software (version 12.0, Chicago, IL). The χ^2 test was used to evaluate whether or not there was a difference in lesions types between responders and non-responders on the basis of assessments made using mammography, sonography and MRI. Receiver operating characteristic (ROC) curve analysis was used to evaluate the diagnostic accuracy, in terms of predicting pCR or near pCR, of change in the largest diameter (LD) between baseline and mid-treatment (LD-1) and between baseline and final study (LD-2).

Results

Patient characteristics are summarised in Table 1. 37 patients underwent mastectomy whereas the remaining 6 patients underwent breast-conserving surgery. The mean delay between the final imaging study and surgery was 18.8 days (median, 14 days; range, 5–50 days). Surgery was targeted at least 2–3 weeks after the final chemotherapy session because the patients needed time to recover their white blood cell counts before surgery. The delay was longer for patients who had complications from chemotherapy, such as fever and neutropenia.

The histological tumour types were 40 invasive ductal carcinomas and 3 micropapillary carcinomas. The mean tumour size on initial presentation was 9.0 cm (range, 2–16 cm) as measured by clinical examination, 7.7 cm (range, 1.8–15 cm) as measured by mammography, 7.3 cm (range, 0.6–17 cm) as measured by sonography and 7.6 cm (range, 2–15 cm) as measured by MRI. On pathological examination after surgery, the mean residual tumour size was 3.3 cm and the range was from 0 to 12 cm. Lymph node metastasis was present in 25 patients, and among these, 13 patients showed 3 or more lymph node metastases.

On pathological examination after surgery, 8 of 43 patients (19%) achieved pCR and 4 (9%) achieved near pCR. In the four patients with near pCR, the pathological residual lesion measured 0.1 cm in three patients and 0.2 cm in one patient. Of the 12 patients with pCR or near pCR, mammography accurately predicted pCR or near pCR in 5 cases (42%; 95% confidence interval (CI), 15–72%), sonography in just 2 cases (17%; 95% CI,

Table 1. Pre-treatment patient characteristics (*n*=43)

Characteristic	
Age in years, mean (range)	42.7 (25–62)
Tumour size (longest dimension) at baseline (cm), mean (range)	
By clinical examination	9.0 (2–16)
By mammography	7.7 (1.8–15)
By sonography	7.3 (0.6–17)
By MRI	7.6 (2–15)
Histological subtype	
Invasive ductal carcinoma	40
Micropapillary carcinoma	3
Her-2/neu status	
Negative	29
Overexpression present	14
Oestrogen receptor status	
Positive	17
Negative	26
Lymph node metastasis	
Negative	18
Positive	25

2–48%), clinical examination in 7 cases (58%; 95% CI, 28–85%) and MRI in 9 cases (75%; 95% CI, 43–95%). Of the 8 patients with pCR, mammography accurately predicted pCR in 3 cases (38%; 95% CI, 10–74%), sonography in just 1 patient (13%; 95% CI, 7–53%), clinical examination in 4 cases (50%; 95% CI, 17–83%) and MRI in 6 cases (75%; 95% CI, 36–96%).

There was no significant difference between responders and non-responders in terms of the morphology on mammography (*p*=0.437), sonography (*p*=0.302) or MRI (*p*=0.583). There was also no significant difference in the frequencies of inflammatory and non-inflammatory breast cancers between responders and non-responders (*p*=0.270).

Mid-treatment predicted response

For all patients with locally advanced or inflammatory breast cancer, the agreement between the mid-treatment predicted responses and the actual pathological responses are shown for each assessment method in Table 2. The predicted treatment response agreed with the pathologically determined treatment response for 58% of predictions made on the basis of clinical examination, 56% of predictions based on mammography, 63% of predictions based on sonography and 67% of predictions based on MRI. The kappa values were 0.28 (95% CI, 0.03–0.54) for clinical examination, 0.32 (95% CI, 0.07–0.57) for mammography, 0.46 (95% CI, 0.13–0.93) for sonography and 0.68 (95% CI, 0.18–0.99) for MRI.

For the 24 patients with inflammatory breast cancer, the predicted treatment response agreed with the pathologically determined treatment response for 46% of predictions made on the basis of clinical examination, 50% of predictions based on mammography, 63% of predictions based on sonography and 67% of predictions based on MRI. The kappa values were 0.23 (95% CI, 0.03–0.42) for clinical examination, 0.26 (95% CI, 0.06–0.45) for mammography, 0.37 (95% CI, 0.09–0.65) for sonography and 0.71 (95% CI, 0.17–0.99) for MRI. MRI showed the best overall agreement between the mid-treatment predicted response and the pathological response.

Final predicted response

The mid-treatment predicted responses and the actual pathological responses are shown for each assessment method in Table 3. Of the total 43 study patients, 3 underwent only 4 cycles of chemotherapy because their estimated response on the basis of clinical examination or imaging studies was stable or progressive disease, or because the patient's condition was not appropriate for

Table 2. Agreement between the mid-treatment predicted responses and the pathologically determined responses

Predicted response	Pathological response[a]			
	CR	PR	SD	PD
Measured by CE				
CR	7	1	0	0
PR	4	17	4	0
SD	1	7	1	1
PD	0	0	0	0
Measured by MG				
CR	3	1	0	0
PR	6	16	1	0
SD	3	8	5	0
PD	0	0	0	0
Measured by US				
CR	2	0	0	0
PR	9	17	1	1
SD	1	4	8	0
PD	0	0	0	0
Measured by MRI				
CR	4	0	0	0
PR	8	18	1	0
SD	0	5	6	0
PD	0	0	0	1

CE, clinical examination; CR, complete response; MG, mammography; PD, progressive disease; PR, partial response; SD, stable disease; US, sonography.
[a]Pathological response was evaluated on the basis of the clinical examination or imaging findings at baseline, which differed according to the method used to predict residual tumour size. Hence, the number of patients in each of the pathological response categories (PR, SD and PD) is not the same for each prediction method.

Table 3. Agreement between the final predicted responses and the pathologically determined responses

Final predicted response	Pathological response[a]			
	CR	PR	SD	PD
Measured by CE				
CR	7	7	0	0
PR	5	13	3	0
SD	0	4	0	0
PD	0	0	1	0
Measured by MG				
CR	5	0	0	0
PR	5	19	0	0
SD	2	6	3	0
PD	0	0	0	0
Measured by US				
CR	2	0	0	0
PR	8	18	1	0
SD	2	3	5	0
PD	0	0	0	1
Measured by MRI				
CR	9	1	0	0
PR	3	20	0	0
SD	0	2	5	0
PD	0	0	0	0

CE, clinical examination; CR, complete response; MG, mammography; PD, progressive disease; PR, partial response; SD, stable disease; US, sonography.

[a]Pathological response was evaluated on the basis of the clinical examination or imaging findings at baseline, which differed according to the method used to predict residual tumour size. Hence, the number of patients in each of the pathological response categories (PR, SD and PD) is not the same for each prediction method.

them to receive chemotherapy. Therefore, the final predicted response was reviewed for 40 patients. The predicted treatment response agreed with the pathologically determined treatment response for 50% of predictions made on the basis of clinical examination, 68% of predictions based on mammography, 65% of predictions based on sonography and 85% of predictions based on MRI. The kappa values were 0.43 (95% CI, 0.02–0.86) for clinical examination, 0.44 (95% CI, 0.07–0.79) for mammography, 0.50 (95% CI, 0.11–0.79) for sonography and 0.82 (95% CI, 0.51–0.98) for MRI.

Of 24 patients with inflammatory breast cancer, 2 patients underwent only 4 cycles of chemotherapy. For these patients, the predicted treatment response agreed with the pathologically determined treatment response for 41% of predictions made on the basis of clinical examination, 64% of predictions based on mammography, 73% of predictions based on sonography and 82% of predictions based on MRI. The kappa values were 0.44 (95% CI, 0.03–0.96) for clinical examination, 0.42 (95% CI, 0.06–0.87) for mammography, 0.62 (95% CI, 0.09–0.99) for sonography and 0.82 (95% CI, 0.48–0.99) for MRI. Once again, MRI showed the best overall agreement between the final predicted response and the pathological response.

Evaluation of residual tumour size

When the residual tumour sizes measured by clinical examination and imaging studies were compared with the residual tumour size measured upon pathological examination, the intraclass correlation coefficient value was 0.65 for clinical examination, 0.69 for mammography, 0.78 for sonography and 0.97 for MRI (Table 4). MRI provided the best agreement. In the prediction of pCR *vs* residual disease, clinical examination had a sensitivity of 50% (95% CI, 17–83%) and a specificity of 71%. The sensitivity and specificity of each of the imaging modalities were as follows: sensitivity 38% (95% CI, 10–74%) and specificity 94% for mammography, sensitivity 13% (95% CI, 7–53%) and specificity 97% for sonography, and sensitivity 75% (95% CI, 36–96%) and specificity 89% for MRI. The accuracy was 67% for clinical examination, 84% for mammography, 81% for sonography and 86% for MRI. MRI showed greater sensitivity in predicting CR than clinical examination, mammography or sonography (Figure 1). In the prediction of pCR and near pCR *vs* residual disease, the sensitivity was 58% (95% CI, 28–85%) and specificity 77% for clinical examination, sensitivity 42% (95% CI, 15–72%) and specificity 100% for mammography, sensitivity 17% (95% CI, 2–48%) and specificity 100% for sonography, and sensitivity 75% (95% CI, 43%–95%) and specificity 97% for MRI. The accuracy was 72% for clinical examination, 84% for mammography, 77% for sonography and 91% for MRI. MRI showed greater accuracy and sensitivity for predicting pCR and near pCR than clinical examination, mammography or sonography.

Predictions of tumour size were estimated as equal to the measurements made on pathology for 21% (9/43) of predictions based on clinical examination, 28% (12/43) of predictions based on mammography, 40% (17/43) of predictions based on sonography and 74% (32/43) of predictions based on MRI. Clinical examination underestimated the pathologically determined size of the residual tumour in 37% (16/43) of patients and overestimated it in 42% (18/43). Mammography underestimated the size of the residual tumour in 14% (6/43) of patients and overestimated it in 58% (25/43). Sonography underestimated the size of the residual tumour in 16% (7/43) of patients and overestimated it in 44% (19/43). MRI underestimated the size of the residual tumour in 7% (3/43) of patients and overestimated it in 19% (8/43).

Table 5 shows the results of the ROC curve analysis of the ability of change in the LD to predict pCR and near pCR. Changes of LD predicted by MRI showed good (A_z=0.88) and excellent (A_z=0.92) performance in predicting pCR or near pCR, respectively. On mammography and sonography, LD-1 showed better accuracy in predicting pCR or near pCR than LD-2. A reduction in LD between baseline and mid-treatment of less than 43% on mammography or of less than 60% on sonography indicated the presence of residual tumour. On MRI, a reduction in LD of less than 35% at LD-1 or of less than 85% reduction at LD-2 were indicative of residual tumour.

Discussion

Neoadjuvant chemotherapy is increasingly used to treat locally advanced breast cancer, thus allowing more breast-conserving surgery to be performed by shrinking

Table 4. Agreement between residual tumour size as predicted by physical examination, mammography, sonography or MRI and residual tumour size determined by pathology

Residual tumour size	ICC	95% CI
Predicted by CE	0.65	0.51–0.72
Predicted by MG	0.69	0.49–0.82
Predicted by US	0.78	0.63–0.87
Predicted by MRI	0.97	0.95–0.98

CE, clinical examination; CI, confidence interval; ICC, intra-class correlation coefficient value; MG, mammography; US, sonography.

larger tumours [6, 32]. A sensitive and specific method to identify tumour responses to neoadjuvant chemotherapy is needed because early recognition of non-responders facilitates an earlier change to a more effective regime, thereby minimising toxicity and optimising the timing of surgery. In addition, lack of response to a particular regime *in vivo* may guide additional chemotherapy after surgery.

Physical examination has served as the gold standard for assessing the clinical response to chemotherapy. In previous studies, correlation with the pathologically assessed residual tumour size ranged from 0.42 to 0.68 for tumour sizes assessed by clinical examination, from 0.33 to 0.84 for tumour sizes assessed by mammography and from 0.29 to 0.89 for tumour sizes assessed by sonography [14, 32–34]. Clinical examination had been found to have a limited value in predicting residual tumour size after neoadjuvant chemotherapy [14, 27, 33–35].

Dense breast tissue and the infiltrating nature of the growth of locally advanced or inflammatory breast cancer are two major factors that might make it difficult to evaluate exact tumour size and response rate after neoadjuvant chemotherapy on mammography. Dense breast tissue often obscures the tumour margin on mammography, thus making size determination difficult. In our study, we determined the extent of the tumour by evaluating a combination of findings such as asymmetric increased density, bulging contour and associated calcifications. In cases where the whole breast parenchyma is involved on initial mammography, however, it was difficult to evaluate the exact extent of the residual tumour after neoadjuvant chemotherapy. In our study, the predicted response made at mid-treatment agreed with the pathologically determined response in 58%, 56% and 63% of cases when predictions were based on clinical examination, mammography and sonography, respectively. For the final predicted response, these percentages were 50%, 68% and 65%, respectively. The kappa values for response evaluations based on clinical examination, mammography and sonography were lower than that for MRI.

In their study of 162 patients, Peintinger et al [9] showed that a combination of mammography and sonography provided a high accuracy for predicting pCR and a moderate agreement in predicting pathological residual tumour size after neoadjuvant chemotherapy. In their study, the accuracies of mammography and sonography were reduced for the invasive lobular histological tumour type, which was associated with an underestimation of residual tumour size [9]. In our study, pCR and near pCR were accurately predicted by clinical examination in 58% of our study patients, by

mammography in 42%, by sonography in 17% and by MRI in 75%. As regards ability to predict pCR and near pCR, the sensitivity was 58% for clinical examination, 42% for mammography, 17% for sonography and 75% for MRI, whereas the specificity was 77% for clinical examination, 100% for mammography, 100% for sonography and 97% for MRI. MRI showed greater sensitivity in predicting pCR and near pCR than did clinical examination, mammography or sonography.

In our study, predictions made on the basis of MRI showed a better correlation with the pathological response and pathological residual size after neoadjuvant chemotherapy than did estimations made on the basis of clinical examination, mammography or sonography. Bonadonna et al [32] suggested that it is difficult to interpret the published results of studies on neoadjuvant chemotherapy because of differences in the assessment or lack of information on the method of assessment of tumour response. Several studies have shown that MRI more frequently underestimated than overestimated residual disease, but these studies used both volumetric assessments and less sensitive imaging techniques [20]. The recent study of Yeh et al [27] also showed that MRI underestimated residual disease in 23% of their study patients. By contrast, Partridge et al [24] described a high frequency of false-positive MRI results in responders to chemotherapy when using current, high-sensitivity MRI protocols. In addition, Rosen et al [25] reported that MRI resulted in an overestimation of residual disease in 33% of their study patients and an underestimation in only 5%.

Recent improvements have been achieved by the introduction of newer MRI techniques and more standardised interpretation criteria. For example, Wasser et al [36] and Martincich et al [37] used high temporal resolution dynamic contrast-enhanced MRI. They reported that certain patterns of parameterised uptake in breast cancer show correlation with vascular permeability and expression of the vascular endothelial growth factor. More recently, Kwong et al [10] reported that MRI frequently overestimated residual disease in responders to chemotherapy treatment. In our study, MRI overestimated residual disease in 19% and underestimated residual disease in 7% of our study patients. When only inflammatory breast cancer is considered, the response predicted by MRI was that which best correlated with pathological response.

There are several limitations to our study. First, our study included relatively small numbers of patients and the number of pCR was also small. As a result, our assessments of sensitivity for estimating pCR or near pCR, in particular, are likely to be imprecise and hence we have provided 95% confidence intervals for sensitivity estimates. However, the proportion of inflammatory breast cancers in this study relatively high (56%, 24/43) compared with the incidence of this disease type (1–5%). Hence the data on inflammatory breast cancers presented here might add some valuable information to that provided by previous studies. To our knowledge, there have been only a few studies on neoadjuvant chemotherapy for inflammatory breast cancer. Second, we used two different neoadjuvant chemotherapy protocols, and we believe that chemotherapeutic agents may affect the imaging appearance, especially with regard to
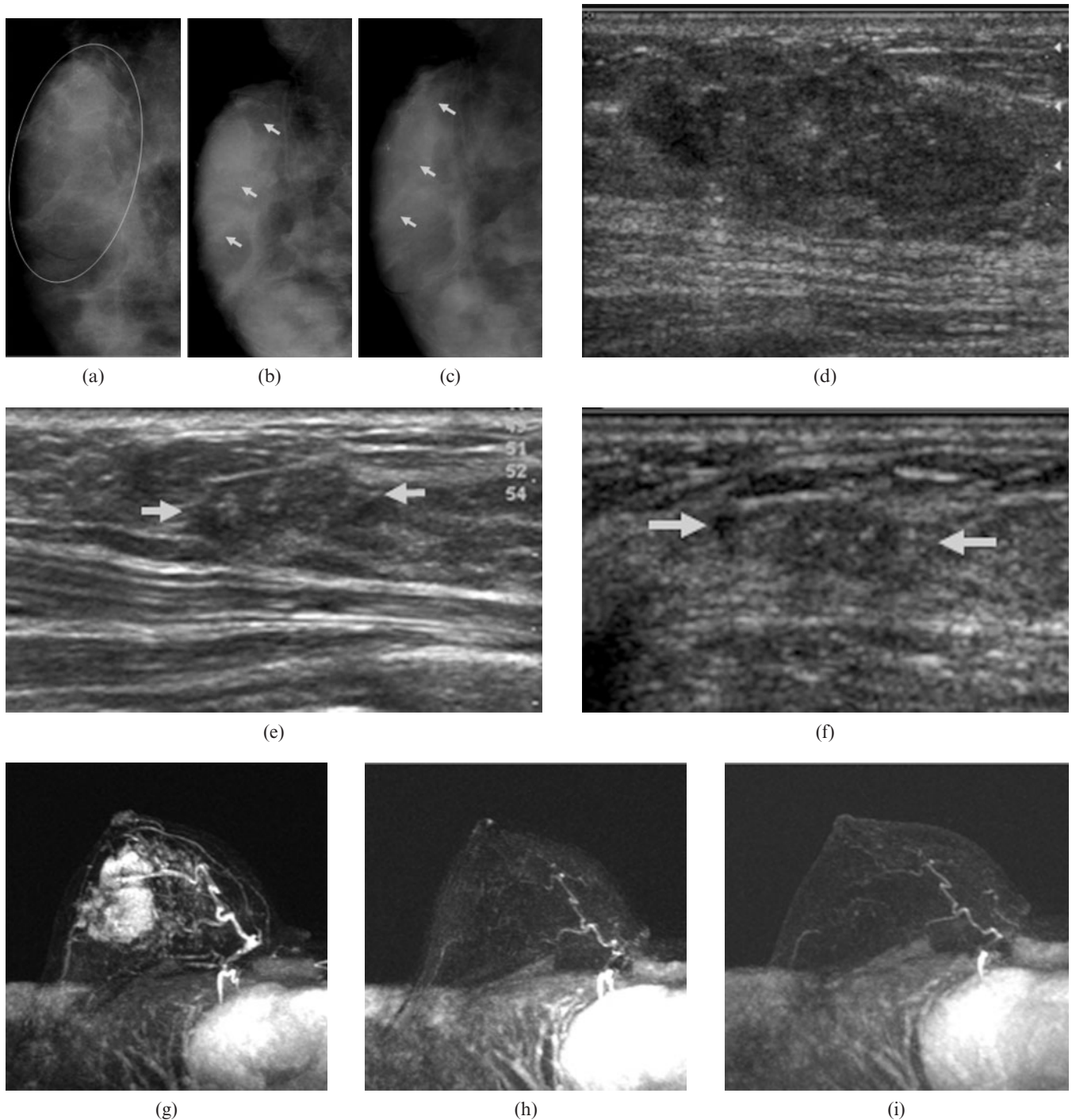
Figure 1. A 30-year-old woman with a palpable lump in the upper outer quadrant of the right breast. (a) The initial mediolateral oblique mammogram shows an approximately 6 cm mass with pleomorphic calcifications (encircled) in the upper outer quadrant of the right breast. Serial follow-up mammograms after (b) four and (c) eight cycles of chemotherapy show the decreased size of the large malignant mass, but also show remaining pleomorphic calcifications with segmental distribution (arrows). This was considered to be a partial response. (d) The initial sonogram shows a 4.2 cm mass with calcifications in the right breast. Two serial follow-up sonograms after (e) four and (f) eight cycles of chemotherapy show that the size of the mass has decreased markedly, but a vague mass with calcifications remains (arrows). This was considered to be a partial response. (g) A maximum intensity projection image on the initial MRI shows an approximately 5 cm enhancing mass in the right breast. Two serial follow-up MRI images after (h) four and (i) eight cycles of chemotherapy show no residual enhancing lesion in the right breast. This was interpreted as complete response on MRI. Surgery confirmed a 0.1-cm ductal carcinoma *in situ*, which indicated a pathological complete response.

the contrast uptake on MRI [38]. Third, two experienced radiologists evaluated the imaging studies in consensus, so we were not able to assess intra- and interobserver variability. Fourth, the gold standard measurement of

response was based on the baseline results for the method under assessment. This may create a bias but there was no alternative reliable baseline measurement of tumour size.

Table 5. Ability of change in the largest tumour diameter, as predicted by clinical examination, mammography, sonography and MRI, to indicate complete[a] and incomplete responses at pathology

Change of the largest tumour diameter		$A_z{}^b$	95% CI
Between baseline and mid-treatment study (LD-1)	Predicted by CE	0.695	0.54–0.83
	Predicted by MG	0.805	0.66–0.91
	Predicted by US	0.839	0.70–0.93
	Predicted by MRI	0.878	0.78–0.98
Between baseline and final study (LD-2)	Predicted by CE	0.743	0.58–0.91
	Predicted by MG	0.744	0.58–0.87
	Predicted by US	0.759	0.59–0.88
	Predicted by MRI	0.924	0.83–0.98

CE, clinical examination; MG, mammography; US, sonography; CI, confidence interval.
[a]Complete responses include pathological complete response (pCR) and near pCR.
[b]A_z values represent the diagnostic performance of CE, MG, US and MRI in predicting pCR and near pCR.

Conclusion

In patients undergoing neoadjuvant chemotherapy, predictions of treatment response evaluated on the basis of MRI either at mid-treatment or just before surgery and estimates of residual tumour size made on the basis of MRI just before surgery appear to better correlate with pathological results than estimates or predictions based on mammography, sonography or clinical examination. MRI is not, however, perfect. It may overestimate or underestimate residual disease in some patients. Further studies are needed to assess the value of MRI for evaluating the response and thus the efficacy of chemotherapy earlier in a course of neoadjuvant chemotherapy.

References

1. Bonadonna G, Valagussa P. Primary chemotherapy in operable breast cancer. Semin Oncol 1996;23:464–74.
2. Newman L, Buzdar A, Singletary S, Kuerer H, Buchholz T, Ames F, et al. A prospective trial of preoperative chemotherapy in resectable breast cancer: predictors of breast-conservation therapy feasibility. Ann Surg Oncol 2002;9:228–34.
3. Kuerer H, Singletary S, Buzdar A, Ames F, Valero V, Buchholz T, et al. Surgical conservation planning after neoadjuvant chemotherapy for stage II and operable stage III breast carcinoma. Am J Surg 2001;182:601–8.
4. Vlastos G, Mirza N, Lenert J, Hunt K, Ames F, Feig B, et al. The feasibility of minimally invasive surgery for stage IIA, IIB, and IIIA breast carcinoma patients after tumour downstaging with induction chemotherapy. Cancer 2000; 88:1417–24.
5. Booser D, Hortobagyi G. Treatment of locally advanced breast cancer. Semin Oncol 1992;19:278–85.
6. Fisher B, Bryant J, Wolmark N, Mamounas E, Brown A, Fisher E, et al. Effect of preoperative chemotherapy on the outcome of women with operable breast cancer. J Clin Oncol 1998;16:2672–85.
7. Kaufmann M, von Minckwitz G, Smith R, Valero V, Gianni L, Eiermann W, et al. International expert panel on the use of primary (preoperative) systemic treatment of operable breast cancer: review and recommendations. J Clin Oncol 2003;21:2600–8.
8. Kuerer H, Newman L, Buzdar A, Hunt K, Dhingra K, Buchholz T, et al. Residual metastatic axillary lymph nodes following neoadjuvant chemotherapy predict disease-free survival in patients with locally advanced breast cancer. Am J Surg 1998;176:502–9.
9. Peintinger F, Kuerer H, Anderson K, Boughey J, Meric-Bernstam F, Singletary S, et al. Accuracy of the combination of mammography and sonography in predicting tumour response in breast cancer patients after neoadjuvant chemotherapy. Ann Surg Oncol 2006;13:1443–9.
10. Kwong M, Chung G, Horvath L, Ward B, Hsu A, Carter D, et al. Postchemotherapy MRI overestimates residual disease compared with histopathology in responders to neoadjuvant therapy for locally advanced breast cancer. Cancer J 2006;12:212–21.
11. Singletary S, McNeese M, Hortobagyi G. Feasibility of breast-conservation surgery after induction chemotherapy for locally advanced breast carcinoma. Cancer 1992;69: 2849–52.
12. Segel M, Paulus D, Hortobagyi G. Advanced primary breast cancer: assessment at mammography of response to induction chemotherapy. Radiology 1988;169:49–54.
13. Swain S, Sorace R, Bagley C, Danforth DJ, Bader J, Wesley M, et al. Neoadjuvant chemotherapy in the combined modality approach of locally advanced nonmetastatic breast cancer. Cancer Res 1987;47:3889–94.
14. Herrada J, Iyer R, Atkinson E, Sneige N, Buzdar A, Hortobagyi G. Relative value of physical examination, mammography, and breast sonography in evaluating the size of the primary tumour and regional lymph node metastases in women receiving neoadjuvant chemotherapy for locally advanced breast carcinoma. Clin Cancer Res 1997;3:1565–9.
15. Weatherall P, Evans G, Metzger G, Saborrian M, Leitch A. MRI vs. histologic measurement of breast cancer following chemotherapy: comparison with X-ray mammography and palpation. J Magn Reson Imaging 2001;13:868–75.
16. Moskovic E, Mansi J, King D, Murch C, Smith I. Mammography in the assessment of response to medical treatment of large primary breast cancer. Clin Radiol 1993;47:339–44.
17. Vinnicombe S, MacVicar A, Guy R, Sloane J, Powles T, Knee G, et al. Primary breast cancer: mammographic changes after neoadjuvant chemotherapy, with pathologic correlation. Radiology 1996;198:333–40.
18. Gilles R, Guinebretiere J, Toussaint C, Spielman M, Rietjens M, Petit J, et al. Locally advanced breast cancer: contrast-enhanced subtraction MR imaging of response to preoperative chemotherapy. Radiology 1994;191:633–8.
19. Abraham D, Jones R, Jones S, Cheek J, Peters G, Knox S, et al. Evaluation of neoadjuvant chemotherapeutic response of locally advanced breast cancer by magnetic resonance imaging. Cancer 1996;78:91–100.
20. Mumtaz H, Davidson T, Hall-Craggs M, Payley M, Walmsley K, Cowley G, et al. Comparison of magnetic resonance imaging and conventional triple assessment in locally recurrent breast cancer. Br J Surg 1997;84:1147–51.
21. Rieber A, Zeitler H, Rosenthal H, Gorich J, Kreienberg R, Brambs H, et al. MRI of breast cancer: influence of chemotherapy on sensitivity. Br J Radiol 1997;70:452–8.

22. Drew P, Kerin M, Mahapatra T, Malone C, Monson J, Turnbull L, et al. Evaluation of response to neoadjuvant chemoradiotherapy for locally advanced breast cancer with dynamic contrast-enhanced MRI of the breast. Eur J Surg Oncol 2001;27:617–20.

23. Balu-Maestro C, Chapellier C, Bleuse A, Chanalet I, Chauvel C, Largillier R. Imaging in evaluation of response to neoadjuvant breast cancer treatment benefits of MRI. Breast Cancer Res Treat 2002;72:145–52.

24. Partridge S, Gibbs J, Lu Y, Esserman L, Sudilovsky D, Hylton N. Accuracy of MR imaging for revealing residual breast cancer in patients who have undergone neoadjuvant chemotherapy. AJR Am J Roentgenol 2002;179:1193–9.

25. Rosen E, Blackwell K, Baker J, Soo M, Bentley R, Yu D, et al. Accuracy of MRI in the detection of residual breast cancer after neoadjuvant chemotherapy. AJR Am J Roentgenol 2003;181:1275–82.

26. Partridge S, Gibbs J, Lu Y, Esserman L, Tripathy D, Wolverton D, et al. MRI measurements of breast tumour volume predict response to neoadjuvant chemotherapy and recurrence-free survival. AJR Am J Roentgenol 2005;184:1774–81.

27. Yeh E, Slanetz P, Kopans D, Rafferty E, Georgian-Smith D, Moy L, et al. Prospective comparison of mammography, sonography, and MRI in patients undergoing neoadjuvant chemotherapy for palpable breast cancer. AJR Am J Roentgenol 2005;184:868–77.

28. Chen JH, Mehta RS, Nalcioglu O, Su MY. Inflammatory breast cancer after neoadjuvant chemotherapy: can magnetic resonance imaging precisely diagnose the final pathological response? Ann Surg Oncol 2008;15:3609–13.

29. Yang WT, Le-Petross HT, Macapinlac H, Carkaci S, Gonzalez-Angulo AM, Dawood S, et al. Inflammatory breast cancer: PET/CT, MRI, mammography, and sonography findings. Breast Cancer Res Treat 2008;109:417–26.

30. Therasse P, Arbuck S, Eisenhauer E, Wanders J, Kaplan R, Rubinstein L, et al. New guidelines to evaluate the response to treatment in solid tumors. European Organization for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. J Natl Cancer Inst 2000;92:205–16.

31. Landis JR, Koch GG. An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. Biometrics 1977;33:363–74.

32. Bonadonna G, Valagussa P, Brambilla C, Ferrari L, Moliterni A, Terenziani M, et al. Primary chemotherapy in operable breast cancer: eight-year experience at the Milan Cancer Institute. J Clin Oncol 1998;16:93–100.

33. Helvie M, Joynt L, Cody R, Pierce L, Adler D, Merajver S. Locally advanced breast carcinoma: accuracy of mammography versus clinical examination in the prediction of residual disease after chemotherapy. Radiology 1996;198:327–32.

34. Dershaw D, Drossman S, Liberman L, Abramson A. Assessment of response to therapy of primary breast cancer by mammography and physical examination. Cancer 1995;75:2093–8.

35. Chagpar A, Middleton L, Sahin A, Dempsey P, Buzdar A, Mirza A, et al. Accuracy of physical examination, ultrasonography, and mammography in predicting residual pathologic tumor size in patients treated with neoadjuvant chemotherapy. Ann Surg 2006;243:257–64.

36. Wasser K, Klein S, Fink C, Junkermann H, Sinn H, Zuna I, et al. Evaluation of neoadjuvant chemotherapeutic response of breast cancer using dynamic MRI with high temporal resolution. Eur Radiol 2003;13:80–7.

37. Martincich L, Montemurro F, De Rosa G, Marra V, Ponzone R, Cirillo S, et al. Monitoring response to primary chemotherapy in breast cancer using dynamic contrast-enhanced magnetic resonance imaging. Breast Cancer Res Treat 2004;83:67–76.

38. Delille J, Slanetz P, Yeh E, Halpern E, Kopans D, Garrido L. Invasive ductal breast carcinoma response to neoadjuvant chemotherapy: noninvasive monitoring with functional MR imaging pilot study. Radiology 2003;228:63–9.