Neoadjuvant chemotherapy in breast cancer: early response prediction with quantitative MR imaging and spectroscopy

A prospective study was undertaken in women undergoing neoadjuvant chemotherapy for locally advanced breast cancer in order to determine the ability of quantitative magnetic resonance imaging (MRI) and proton spectroscopy (MRS) to predict ultimate tumour response (percentage decrease in volume) or to detect early response. Magnetic resonance imaging and MRS were carried out before treatment and after the second of six treatment cycles. Pharmacokinetic parameters were derived from $T_1$-weighted dynamic contrast-enhanced MRI, water apparent diffusion coefficient (ADC) was measured, and tissue water:fat peak area ratios and water:fat ratios were measured using unsuppressed one-dimensional proton spectroscopic imaging (30 and 135 ms echo times). Pharmacokinetic parameters and ADC did not detect early response; however, early changes in water:fat ratios and water $T_2$ (after cycle two) demonstrated substantial prognostic efficacy. Larger decreases in water $T_2$ accurately predicted final volume response in 69% of cases (11/16) while maintaining 100% specificity and positive predictive value. Small/absent decreases in water:fat ratios accurately predicted final volume non-response in 50% of cases (3/6) while maintaining 100% sensitivity and negative predictive value. This level of accuracy might permit clinical application where early, accurate prediction of non-response would permit an early change to second-line treatment, thus sparing patients unnecessary toxicity, psychological morbidity and delay of initiation of effective treatment.

British Journal of Cancer (2006) 94, 427 – 435. doi:10.1038/sj.bjc.6602948 www.bjcancer.com
© 2006 Cancer Research UK

Keywords: breast neoplasms; chemotherapy; magnetic resonance; MR; spectroscopy; contrast enhancement; diffusion study

Recent statistics confirm that breast cancer remains one of the most prevalent and serious forms of neoplastic disease in the UK. Breast cancer accounts for more than one in four female cancer cases and currently demonstrates age-standardised incidence and mortality rates of 117 and 31 per 100,000, respectively (Toms, 2003).

Some primary breast cancers are considered inoperable at diagnosis owing to their problematical location and size, with tumours greater than 3 cm in diameter being associated with an increased risk of disseminated disease. In such cases, neoadjuvant chemotherapy is routinely used before surgery to increase the chances of a successful outcome. Assessing tumour response to chemotherapy is crucial to patient management and is currently achieved by monitoring changes in tumour size using clinical examination backed up by longitudinal X-ray or ultrasound mammography. A poor response to the primary treatment regime usually prompts either change of chemotherapy regime or an early resort to surgery. Poor response might also require a greater degree of surgical intervention.

Repeated X-ray mammography (XRM) has several drawbacks: discomfort, radiation exposure, geometric distortion owing to compression and magnification render tumour volume measurements less accurate than magnetic resonance imaging (MRI), and tumour may be impossible to distinguish from dense glandular tissue and fibrosis. Magnetic resonance imaging has been shown to correspond better with pathological size measurement than XRM (Esserman et al, 1999; Drew et al, 2001) and to be less obscured by dense breast diseases. Ultrasound measurement is more accurate than XRM, but fails when the tumour to be measured is larger than the field of view or complex in shape. Magnetic resonance imaging is also more accurate than ultrasound at detecting small volume residual tumour.

However, even if tumour volume changes can be accurately assessed by MRI, they may manifest themselves later than changes in underlying tumour function such as vascular density or permeability (Wasser et al, 2003). Therefore, vascular or metabolic parameters might provide a more sensitive indicator of early tumour response, thus permitting individual treatment regimes to be adjusted more rapidly, and sparing patients unnecessary morbidity, expense and delay in initiation of effective treatment. Treatment-induced changes in metabolism can be investigated using both positron emission tomography (Byrne et al, 2004; Kumar and Alavi, 2004; Weber, 2005) and single photon emission computed tomography (Buscombe et al, 1997), but the clinical utility of such examinations is reduced by the need to limit repeated radiation doses.

Treatment-induced changes in tumour neovascularity can be assessed using dynamic contrast-enhanced MRI (DCE-MRI) and...
quantified using pharmacokinetic (PK) modelling (Padhani, 2002; Knopp et al., 2003). Diffusion-weighted MRI can be used to detect changes in the apparent diffusion coefficient (ADC) for tissue water associated with changes in tissue and intracellular structure (Zhao et al., 1996; Gallons et al., 1999; Ross et al., 2003; Moffat et al., 2005). Proton magnetic resonance spectroscopy (MRS) can be used to characterise breast lesions through differences in the ratio of fat and water signals (Chu et al., 1987; Sijens et al., 1988) or the intensity of signal from choline-containing compounds (Bradamante et al., 1988; Katz-Brull et al., 2002; Bolan et al., 2003; Jacobs et al., 2004; Jacobs et al., 2005). Both fat:water ratios (Jagannathan et al., 1998) and choline levels (Freud et al., 2000; Jagannathan et al., 2001; Schwarz et al., 2002; Meisamy et al., 2004) have also been used to monitor treatment-induced changes, and phosphorous MRS has been used to predict response using differences in the levels of phosphocholine or phosphomonoesters (Shukla-Dave et al., 2002a, b; Arias-Mendoza et al., 2004). Sodium MRI might also be useful in monitoring treatment response (Kline et al., 2000; Scheppin et al., 2005).

A prospective study was, therefore, undertaken whereby three of these techniques (DCE-MRI, ADC measurement and MRS fat:water ratios) were carried out in the same patients to determine in each technique the relative prognostic utility. The study hypothesis was that a combination of the quantitative data from the three techniques would lead to a synergistic increase in prognostic accuracy and, therefore, provide an accurate, reliable and non-invasive early indication of tumour response to chemotherapy, which could potentially have a positive influence on patient management.

MATERIALS AND METHODS

Chemotherapy regime and relative timing of MRI

All women who were scheduled to undergo neoadjuvant chemotherapy for primary inoperable locally advanced breast cancer were approached to join the study, which had received Local Research Ethics Committee approval (reference number 03/00/038). Longitudinal MRI was carried out using a 1.5 T Signa Advantage clinical MRI scanner (International General Electric, Milwaukee, WI, USA) using a dedicated bilateral breast coil (Machnet BV, Eelde, Netherlands). A standard dosage chemotherapy regime and relative timing of MRI

MR imaging and spectroscopy

All the MRI data at each time point were acquired in a single imaging session with an average length of 1 h. Dynamic contrast-enhanced MRI was carried out first, during which gadopentetate dimeglumine (formerly known as Gd-DTPA; Schering Health Care, Burgess Hill, UK) was injected as a bolus at a dose of 0.1 mmol kg\(^{-1}\) body weight, using a two-dimensional, \(T_1\)-weighted multislice, fast RF spoiled gradient echo (FSPGR) sequence. In order to attain adequate tissue coverage in all cases, imaging was carried out in either the coronal (30 cm field-of-view, FOV) or sagittal (20 cm FOV) planes (256 \(\times\) 128 matrix in all cases) with five, seven or nine slices (4–9 mm thick, 2 mm gaps). Scan timing was that a combination of the quantitative data from the three techniques the relative prognostic utility. The study hypothesis was that a combination of the quantitative data from the three techniques would lead to a synergistic increase in prognostic accuracy and, therefore, provide an accurate, reliable and non-invasive early indication of tumour response to chemotherapy, which could potentially have a positive influence on patient management.

A prospective study was, therefore, undertaken whereby three of these techniques (DCE-MRI, ADC measurement and MRS fat:water ratios) were carried out in the same patients to determine in each technique the relative prognostic utility. The study hypothesis was that a combination of the quantitative data from the three techniques would lead to a synergistic increase in prognostic accuracy and, therefore, provide an accurate, reliable and non-invasive early indication of tumour response to chemotherapy, which could potentially have a positive influence on patient management.

A prospective study was, therefore, undertaken whereby three of these techniques (DCE-MRI, ADC measurement and MRS fat:water ratios) were carried out in the same patients to determine in each technique the relative prognostic utility. The study hypothesis was that a combination of the quantitative data from the three techniques would lead to a synergistic increase in prognostic accuracy and, therefore, provide an accurate, reliable and non-invasive early indication of tumour response to chemotherapy, which could potentially have a positive influence on patient management.

A prospective study was, therefore, undertaken whereby three of these techniques (DCE-MRI, ADC measurement and MRS fat:water ratios) were carried out in the same patients to determine in each technique the relative prognostic utility. The study hypothesis was that a combination of the quantitative data from the three techniques would lead to a synergistic increase in prognostic accuracy and, therefore, provide an accurate, reliable and non-invasive early indication of tumour response to chemotherapy, which could potentially have a positive influence on patient management.

A prospective study was, therefore, undertaken whereby three of these techniques (DCE-MRI, ADC measurement and MRS fat:water ratios) were carried out in the same patients to determine in each technique the relative prognostic utility. The study hypothesis was that a combination of the quantitative data from the three techniques would lead to a synergistic increase in prognostic accuracy and, therefore, provide an accurate, reliable and non-invasive early indication of tumour response to chemotherapy, which could potentially have a positive influence on patient management.

A prospective study was, therefore, undertaken whereby three of these techniques (DCE-MRI, ADC measurement and MRS fat:water ratios) were carried out in the same patients to determine in each technique the relative prognostic utility. The study hypothesis was that a combination of the quantitative data from the three techniques would lead to a synergistic increase in prognostic accuracy and, therefore, provide an accurate, reliable and non-invasive early indication of tumour response to chemotherapy, which could potentially have a positive influence on patient management.
following diffusion weightings: 0, 14, 55, 125, 222, 346, 499 and 680 s mm$^{-2}$. The ADC values were calculated on a whole ROI basis (Gibbs et al, 2001), which included a correction for the bias inherent in magnitude image reconstruction at low signal-to-noise ratios (especially true for higher $b$-values) (Miller and Joseph, 1993).

The relative proportions of spectroscopic signal arising from tissue water and lipids were measured at echo times of 30 and 135 ms using a one-dimensional stimulated echo (STEAM) spectroscopic imaging sequence (without chemical shift selective suppression of any moieties) and a repetition time of 3 s. Thirty-two phase-encode steps were used over a 16 cm superior–inferior FOV with STEAM excitation limited to seven adjacent voxels (1.0 × 0.5 cm in the anterior–posterior and left–right directions, respectively, giving a nominal voxel volume of 0.25 ml). Figure 1 shows the seven spectroscopic voxels in a representative tumour before chemotherapy. To ensure that the same region of tissue was examined in follow-up studies, MRS voxels were relocated, by an experienced radiologist, after careful comparison of the breast architecture surrounding the lesion with the architecture demonstrated on hard copies of previous voxel locations. Changes in tumour choline were not measured, as the predictive utility of the technique had not been established in the literature at the time of study design.

Spectral analysis was carried out using the SAGE IDL package (International General Electric, Milwaukee, WI, USA). In order to decrease the degree of inter-voxel signal contamination, mild spatial apodisation was applied (Fermi filter: diameter/width = 90%/5%), which resulted in an effective voxel volume of 0.33 ml (Jackson et al, 1994). To minimise partial volume contamination from adipose tissues outside the tumours, only signals from those voxels contained wholly within the tumour were averaged; then, water and fat signal intensities were measured by frequency-domain fitting of the water peak (4.7 p.p.m.) and the dominant lipid peaks at 1.3 and 0.9 p.p.m. (representing the methylene –$(\text{CH}_2)$– and terminal methyl –CH$_3$ lipid moieties, respectively). Spectroscopic water content (%$W_{\text{MRS}}$) was then quantified using the ratio of the water signal to that of the sum of water and the two dominant fat signals. Water $T_2$ was also estimated using a crude two-point method: $T_2 \approx (\text{TE}_{135} − \text{TE}_{135})/\ln(\text{area}_{30}/\text{area}_{135})$.

Biopsy data

Patients’ notes were reviewed to collate data obtained from histopathological analysis of specimens acquired via pre-chemotherapy core biopsy and final surgery. Parameters of interest were histological tumour subtype, tumour grade (pre-chemotherapy), the presence or absence of ductal carcinoma in situ (DCIS) and pre-chemotherapy oestrogen and progesterone receptor status (ERS and PRS) scores. Oestrogen and progesterone receptor status scores were recorded in the notes on either a six- or eight-point scale and converted to continuous data by calculating the ratio of the score to maximum possible score. The maximum diameter of the resected tumour, as calculated by the pathologist, was also collated and compared to the maximum diameter recorded in the final clinical MRI report using the Limits of Agreement method proposed by Bland and Altman (1986).

Quantifying and correlating parameter changes and tumour response

Changes in parameter values between TP$_0$ and TP$_2$ were calculated either as absolute differences, $D_{\text{TP}}(X) = X_{\text{TP}} − X_{\text{TP}_0}$, or percentage change, $P_{\text{TP}_2}(X) = X_{\text{TP}}/X_{\text{TP}_0} − 1$, as deemed most appropriate. A change in tumour volume at the end of treatment, $P_{\text{TP}_2}(V)$, exceeding $−65\%$ was taken as the cutoff between partial response (PR) and stable disease (SD), this being equivalent to a decrease in cross-sectional area of $50\%$, itself broadly equivalent to the RECIST criterion of a $50\%$ decrease in the product of maximum orthogonal diameters (Therasse et al, 2000). No attempt was made to predict final tumour size in this study, as this is not likely to alter patient management before surgery.

All statistical analyses were carried out using the SPSS package (SPSS Inc., Chicago, IL, USA). Correlation between $D_{\text{TP}}(X)$, $P_{\text{TP}_2}(X)$ or continuous/ranked pathology data and $P_{\text{TP}_2}(V)$ was assessed using the Spearman non-parametric test (two-tailed). Correlation between dichotomous pathology data (presence/absence of DCIS) and categorical pathology data (cancer type) was assessed using the Mann–Whitney and Kruskal–Wallis non-parametric tests (two-tailed), respectively. Prognostic efficacy was assessed using receiver–operator characteristic (ROC) curves (Altman and Bland, 1994), with PR as the positive result, and the areas contained underneath them (AUC) calculated.

Variables were combined, in the hope of attaining a synergistic increase in prognostic efficacy, using logistic regression analysis (LRA) modelling, a statistical technique that maximises binary classification accuracy using a linear combination of weighted input variables plus a constant term (Hosmer and Lemeshow, 1989). Backwards conditional elimination of input variables ($P > 0.10$) was used in order to prevent over-parameterisation of models, and ordinal variables were treated as categorical with the most benign category used as the indicator reference. Some statistical confidence intervals (CIs) were obtained using tabulated values (Anonymous, 1982).

RESULTS

Patient recruitment, withdrawal and other excluding factors

A total of 46 women were recruited into the study over a period of approximately 2 years; however, six (13%) of these subsequently
withdrew. One case of atypical, inflammatory lobular cancer was excluded from data analysis on the grounds that it is impossible to measure tumour volume accurately in such tumours. One case in which histology only revealed DCIS was also excluded on the grounds that DCIS is too dissimilar to the other tumour types present, and chemotherapy is not the treatment of choice for it (2/40 = 5%).

In three of the remaining 38 cases, the second MRI examination was severely delayed owing to clinical complications and was carried out after the chosen cutoff of 55 days, thus necessitating the exclusion of the TP₂ (but not the TP₀) data. In addition, it was not possible to acquire ADC and MRS data on every occasion because of a number of problems including hardware failure and an inability to acquire sufficient signal to permit sequence optimisation. Such failures occurred during the TP₀ scan, necessitating complete exclusion of the patient from data analysis, in four cases, and at TP₂, necessitating removal of TP₂ data only, in nine cases. A total of 13 failures in the 73 examinations (38 at TP₀ and 35 at TP₂) represent a failure rate of 18%. Results will, therefore, be presented for 22 cases for which both TP₀ and TP₂ data were available and an additional 12 cases for which data were available at TP₀ only (see Figure 2).

Neither tumour volume measurements nor PK analysis were affected at any time point, as these were based on threedimensional and DCE-MRI images acquired before ADC and MRS were attempted. Therefore, tumour volume response, PC₀F(V), could be quantified in all cases. Pharmacokinetic data and tumour volume data calculated in cases where ADC and MRS data were not acquired were voluntarily excluded, so as not to bias statistical power in favour of PK/volume parameters. In the 22 cases for which valid TP₀ data were available, 18 (82%) of the women had received two courses of chemotherapy at the time of MRI and four had received three courses. In three cases, it was decided to cease chemotherapy before a full six cycles were completed, as poor clinical response was being demonstrated. In these cases, MRI was, however, carried out after the final course of treatment, thus permitting a final measurement of tumour volume to be made and, thus, quantifying the relatively poor response adequately.

All five pathology variables (i.e. cancer type, grade, presence/absence of DCIS, and ERS and PRS scores) were available in 35 of the 38 eligible cases. These included 32 of the 34 valid TP₀ MRI cases, all 22 of the valid TP₀ MRI cases and three cases in which MRI data were unusable (see Figure 2). The age at TP₀ for the 37 women whose data are used in this study ranged from 26 years, 6 months to 75 years, 7 months, with a median of 48 years, 6 months.

The pathological cancer subtypes of these 37 cases were 19 ductal/not specified (12 with DCIS), 15 (nine) ductal and three (one) lobular. The number of women with cancers of grade one, two and three were nine, 18 and eight, respectively (cancer grade was not reported by the pathologist in two cases). Two women did not undergo surgery because of the presence of metastatic disease and one woman did not undergo surgery as post-chemotherapy core biopsy confirmed the absence of residual cancer. In the remaining cases, 17 women underwent mastectomy and 17 underwent wide local excision. The length of time between the final MRI scan and surgery ranged from 14 to 117 days, with a median of 28 days. Maximum tumour diameter was recorded in the patient’s notes by both the pathologist and the MRI radiologist in 21 cases.

**Agreement between MRI and pathology measurements**

Calculation of the limits of agreement between the maximum tumour diameters as measured by the pathologists and the MRI radiologist demonstrated a mean difference (MRI – pathology) of −2.1 mm with a 95% CI (1.96 × standard deviation) of ±18.6 mm. The limits of agreement were, therefore, −20.7 to 16.5 mm. A paired t-test of the two sets of measurements showed that the mean difference was not statistically significant (P = 0.32), thus indicating that MRI is an unbiased estimator of true tumour size and, therefore, suitable for quantifying tumour response. The maximum diameter as measured by the MRI radiologist was also compared to the cube root of the final tumour volume, and these two variables demonstrated excellent linear correlation (P = 0.0001; R² = 0.56).

**Predicting response using individual variables**

The ranges of all predictive MRI variables are presented in Table 1, and the results of correlating the predictive MRI and pathology variables with PC₀F(V) are given in Table 2. Six MRI variables demonstrated P-values less than 0.07 and the data for these, divided into SD and PR subsets, are presented in Figure 3. These six variables along with four histopathology variables for which the P-values were also less than 0.07 underwent ROC curve analysis, and the results of this are also given in Table 2. The slight differences in the patterns of statistical significance between the correlation and ROC analyses can be attributed to the different variables quantifying response in the two types of tests (specifically that PC₀F(V) is continuous whereas PR/SD is dichotomous). Four variables demonstrated an ROC AUC significantly larger than 0.5: water T₂ at TP₀ and changes in %WMS at 135 ms, water T₂ and tumour volume between TP₀ and TP₂. The ROC plots for these are shown in Figure 4.

**Table 1** The range of values for all predictive MRI variables used in this study, as grouped by clinical time point (TP₀ being before treatment and TP₂ being after the second course of chemotherapy)

| Parameter (units) | TP₀ (n = 34) | TP₂ (n = 22) |
|-------------------|-------------|-------------|
| Volume (ml)       | 3.7–240.0   | 1.3–95.0    |
| Kᵣ⁰ (per unit time)| 1.4–18.0    | 0.6–8.3     |
| Kₑ (per unit time)| 0.4–12.4    | 0.4–3.9     |
| Vₑ (arbitrary)   | 0.6–3.5     | 1.2–3.1     |
| MEF (arbitrary)   | 0.8–3.5     | 1.3–3.6     |
| ADC (mm²/s)       | 0.3–5.9     | 0.9–4.3     |
| %WMS at 135 ms    | 68–100%     | 30–100%     |
| %WMS at 30 ms     | 73–100%     | 49–100%     |
| Water T₂ (ms)     | 47–130      | 24–104      |

ADC = apparent diffusion coefficient; MEF = maximum enhancement factor; MRI = magnetic resonance imaging.
Table 2  Statistical significance of non-parametric correlation analyses between the variables indicated and final tumour volume response \( PC_{0F}(V) \)

| No. of cases | Correlation P (sense) | ROC curve | |
|--------------|----------------------|-----------|-----|
|              | PR+SD                |           |     |
| MRI data at TP0 |          |           |     |
| \%W_MRS at 135 ms | 24+10 | S: 0.034 (–) | 0.14 | 0.66 |
| Water \( T_2 \) | 24+10 | S: 0.055 (–) | 0.03 | 0.73 |
| Tumour volume | 24+10 | S: 0.059 (–) | 0.20 | 0.64 |
| \%W_MRS at 30 ms | 24+10 | S: 0.579 | 0.56 to 0.91 |
| \( K_{trans} \) | 24+10 | 0.686 | 0.56 to 0.91 |
| MEF | 24+10 | S: 0.720 | 0.56 to 0.91 |
| Water ADC | 24+10 | S: 0.808 | 0.56 to 0.91 |

| MRI data at TP2 |          |           |     |
| \( PC_{0F}(\text{tumour volume}) \) [2] | 16+6 | S: 0.001 (+) | 0.03 | 0.80 |
| \( \Delta D_0(\text{water } T_2) \) [3] | 16+6 | S: 0.006 (+) | 0.04 | 0.79 |
| \( PC_{0F}(\%W_MRS \text{ at 135 ms}) \) [4] | 16+6 | S: 0.025 (+) | 0.02 | 0.83 |
| \( \Delta D_0(\%W_MRS \text{ at 30 ms}) \) | 16+6 | S: 0.055 | 0.03 | 0.73 |
| \( K_{ep} \) | 16+6 | S: 0.05 | 0.03 | 0.73 |
| \( MEF \) | 16+6 | S: 0.057 | 0.03 | 0.73 |
| \( \Delta D_0(\text{water ADC}) \) | 16+6 | S: 0.0736 | 0.03 | 0.73 |
| Other data |          |           |     |
| Age at TP0 | 26+11 | S: 0.958 | 0.08 | 0.69 |
| Cancer grade [5] | 24+11 | S: 0.003 (–) | 0.14 | 0.79 |
| Presence of DCIS | 24+11 | M: 0.006 (+) | 0.09 | 0.68 |
| PRS score [6] | 24+11 | S: 0.008 (+) | 0.17 | 0.65 |
| ERS score | 24+11 | S: 0.061 (+) | 0.17 | 0.65 |
| Cancer type | 24+11 | K: 0.740 | 0.17 | 0.65 |

LRA models (variables included)

|                               | MRI only [1,2,4] | MRI and histopathology [1 to 6] |
|-------------------------------|-----------------|------------------------------|
| No. of cases | 16+6 | 16+6 | 16+6 | 16+6 |
| Correlation P (sense) | 0.0032 | 0.0004 | 0.0032 | 0.0004 |
| AUC | 0.92 | 1.00 | 0.92 | 1.00 |
| 95% CI | 0.79 to 1.00 | 1.00 to 1.00 | 0.79 to 1.00 | 1.00 to 1.00 |

The sense of the correlations is also shown. Results of ROC curve analyses (with PR/SD as positive/negative results) are also shown for selected variables. Numbers in square brackets indicate those variables chosen as inputs to LRA modelling. S = Spearman rank correlation; M = Mann–Whitney test; K = Kruskal–Wallis test; (+) a positive correlation with lower, or more negative values being associated with more negative values of \( PC_{0F}(V) \), that is, PR; (–) a negative correlation with higher, or more positive values being associated with more negative values of \( PC_{0F}(V) \), that is, PR. AUC = area under the curve; ADC = apparent diffusion coefficient; DCIS = ductal carcinoma in situ; ERS = oestrogen receptor status; LRA = logistic regression analysis; MEF = maximum enhancement factor; MRI = magnetic resonance imaging; PR = partial response; PRS = progesterone receptor status; ROC = receiver–operator characteristic; SD = stable disease.

Figure 3  Dot plots showing the PR and SD data for six MRI variables that when correlated with a final change in tumour volume demonstrated \( P \)-values less than 0.07. The medians for both SD and PR subsets are also shown.
When specificity is set at 100% in Figure 4, the sensitivity for the change in water $T_2$ at TP2 is 69% (11/16; 95% CI = 41–89%), which suggests that this fraction of PR cases can be accurately predicted, that is, at 100% positive predictive value (PPV). Similarly, when sensitivity is set to 100%, the specificity displayed by the change in $W_{MRS}$ at TP2 is 50% (3/6; 95% CI = 12–88%), which suggests that this fraction of SD cases can be accurately predicted, that is, at 100% negative predictive value (NPV) by this variable. The results of Table 2 also suggest that the dominant factor in driving the observed changes in $W_{MRS}$ at 135 ms ($P = 0.025$) is the change in water $T_2$ ($P = 0.006$) as opposed to changes in the relative water and lipid concentrations, which would be reflected by changes in $W_{MRS}$ at 30 ms ($P = 0.788$) for which the effects of $T_2$ would be much less.

**Predicting response using a combination of variables**

All variables demonstrating an ROC $P$-value less than 0.10 (Table 2) were presented as inputs to LRA modelling. Two separate analyses were conducted: one with MRI variables only (water $T_2$ at TP0, PC20(tumour volume), $D_{20}$(water $T_2$) and PC20($W_{MRS}$ at 135 ms)) and one with both MRI and pathology data (PRS score and grade). The number of cases that could be included in LRA modelling was limited to those 22 for which all of the above variables were available (16 PR and six SD). The logits for the two final models were as follows:

Logit 1 = $114.9 \times [T_2 \text{ at TP}_0] - 10.1 \times PC_{20}(\text{tumour volume})$

- $10.9 \times PC_{20}(\% W \text{ at 135ms}) - 11.1$

Logit 2 = $19709.5 \times [T_2 \text{ at TP}_0]$

- $1890.1 \times PC_{20}(\text{tumour volume})$

- $1877.9 \times PC_{20}(\% W \text{ at 135ms})$

+ $6330.0 \times D_{20}$(water $T_2$)

- $785.6 \times [PRS \text{ score}] - 310.4[IF\ grade\ 1]$

- $111.0[IF\ grade\ 3] - 1245.6$

The ROC AUC values for the above logits are shown in Table 2. Both models provide an increase in prognostic accuracy over all of the individual variables, and the second model is able to classify all 22 cases accurately. Both PPV and NPV are 100% for this data set and the 95% CI lower bounds are 79% (16/16) and 54% (6/6), respectively.

**DISCUSSION**

To the authors’ knowledge, this paper represents the first study whereby PK parameter mapping, ADC mapping and unsuppressed proton spectroscopy (water:fat ratios) have all been investigated in the same patient examinations, thus permitting their relative abilities to predict chemotherapy response to be ascertained rigorously. The use of spectroscopic imaging at both short and long echo times represents a significant extension to the methodology of Jagannathan et al (1998): firstly through permitting greater post-examination control over intra-voxel partial volume effects and secondly by permitting the effects of $T_2$ relaxation to be revealed. It was also hoped that the 30 ms TE spectra would permit investigation of the relative proportions of the 1.3 and 0.9 p.p.m. lipid peaks, along with the contributions of other, minor lipid moieties, which are more conspicuous at short TE. This analysis is not possible with quantitative chemical shift imaging (Daniel et al, 1998), as this technique, while having much higher spatial resolution, only provides images for water and total fat signal. It did not prove possible to resolve the separate lipid components adequately in a sufficient number of cases to permit this; therefore, chemical shift imaging may be a suitable alternative to spectroscopic imaging, permitting increased spatial resolution in future studies.

This study shows that MRS parameters provide substantial prognostic information, and slightly more than that provided by volume measurements alone, thus supporting the conclusions of Jagannathan et al (1998). This study also suggests that water $T_2$ plays a dominant part in driving the observed treatment changes in water:fat ratios (at TP2) and it is, therefore, possible that image-based $T_2$ mapping techniques (Liney et al, 1996) might provide the same prognostic information as $W_{MRS}$ but with the combined benefits of greatly increased spatial resolution, wider availability in clinical MRI centres and a reduced likelihood of technical failure in a given examination. Quantitative $T_2$ mapping has been applied to monitoring chemotherapy in one animal tumour study, but the increases in $T_2$ observed 3 days post-therapy were not statistically significant (Duvvuri et al, 2001). Significant differences were observed for $T_{1π}$ however – a parameter not measured in this study because of time and technical constraints. Whether chemotherapy induces cell death via necrosis, apoptosis or a combination of the two, profound changes will certainly occur in the molecular environment of tissue water, leading to changes in its measured $T_2$. This molecular environment is known to be complex and multicompartmental, which precludes a simple rationale for the results of this study.

Although water:fat ratios have been shown to contain substantial prognostic power, in both this and another study (Jagannathan et al, 1998), it is noted that a number of groups are now focusing on measurements of choline-containing compounds using water-suppressed MRS, as the choline moiety is potentially a more sensitive, and probably a more biochemically relevant marker, of cancer cell viability. Such studies have recently been carried out in breast cancer (Jagannathan et al, 2001; Meisamy et al, 2004), lymphoma and germ cell tumours (Schwarz et al, 2002), and glioma (Preul et al, 2000), where positive treatment response has been associated with reduction in the levels of total choline. Similar studies using phosphorous spectroscopy to detect changes in phosphocholine or phosphomonoesters have also been...
carried out in lymphoma (Arias-Mendoza et al, 2004) and head and neck cancer (Shukla-Dave et al, 2002a, b). It is noteworthy that these studies attempted to predict response before treatment commencement, always a more valuable prognostic test, sparing the patients all unnecessary morbidity and time, than reliably detecting early response, which the methods presented herein achieve.

This study also demonstrates that histopathology data routinely available from pre-chemotherapy core biopsies also provide substantial prognostic power (especially so for PRS score and grade). This conclusion is supported by other laboratory studies, which have included investigation of expression of the HER2 oncogene (Penault-Llorca et al, 2003) and the cellular proliferation marker Ki-67 (Aas et al, 2003). Whereas laboratory analysis of actual tissue is inherently superior to imaging alone, it should be noted that the core biopsy technique is always prone to sampling errors, whereas imaging has the advantage of covering the whole of the lesion.

This study also demonstrates that a combination of prognostic variables, via LRA, can be used to provide a synergistic increase in prognostic accuracy, thus proving the study hypothesis. Although the perfect prognostic accuracy achieved using one of the LRA models indicates that the methods used in this study have substantial prognostic power and could be applied clinically, leading to benefits for patients, the true accuracy will be less and will need to be ascertained in larger studies. A multicentre study would also allow investigation of reproducibility. It might also be beneficial to split the data acquired in subsequent studies into separate test and validation subsets, thus allowing optimal model coefficients to be set with the former subset and their true prognostic accuracy determined through application of the model to the completely independent data in the latter.

Early, accurate detection of non-response to chemotherapy would permit an early change to second-line treatment and thus spare patients unnecessary toxicity, cost and delay of initiation of effective treatment. The cessation of treatment that would ultimately prove ineffective might also be beneficial in terms of health economics, allowing resources to be applied more efficiently. Conversely, early, accurate detection of response to chemotherapy might have positive benefits for the psychological well-being and quality-of-life of patients through giving them increased hope. The high prognostic accuracy might be bolstered by including variables that quantify lesion shape and texture, as obtained from the post-contrast three-dimensional images, in the LRA modelling (Esserman et al, 2001).

This study shows that ADC mapping after the second course of chemotherapy does not contribute significantly towards detecting early response. Other studies have shown that treatment-induced changes in ADC in human breast cancer and breast cancer metastasised to liver are most marked a few days after the first dose of chemotherapy (Theilmann et al, 2003). Therefore, our study protocol, with the predictive scans being carried out after two cycles of chemotherapy rather than one (Gibbs et al, 2003), was perhaps not well suited to elucidating the relative merits of ADC mapping.

In contrast to our results, other studies have indicated that PK parameters derived from DCE-MRI contain prognostic information during chemotherapy treatment in a number of locations in the human body (with reductions in $K_{trans}$ being associated with positive response) as reviewed by Padhani (2002) and Knopp et al (2003). This apparent discrepancy regarding the efficacy of PK parameters may be due to the relatively small number of cases included in various studies, or the wide variety of DCE-MRI acquisition and PK analysis techniques used. Intra-tumour vascular heterogeneity might also be significant in some breast cancers (Hayes et al, 2002), leading to the possible breakdown of PK model assumptions such as the fast exchange limit within better-perfused regions (usually located at the tumour rim) (Zhou et al, 2004; Li et al, 2005). This could present a complicating factor in obtaining reliable results, an issue not addressed in this study, which utilised a whole ROI approach so as to minimise the effects of intra-scan patient motion.

Some DCE-MRI studies of breast cancer have attempted to improve the accuracy of PK modelling by imaging in the axial plane so as to allow a plasma concentration time-course to be measured in the aorta, thus permitting the calculation of an arterial input function (AIF) for the tumour. It is debatable, however, whether the benefits of including an AIF measured so distantly from the tissue of interest (i.e. one which does not allow for bolus dispersion effects) outweigh the reduction in spatial resolution and tissue coverage necessitated by axial imaging, especially when the DCE-MRI images are also used for clinical purposes. A consensus on the adoption of standardised techniques and the establishment of large, multicentre studies would, therefore, seem appropriate.

ACKNOWLEDGEMENTS

We gratefully acknowledge the financial support of Yorkshire Cancer Research and also Catherine Bowker, Lynn Cawkwell, John Greenman, John Read and Irvine Long for assistance with study design, patient recruitment, data acquisition and data analysis.

REFERENCES

Aas T, Geisler S, Eide GE, Haugen DF, Varhaug JE, Bassoe AM, Thorsen T, Bernsen H, Borresen-Dale AL, Akslen LA, Lonning PE (2003) Predictive value of tumour cell proliferation in locally advanced breast cancer treated with neoadjuvant chemotherapy. Lancet Oncol 4: 310–317
Altman DG, Bland JM (1994) Diagnostic tests 3 – receiver operating characteristic plots. BMJ 309: 168
Anonymous (1982) Geigy Scientific Tables. Basle: Geigy
Arias-Mendoza F, Smith MR, Brown TR (2004) Predicting treatment response in non-Hodgkin’s lymphoma from the pretreatment tumor content of phosphoethanolamine plus phosphocholine. Acad Radiol 11: 368–376
Bland JM, Altman DG (1994) Diagnostic tests 3 – receiver operating characteristic plots. BMJ 309: 168
Buckley DL, Kerslake RW, Blackband SJ, Horsman A (1997) Scintigraphic imaging of breast cancer: a review. Nucl Med Commun 18: 698–709
Byrne AM, Hill ADK, Skehan SJ, McDermott EW, O’Higgins NJ (2004) Positron emission tomography in the staging and management of breast cancer. Br J Surg 91: 1398–1409
Chu DJZ, Yamanashi WS, Frazer J, Hazlewood CF, Gallager HS, Boddie AW, Martin RG (1987) Proton NMR of human breast tumors – correlation with clinical prognostic parameters. J Surg Oncol 36: 1–4
Dobrindt S, Barchiesi E, Pilotti S, Borasi G (1988) High-resolution H-1 NMR spectroscopy in the diagnosis of breast-cancer. Magn Reson Med 8: 440–449
Buckley DL, Kerslake RW, Blackband SJ, Horsman A (1997) Scintigraphic imaging of breast cancer: a review. Nucl Med Commun 18: 698–709
Byrne AM, Hill ADK, Skehan SJ, McDermott EW, O’Higgins NJ (2004) Positron emission tomography in the staging and management of breast cancer. Br J Surg 91: 1398–1409
Chu DJZ, Yamanashi WS, Frazer J, Hazlewood CF, Gallager HS, Boddie AW, Martin RG (1987) Proton NMR of human breast tumors – correlation with clinical prognostic parameters. J Surg Oncol 36: 1–4

© 2006 Cancer Research UK
British Journal of Cancer (2006) 94, 427 – 435

Molecular Diagnostics
Predicting chemotherapy response with quantitative MRI

DJ Manton et al

Daniel BL, Butts K, Glover GH, Cooper C, Herfkens RJ (1998) Breast cancer: gadolinium-enhanced MR imaging with a 0.5 T open imager and three-point Dixon technique. Radiology 207: 183 – 190

Drew PJ, Kerin MJ, Mahapatra T, Malone C, Monson JRT, Turnbull LW, Fox NJ (2001) Evaluation of response to neoadjuvant chemoradiotherapy for locally advanced breast cancer with dynamic contrast-enhanced MRI of the breast. Eur J Surg Oncol 27: 617 – 620

Duvvuri U, Poptani H, Feldman M, Nadal-desbarats L, Gee MS, Lee WMF, Reddy R, Leigh JS, Glickson JD (2001) Quantitative T1- rho magnetic resonance imaging of RIF-1 tumors in vivo: detection of early response to cyclophosphamide therapy. Cancer Res 61: 7747 – 7753

Esser man L, Hylton N, Yassa L, Barclay J, Frankel S, Sickles E (1999) Utility of magnetic resonance imaging in the management of breast cancer: evidence for improved preoperative staging. J Clin Oncol 17: 110 – 119

Esser man L, Kaplan E, Partridge S, Tripathy D, Rugo H, Park J, Hwang S, Kuerer H, Sudilovsky D, Lu Y, Hylton N (2001) MRI phenotype is associated with response to doxorubicin and cyclophosphamide neoadjuvant chemotherapy in stage III breast cancer. Ann Surg Oncol 8: 549 – 559

Gallons JP, Altbach G, Paine-Murrieta C, Taylor C, Gillies RJ (1999) Early increases in breast tumor xenograft water mobility in response to paclitaxel therapy detected by non-invasive diffusion magnetic resonance imaging. Neoplasia 1: 113 – 117

Gibbs JE, Partridge SC, Hylton NM (2003) Incremental value of ADC as an indicator of treatment response in patients undergoing neoadjuvant chemotherapy for locally advanced breast cancer. Proc Int Soc Magn Reson Med 11: 291

Gibbs P, Tozer DJ, Liney GP, Turnbull LW (2001) Comparison of quantitative T2 mapping and diffusion-weighted imaging in the normal and pathologic prostate. Magn Reson Med 46: 1054 – 1058

Hayes C, Padhani AR, Leach MO (2002) Assessing changes in tumour vascular function using dynamic contrast-enhanced magnetic resonance imaging. NMR Biomed 15: 154 – 163

Hittmair K, Gomiscek G, Langenberger K, Recht M, Imhof H, Kramer J (1994) Method for the quantitative assessment of contrast agent uptake in dynamic contrast-enhanced MRI. Magn Reson Med 31: 567 – 571

Hosmer DW, Lemeshow S (1989) Applied Logistic Regression. Chichester: Wiley

Jackson EF, Doyle TJ, Wolinsky JS, Narayana PA (1994) Short TE H-1 spectroscopic MR-imaging of normal human brain – reproducibility studies. JMRI J Magn Reson Imaging 4: 545 – 551

Jacobs MA, Barker PB, Argani P, Ouwerkerk R, Blujwawa ZM, Blumeka DA (2005) Combined dynamic contrast enhanced breast MR and proton spectroscopic imaging: a feasibility study. J Magn Reson Imaging 21: 141 – 146

Jacobs MA, Barker PB, Bottomley PA, Blujwawa Z, Blumeka DA (2004) Proton magnetic resonance spectroscopic imaging of human breast cancer: a preliminary study. J Magn Reson Imaging 19: 68 – 75

Jagannathan NR, Kumar M, Seenu V, Coschic O, Dwivedi SN, Julka PK, Srivastava A, Rath GK (2001) Evaluation of total choline from in vivo volume localized proton MR spectroscopy and its response to neoadjuvant chemotherapy in locally advanced breast cancer. Br J Cancer 84: 1016 – 1022

Jagannathan NR, Singh M, Govindaraju V, Raghunathan P, Coschic O, Julka PK, Rath GK (1998) Volume localized in vivo proton MR spectroscopy of breast carcinoma: variation of water–fat ratio in patients receiving chemoradiography. NMR Biomed 11: 414 – 422

Katz-Brull R, Lavin PT, Lenkinski RE (2002) Clinical utility of proton magnetic resonance spectroscopy in characterizing breast lesions. J Natl Cancer Inst 94: 1197 – 1203

Kline RP, Wu EX, Petrylak DP, Szabolcs M, Alderson PO, Weisfeil ML, Cannon P, Katz J (2000) Rapid in vivo monitoring of chemotherapy responsive using weighted sodium magnetic resonance imaging. Clin Cancer Res 6: 2146 – 2156

Knopp MV, von Tengg-Kobligk H, Choyke PL (2003) Functional magnetic resonance imaging in oncology for diagnosis and therapy monitoring. Mol Cancer Ther 2: 419 – 426

Kumar R, Alavi A (2004) Fluorodeoxyglucose-PET in the management of breast cancer. Radiol Clin N Am 42: 1113 – 1122

Li X, Huang W, Yankeelov TE, Tudorica A, Rooney WD, Springer CS (2005) Shutter-speed analysis of contrast reagent bolus-tracking data: preliminary observations in benign and malignant breast disease. Magn Reson Med 53: 724 – 729

Liney GP, Knowles AJ, Manton DJ, Turnbull LW, Blackband SJ, Horsman A (1996) Comparison of conventional single echo and multi-echo sequences with a fast spin echo sequence for quantitative T2 mapping: application to the prostate. JMRI J Magn Reson Imaging 6: 603 – 607

Meisamy S, Bolan PJ, Baker EH, Bliss RL, Gulbahce E, Everson LI, Nelson MT, Emory TH, Tuttle TM, Yee D, Garwood M (2004) Neoadjuvant chemotherapy of locally advanced breast cancer: predicting with in vivo H-1 MR spectroscopy – a pilot study. Radiology 233: 424 – 431

Miller AJ, Joseph PM (1993) The use of power images to perform quantitative analysis on low SNR MR images. Magn Reson Imaging 11: 1051 – 1056

Mofatt BA, Chenevert TL, Lawrence TS, Meyer GR, Johnson TD, Dong Q, Tsien C, Mukherji S, Quint DJ, Gebarski SS, Robertson PL, Junc K, Rehemtulla A, Ross BD (2005) Functional diffusion map: a noninvasive MRI biomarker for early stratification of clinical brain tumor response. Proc Natl Acad Sci USA 102: 5524 – 5529

Padhani AR (2002) Dynamic contrast-enhanced MRI in clinical oncology: current status and future directions. J Magn Reson Imaging 16: 407 – 422

Partridge SC, Gibbs JE, Lu Y, Esser man L, Tripathy D, Wolverton DS, Rugo HS, Hwang GE, Ewing CA, Hylton NM (2005) MRI measurements of tumor volume predict response to neoadjuvant chemotherapy and recurrence-free survival. Ann J Roentgenol 184: 1774 – 1781

Penault-llorca F, Cayre A, Mishellany BB, Amat S, Feillel V, Le Boudecg, Ferriere JP, De Latour M, Chollet P (2003) Induction chemotherapy for breast cancer: predictive markers and relation with outcome. Int J Oncol 22: 1319 – 1325

Preul MC, Caramanos Z, Villeneuve JG, Shenouda G, LeBlanc R, Langleben D (2000) Using proton magnetic resonance spectroscopic imaging to predict in vivo the response of recurrent malignant gliomas to tamoxifen chemotherapy. Neurosurgery 46: 306 – 318

Ross BD, Mofatt BA, Lawrence TS, Mukherji SK, Gebarski SS, Quint DJ, Johnson TD, Junck L, Robertson PL, Maruszkio KM, Dong Q, Meyer CR, Bland PH, McConville P, Geng HR, Rehemtulla A, Chenevert TL (2003) Evaluation of cancer therapy using diffusion magnetic resonance imaging. Mol Cancer Ther 2: 581 – 587

Scheplin VD, Ross BD, Chenevert TL, Rehemtulla A, Sharma S, Kumar M, Stojanovska J (2005) Sodium magnetic resonance imaging of chemotherapeutic response in a rat glioma. JMRI J Magn Reson Imaging 21: 288 – 292

Schwarz AJ, Maier NE, Collins DJ, Cunningham D, Huddart R, Leach MO (2002) Early in vivo detection of metabolic response: a pilot study of H-1 MR spectroscopy in extracranial lymphoma and germ cell tumours. Br J Radiol 75: 959 – 966

Shukla-Dave A, Poptani H, Loewner LA, Mancuso A, Serrai H, Rosenthal DI, Kliger AM, Nelson DS, Zakin XL, Arias-Mendoza F, Rijipkema M, Koutcher JA, Brown TR, Herorsch A, Gliskon JD (2002a) Prediction of treatment response of head and neck cancers with P-31 MR spectroscopy from pretreatment relative phosphonoester levels. Acad Radiol 9: 688 – 694

Shukla-Dave A, Poptani H, Loewner LA, Mancuso A, Serrai H, Rosenthal DI, Kliger AM, Nelson DS, Zakin XL, Arias-Mendoza F, Rijikpes M, Koutcher JA, Brown TR, Herorsch A, Gliskon JD (2002b) Prediction of treatment response of head and neck cancers with P-31 MR spectroscopy from pretreatment relative phosphonoester levels (vol 9, pg 688, 2002). Acad Radiol 9: 772

Sijens PE, Wijdeman HK, Moerland MA, Bakker CJG, Vermeulen JWAH, Luyten PK (1998) Human breast cancer in vivo – H-1 and P-31 MR spectroscopy at 1.5 T. Radiology 169: 615 – 620

Stojanovska J, Theilmann RJ, Xia G, Steopock A, Outwater E, Gillies RJ (2003) Correlation of chemotherapeutic response with the apparent diffusion coefficient of water (ADCw) of liver metastases. Proc Int Soc Magn Reson Med 11: 533

Therasse P, Arbbcuk SG, Eisenhauer EA, Wanders J, Kaplan RS, Rubinstein L, Verweij J, Van Glabeke M, van Ooosterom AT, Christian MC, Gwerther SG (2000) New guidelines to evaluate the response to treatment in solid tumors. J Natl Cancer Inst 92: 205 – 216

Tobis PS (1997) Modeling tracer kinetics in dynamic Gd-DTPA MR imaging. JMRI J Magn Reson Imaging 7: 91 – 101

Toms JR (ed) (2003) CancerStats Monograph. London: Cancer Research UK

Wasser K, Klein SK, Fink C, Junkermann H, Sinn HP, Zuna I, Knopp MV, Delorme S (2003) Evaluation of neoadjuvant chemotherapeutic response of breast cancer using dynamic MRI with high temporal resolution. Eur Radiol 13: 80 – 87

© 2006 Cancer Research UK
Weber WA (2005) Use of PET for monitoring cancer therapy and for predicting outcome. J Nucl Med 46: 983–995
Zhao M, Pipe JG, Bonnett J, Evelhoch JL (1996) Early detection of treatment response by diffusion-weighted H-1 NMR spectroscopy in a murine tumour in vivo. Br J Cancer 73: 61–64

Zhou R, Pickup S, Yankeelov TE, Springer CS, Glickson JD (2004) Simultaneous measurement of arterial input function and tumor pharmacokinetics in mice by dynamic contrast enhanced imaging: effects of transcytotic and interstitial exchange. Magn Reson Med 52: 248–257