PERGAMON

European Journal of Cancer 38 (2002) 2085–2093

European Journal of Cancer

www.ejconline.com

Magnetic Resonance Spectroscopy of cancer—practicalities of multi-centre trials and early results in non-Hodgkin's lymphoma

J.R. Griffiths, A.R. Tate, F.A. Howe, M. Stubbs*as part of The Multi-Institutional Group on MRS Application to Cancer[1]

Cancer Research Campaign Biomedical Magnetic Resonance Research Group, Department of Biochemistry and Immunology, St. George's Hospital Medical School, Cranmer Terrace, London SW17 ORE, UK

Received 20 September 2001; accepted 10 December 2001

Abstract

This review describes problems and solutions encountered in large scale multicentre trials of Magnetic Resonance Methods for monitoring cancer. It is illustrated with reference to the Multi-Institutional Group on Magnetic Resonance Spectroscopy (MRS) Applications to Cancer which was set up to perform a trial of ^{31}P MRS for monitoring non-invasively chemotherapy of solid tumours. ^{31}P MR spectra of non-Hodgkin's lymphoma (NHL) pre- and posttreatment, across nine Institutions, were acquired on either General Electric (GE) or Siemens 1.5T Clinical MR instruments. Development of the trial protocol, design of the Radio Frequency (RF) coils and Quality Control procedures necessary to ensure that the datasets acquired at each centre were comparable, are described. The data revealed that phosphomonoesters (PME)/nucleotide triphosphates (NTP) ratio decreased significantly after treatment in the Complete ($P<0.001$) and Partial ($P<0.05$) Responders but not in the Non-Responders ($P>0.1$). In addition, the PME/NTP ratio in the pre-treatment spectra correlated with the subsequent outcome of treatment indicating that PME/NTP levels are significant predictors of long-term clinical response and time-to-treatment failure in NHL.

Keywords: Magnetic Resonance Spectroscopy; ^{31}P MRS, Non-Hodgkin's lymphoma; Clinical trial; Cancer; Treatment response; Phosphomonoesters; NTP; Quality control; RF coils; Decoupling

1. Introduction

Magnetic Resonance Spectroscopy (MRS) is a promising non-invasive technology that allows *in vivo* examination of many cancers. It can be used to measure various parameters including (i) grading and staging of (brain) tumours (ii) pharmacodynamic and (when a drug contains MRS-detectable nuclei such as ^{19}F) pharmacokinetic endpoints, in trials of new oncology drugs and (iii) early response to treatment (see Ref. [1] for review). Most hospitals (in developed countries) now have Magnetic Resonance Imaging (MRI) facilities for routine imaging. However, instruments capable of MRS, (a modality which can be installed on most modern MRI scanners with magnets of 1.5T and above) are usually only found in teaching hospitals where academic research programmes are undertaken. MRS-capable instruments from the main manufacturers now come with automated set-up routines for ^{1}H MRS, the most clinically important nucleus. Once a Volume-of-Interest (VOI) has been marked on a scout image, a non-invasive MR spectrum of the metabolites within a tumour can be taken without further operator intervention. An example of this general approach is seen in Fig. 1 showing the orientation of the patient with a brain tumour lying in the magnet, along with proton images of the brain taken in sagittal and axial orientations and a ^{1}H MR spectrum taken from a localised voxel of the brain tumour.

Whereas MRI gives conventional anatomical images, interpretation of the unique chemical information from MRS may involve complex data analysis algorithms. In spite of this, MRS is at last entering clinical practice: several ^{1}H methods are now in regular use, mainly in

* Corresponding author. Tel.: +44-208-725-5852; fax: +44-208-725-2992.

E-mail address: mstubbs@sghms.ac.uk (M. Stubbs).

[1] See the Appendix for a list of the Institutions.

0959-8049/02/$ - see front matter
PII: S0959-8049(02)00389-1

Neuroradiology [2,3], and promising ^{31}P (4,5) and ^{19}F [6,7] MRS techniques have been widely reported. In this age of evidence-based medicine, formal clinical trials of all these methods are needed, but trials of high-technology methods that are constantly changing present formidable problems [8]. Large numbers of patients with closely-defined conditions have to be studied to obtain sufficient statistical power to verify that the new method is better than a conventional one, a process that usually takes several years. This is clearly difficult for MRS, since the exact parameters for any data-acquisition protocols are very dependent on both the software and hardware of the MR system, and research MR instruments are constantly being upgraded. A trial that takes more than a year or two is likely to present its performers with a dilemma: should they freeze the technology of their instrument for an indefinite period or should they accept that the later patients will be scanned using different MR parameters? Even quite subtle changes can make a considerable difference: a sharper slice selection profile could significantly reduce contaminating signals from outside the VOI, for instance. Improved gradient coils could set in train numerous differences in performance, especially as the instrument's manufacturers are likely to provide new pulse programmes that take advantage of the new hardware, and the old ones may no longer work. Many instruments are shared with other workers, so the interests of the clinical trial may not be the first priority in deciding upgrade policy.

In the field of cancer, there is another dimension of complexity. Very few centres see enough cases of a single type of cancer to perform a clinical trial of MRS in a reasonable time. Multicentre trials are therefore needed to accrue sufficient patients. This is relatively simple if, as in some recent multicentre studies using ^{1}H MRS in the brain, all the data were acquired from a single manufacturer's instruments [9–11]. However, in the majority of cases, the instruments used in the collaborating centres are unlikely to be made by a single manufacturer, and even those that are may be different models. Clinicians in each of the collaborating centres are unlikely to use identical treatment protocols for the same cancer and in a trial lasting 5 years, it is unlikely that any of the cancer treatments used at the beginning of the trial would be the same as the treatments offered at the end.

This review is based on our experience of a multicentre trial that has been grappling successfully with these and other problems for more than 6 years. The general solutions it has reached are offered as a starting point for others.

2. The Multi-Institutional Group on MRS Applications to Cancer

The Multi-Institutional Group on MRS Applications to Cancer was set up to perform a trial of ^{31}P MRS for

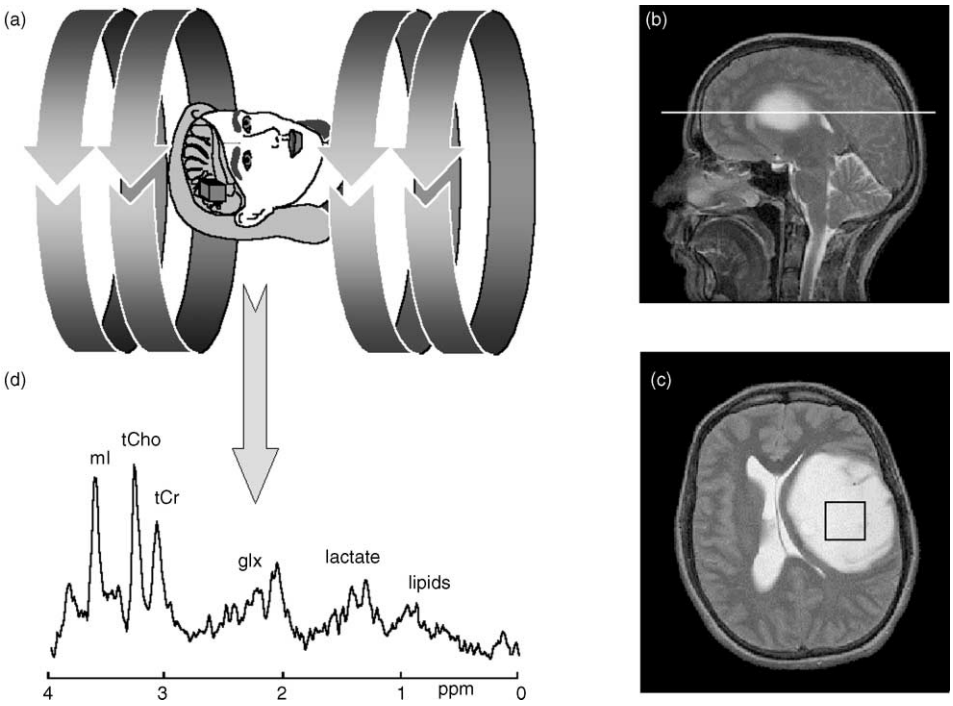

Fig. 1. A schematic diagram (a) showing the orientation of the patient lying in the magnet with proton images of the brain taken in the (b) sagittal (c) axial orientation and a ^{1}H MR spectrum (d) taken from a localised voxel (as defined in (c)) of a Grade II astrocytoma showing resonances from the different metabolites—myoinositol (mI), total choline (tCho), total creatine (tCr), glutamine + glutamate (glx), lactate and lipids.

monitoring chemotherapy (in the widest sense) of solid tumours. ^{31}P MRS is less widely used for patient studies than ^{1}H MRS but it is much simpler to implement in regions of the body outside the brain, and it detects a number of important substances. Several fairly small studies had demonstrated that a wide variety of human cancers show a fall in the phosphomonoesters (PME) peak of the ^{31}P spectrum in the early stages of successful therapy. These data, collected in a meta-analysis by the late Dr Bill Negendank [4], suggested that monitoring tumours by ^{31}P MRS could be used to fine-tune chemotherapy for individual patients, thereby permitting unsuccessful treatments to be halted and new ones tried. The outcome would be improved response, lower toxicity and cost benefit. There was also a chance that ^{31}P MR spectra of the tumours before treatment might contain information that would predict subsequent tumour response to treatment.

The trial is funded by the National Cancer Institute under the leadership of Dr Truman Brown, Director of MR Research in the Department of Radiology and Biomedical Engineering at Columbia University, New York (formerly of Fox Chase Cancer Center (FCC), Philadelphia). Five other hospitals in the USA were involved (Memorial Sloan Kettering (MSK), Dr Jason Koutcher; The Hospital of the University of Pennsylvania, Dr Jerry Glickson (HUP); Wayne State University (WSU), Dr Jeff Evelhoch; Duke University, Dr Cecil Charles; University College of San Francisco, Dr Sarah Nelson; along with two in Britain (St. George's Hospital, Prof John Griffiths (SGH) and the Royal Marsden Hospital (RMH), Prof Martin Leach) and one in The Netherlands (NIJ) (Nijmegen, Dr Arend Heerschap).

3. Choice of tumours

The basic requirements were: (i) That each tumour type studied should be sufficiently common that a reasonable number could be accrued in a 5-year trial; furthermore, common tumours are, of course, clinically more significant. (ii) That each tumour type should respond to chemotherapy (at the time when the decision was made this effectively ruled out brain tumours); (iii) That tumours large enough for ^{31}P MRS in 1.5T instruments ($\geqslant$10 cm^3) should commonly present at superficial sites where current ^{31}P MRS protocols can be used relatively easily. This last requirement ruled out tumours in the bowels, lungs and kidney, all of which were too difficult for routine ^{31}P MRS (i.e. suffering from both motion and susceptibility artefacts and being deep within the body the signal/noise ratio would be too low). The four types finally selected were breast, non-Hodgkin's lymphoma (NHL), sarcomas and head and neck tumours.

4. Development of MRS Protocol

4.1. MR instruments

Three of the collaborating institutions use General Electric (GE) Signa instruments and four have Siemens instruments; all are 1.5T. It was decided early on that it would be impossible to stay with a single level of hardware or software, or to co-ordinate upgrades. Instead, careful quality-control procedures would be performed, to ensure that the basic requirements of the study were always met.

4.2. Design and manufacture of RF coils

Standardised Radio Frequency (RF) coils were built and each hospital was supplied with a set of various sizes, depending on the tumours to be studied. Partly flexible designs were generally favoured over rigid ones to allow the coils to fit snugly to the patients. The standard coil design, scaled to suit tumours in different regions of interest, was a surface coil designed at Fox Chase Cancer Centre by Dr J Murphy-Boesch (see Ref. [12]). It consisted of an 8-cm diameter ^{31}P surface coil surrounded by a 12 by 30 cm ^{1}H butterfly coil with flexible wings which could be used for imaging and proton decoupling (for explanation of this term see below). This coil design is ideal for measurements on larger, superficial tumours such as breast tumours, NHLs and sarcomas. For studies of head and neck tumours and some NHLs, the design was scaled down to give a 5-cm ^{31}P surface coil and corresponding smaller ^{1}H butterfly [13].

4.3. Pulses and fiducial markers

Surface coils give good signal/noise and adiabatic pulse protocols (designed to give uniform signal excitation in tissues at various depths) are needed to optimise signal acquisition. Initial studies were performed with the BIR4 pulse [14], but it was found to give poor results off-resonance—i.e. ^{31}P signals far from the centre frequency, at the extremes of the ^{31}P spectrum (e.g. the PME and β-ATP resonances), were poor. On the advice of Dr Mike Garwood, the BIR-4 phase-cycled or BIRP pulse [15] was therefore chosen and found to give good off-resonance performance. However, the BIRP pulse does not come as part of the standard package on a clinical instrument and until BIRP pulses were programmed-in, non-adiabatic hard pulses were used.

MRI-visible markers were fixed 3.3 mm below the ^{31}P coil to enable the coil position to be determined relative to the tumour in the localising images. Likewise, fiducial markers were also attached to the ^{1}H coil housing. For flip angle calculation (the flip angle is a major determinant of signal/noise), a ^{31}P reference sample (1.9 M

solution of triphenylphosphate (TPP) doped with copper acetoacetate in chloroform to give a stable, short repetition time (TR) of 1 s and thus allow rapid pulse acquisition) was positioned 8.4 mm above the centre of the coil to give a standard signal for pulse flip angle calibration of hard pulses (note: when eventually BIRP pulses were used, precise power level calibrations were not needed for each patient). The molarity of the TPP reference was calibrated against a commercially-available TPP standard (Isotec, Inc. USA; accuracy $\pm 0.05\%$) and in a second independent laboratory (Galbraith Laboratories, Inc., USA) to allow for absolute quantification of the acquired ^{31}P data. The commercially available TPP standard was not suitable for routine studies due to its long T_1 value.

4.4. Decoupling procedures

The phosphomonoester (PME) resonance in the ^{31}P MR spectra of tumours includes signals from both phosphocholine and phosphoethanolamine, which are poorly resolved at 1.5T. This is mainly due to magnetic coupling, a phenomenon in which the magnetic fields of the ^{31}P nuclei in the molecules interact with resonances of adjacent ^{1}H nuclei, producing a complex, multiple peak resonance. This interaction can be suppressed by decoupling, a method in which a precisely modulated, continuous radiofrequency signal, is applied at the ^{1}H resonance frequency. Dr Brown and colleagues had previously demonstrated that ^{1}H decoupling could resolve the two PME peaks [5] so decoupling was made a part of the basic protocol. Proton decoupling was not a standard option on all instruments at that time, so a simple decoupler was designed (by GE for their instruments), constructed and installed at the sites that needed it.

4.5. Safety procedures implemented

The safe performance of *in vivo* MR studies includes compliance with advisory limits on local tissue heating due to RF irradiation. Because high power levels are needed for both decoupling and the BIRP pulse, particular attention had to be paid to ensure that the specific absorption rate (SAR) of the tissue under study was kept below the recommended maximum. To minimise both power requirements and power deposition, the ^{1}H decoupling signal was delivered with a co-planar surface coil consisting of separately tuneable fixed ^{31}P coil with a flexible butterfly component to provide the decoupling pulses. The elaborate simulations required to ensure that SAR standards were not breached at any point within the field of the RF coil are reported in Ref. [16].

4.6. Method of localisation (i.e. selecting the VOI)

There are two main pulse sequences available for this in ^{31}P MRS. Image selected *in vivo* spectroscopy (ISIS) [17], is a simple approach for obtaining a spectrum from a single voxel, but it has an inherent chemical shift artefact: the peaks at each point in the spectrum are obtained from slightly different volumes of tissue, each volume shifted diagonally by an amount proportional to the chemical shift of the resonance. Chemical Shift Imaging (CSI) [18], using a non-selective pulse to excite the whole sensitive volume of the coil has no chemical shift artefact and also allows multiple spectra to be obtained simultaneously from a larger VOI. The result of this is a grid of spectra from individual adjacent voxels within the VOI (see Fig. 2). This grid can be retrospectively adjusted by computational means, so as to position one or more voxels precisely over the tumour.

However, the definition of the localised volume of each voxel is not as accurately defined as indicated by the grid in Fig. 2, due to the point spread function of the CSI technique [12] and there is some contribution of signal from tissue outside the ideal voxel. This problem is exacerbated by patient movement, and two studies published by members of this Multi-Institutional Group on MRS Application to Cancer [19,20] have addressed these issues and evaluated the effects of respiratory motion on the extent of the signal contribution from neighbouring voxels.

5. Quality control

Having standardised the hardware and software to be used, and bearing in mind that the technology was liable to change during the trial, it was necessary to develop reliable Quality Control (QC) procedures to ensure that the studies would be performed safely and consistently over a period of several years and to demonstrate that comparable studies were always performed at each of the participating sites.

5.1. QC of BIRP pulses

All sites established the voltage/power requirements and the off-resonance behaviour across the ^{31}P spectral range of interest for the BIRP pulses on their system, and checked this on a regular basis. This is for comparison with other sites and to ensure that performance is not degraded throughout the course of the trial.

5.2. QC of quantification samples and CSI QC measurements

Once the coils had been designed, manufactured, tested and distributed, a trial was performed on the accuracy of the quantitation at the different sites. All sites were issued with a 1.5-ml sealed glass bulb containing a solution of 1 mM phosphoric acid (PA) in water, doped with 7 mM $NiCl_2$ to give a relaxation time of <0.2 s

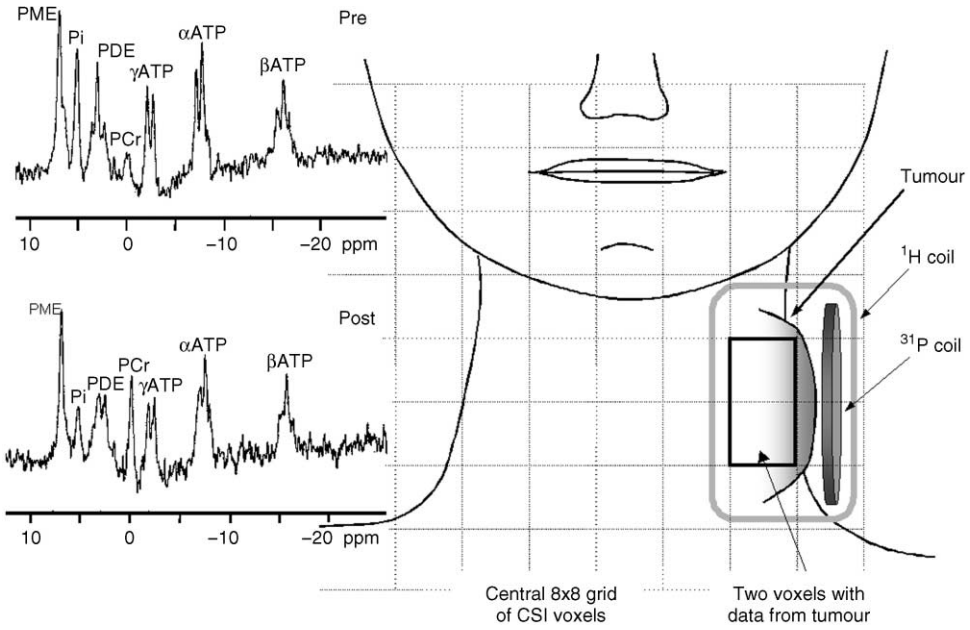

Fig. 2. A schematic diagram of a patient with a non-Hodgkin's lymphoma (NHL) in the neck showing the geometry of the double coil set-up with the fixed ^{31}P coil for acquiring the ^{31}P data and the ^{1}H butterfly coil with flexible wings, which could be used for imaging and proton decoupling. Pre- and posttreatment spectra, each obtained by summing the data from the two CSI voxels indicated are shown on the left. Resonances in the spectra were identified as PME, phosphomonoesters; P_i, inorganic phosphate; PDE, phosphodiesters; PCr, phosphocreatine and three peaks of ATP, α, β and γ. ppm, parts per million; CSI, Chemical Shift Imaging.

(all these phantoms were prepared at one site and the absolute PA concentration in each bulb was accurately known). In preliminary tests, this solution was found to be stable for more than 2 months. When these samples were tested in one institution, the RMS error versus the actual amount of PA in each sample was 2.8%. When one of these PA samples was sent to each institution and the individual sites quantified the PA signal using the TPP standard in their own coil as a reference, the multi-institutional RMS error was determined to be 11.3% ($n=6$) [21]. Quantitation of the PA phantom was performed with both non-localised and localised protocols, to ensure that the CSI datasets acquired at each centre were comparable.

5.3. QC of decoupling performance

All sites were issued with a 2-ml sealed glass bulb containing neat trimethyl phosphate (TMP) to test decoupler performance. These samples had again all been prepared at one specified site. This sample was positioned relative to the coil in an identical fashion to the PA protocol. Forward and reflected power values at the decoupling amplifier output were recorded at the power level necessary for the TMP multiplet to collapse into a singlet. Spectra were acquired at a series of 'off-resonance' decoupler frequencies to ensure that adequate decoupling was obtained across the ^{31}P spectral range of interest. The results of these qualitative QC tests were compared across institutions and found to be satisfactory.

6. Finalised trial protocol

As soon as the protocol and quality control measures were in place the accrual of patients commenced. The finalised trial protocol involved the use of scout images to locate the tumour, followed by standardised BIRP (when eventually implemented) or hard pulses to acquire 3D ^{1}H-decoupled ^{31}P CSI data. Pretreatment scans were first acquired, followed by posttreatment scans at chosen intervals designed to monitor response to treatment in each tumour group. All the sites had obtained ethics approval from the institutional review boards (IRB) or corresponding institutional bodies for the treatment of human subjects.

7. Data transfer

Once data accrual started, transferring large data files of spectra between sites was necessary to permit QC, database development and pattern recognition studies. This turned out to be relatively simple when the files were being transferred between sites with MR instruments

from the same manufacturer. However, to enable transfer from *different* manufacturers' instruments to a common database, initial postprocessing was required to provide a single spectrum from each tumour, pre- and posttreatment. This involved voxel-shifting of the CSI data and summing spectra over a chosen VOI within the tumour. The spectra were then transferred in the VARPRO format [22], a post-processing package for quantifying MR spectra obtained *in vivo* which can convert the different datafile formats of the major manufacturers into a standard format.

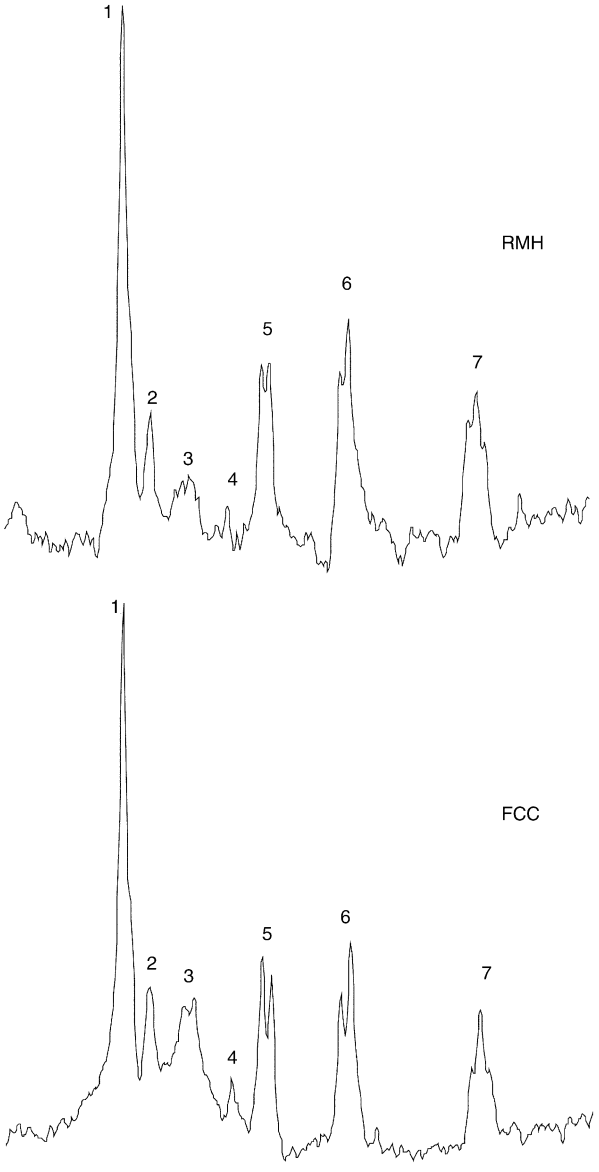

Fig. 3. Mean phased real spectra from non-Hodgkin's lymphoma (NHL) in patients scanned at Royal Marsden Hospital (RMH) ($n=9$) and Fox Chase Cancer Centre (FCC) ($n=12$). Resonances in the spectra were identified as (1) PME, (2) P_i, (3) PDE, (4) PCr, (5) γATP, (6) α ATP and (7) β ATP.

8. Preliminary results on patients in the trial

Due to much improved access to mammography since this trial was initiated, primary breast adenocarcinomas of a size ($>2\times2\times2$ cm) large enough for good signal/noise at 1.5T, are increasingly rare. Consequently, this tumour type was dropped from the patient accrual early on. However, good signal/noise was achieved with NHL, head and neck tumours and sarcomas. NHL spectra were the easiest to accrue, and the results reported below are confined to this tumour type. Fig. 2 shows a schematic of the CSI grid from which the spectra were obtained from a patient with NHL (at St George's Hospital). The appearance of the PCr peak in the posttreatment spectrum is largely due to the significant tumour regression on treatment which means that the tumour spectrum has contributions from other tissues, including muscle.

A fundamental question for trials of this kind is whether data accrued on instruments made by different manufacturers and, inevitably (because of hardware and software constraints), using slightly different data acquisition protocols, can be analysed in a single database. This question was addressed by performing a preliminary pattern recognition study on a small group of spectra obtained from 25 patients at three different sites.

The pretreatment spectra of NHL tumours from the sites look remarkably similar in many ways- the average spectra from FCC ($n=12$) and RMH ($n=9$) are seen in Fig. 3 and a representative one from SGH in seen in Fig. 2. Pattern recognition by principal component analysis (first two principal components of 13 peak variables) of the combined pre-treatment datasets from the three sites (Fig. 4) demonstrated an element of clustering by site, probably due to differences in data

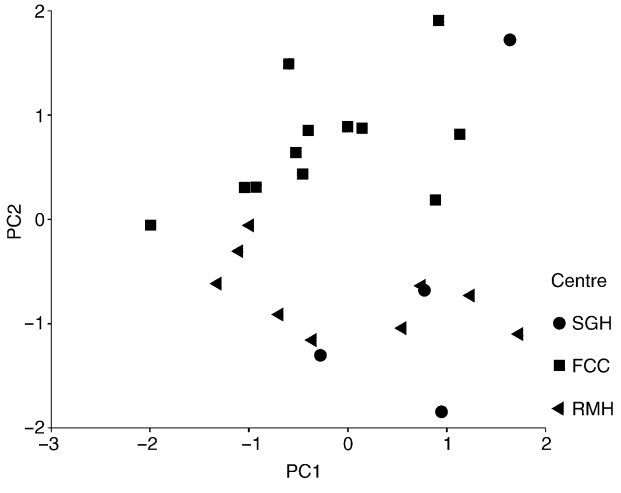

Fig. 4. Principal component analysis of pretreatment spectra from three different institutions St George's Hospital (SGH) ($n=4$), FCC ($n=12$) and RMH ($n=9$) showing some clustering relating to site (Institution). The data plotted are from the first two principal components of 13 spectral variables of the data.

acquisition protocols from two Siemens sites (RMH and FCC) and one GE site (SGH). A different degree of roll in the baseline is observed (Figs. 2 and 3) and the magnitude of the phosphodiester (PDE) signal seems to vary somewhat between the sites. The PDE signals are notoriously difficult to resolve in ^{31}P MRS spectra, probably because some of the substances that give rise to them are partially immobilised in cell membranes and have short T_2 relaxation rates. Differences in adiabatic pulse duration and the length of the phase encoding gradient used at the three sites could have caused T_2 effects, and thus differences in the PDE region of the spectrum.

In spite of these differences, it was still possible to perform a correlation analysis on the combined data set. The pretreatment spectra were differentiated into classes of response (complete responders (CR), partial responders (PR) and non-responders (NR)), and the scatterplot of the summed NTP and PME regions of the pretreatment spectra (from two classes: CR and NR), showed that all but one NR had higher values for the PME region than the CR (Fig. 5). This was an indication that the pre-treatment spectra contained metabolic information that was related to subsequent outcome of treatment. Further analysis of the data obtained so far has been performed using conventional statistics and these are reported in the next section.

NHLs come in a variety of types and grades—from the low-grade indolent through to the high-grade aggressive categories. When the data across all these types were classified into CR, NR and PR, the pretreatment scans showed a positive correlation with the PME/NTP ratio. In a sample of 43 patients from three different sites, a simple mean ± standard error of the mean (SEM) analysis of pre- and posttreatment PME/NTP ratios showed a statistically significant difference between CR and PR and between CR and NR in both pre- and post-treatment spectra, although no significant difference was seen between PR and NR (Table 1). The observation made originally [4,5] that the PME/NTP ratio decreased significantly after treatment in the responding patients was confirmed at $P<0.001$ level of significance for a comparison between the pre- and posttreatment spectra for the CR, $P<0.05$ for the PR and no difference for the NR $P>0.1$ (Table 1). So it appears that in spite of the differences detected by the principal component analysis in the spectra produced from the different institutions, it was still possible to obtain comparable spectra, which when analysed produced consistent and meaningful results.

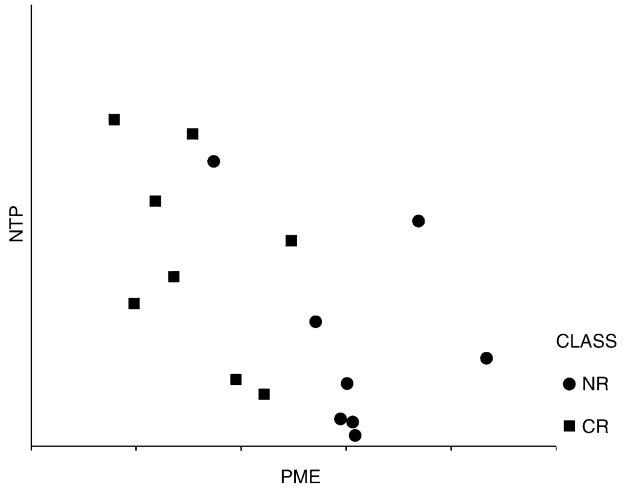

Fig. 5. Scatterplot of the summed nucleotide triphosphate (NTP) and PME regions for pretreatment spectra ($n=16$). The values on the plot are sums of the intensities in the spectra of eight Complete (CR) and eight Non-Responders (NR) from the regions that, in the correlation analysis, differentiated between the classes. Prior to the analysis, the spectra were normalised to the Euclidian norm.

Importantly, as well as confirming the initial hypothesis that PME signals would decline during successful therapy, further conventional statistical analysis by Dr F. Arias-Mendoza [23] has disclosed a very significant association between lower PME, measured prior to treatment, and long-term positive response to that treatment. Interestingly, this result appears to be independent of the type of treatment, although we have not accrued enough cases to truly test this. Our observations, if confirmed in a larger patient population, may significantly influence treatment decisions for individual patients with NHL, and potentially other tumours as well. In addition to the association mentioned, the standard International Prognostic Index (IPI) for NHLs has been integrated with the PME levels to create a combined index. The combined index is a significantly better predictor than the IPI alone, showing a sensitivity of 77% with a specificity of 85% with regard to predicting durable complete response, the only meaningful clinical result in these patients. In addition, the mean time-to-treatment failure as well as the Kaplan–Meier survival curves (Fig. 6) calculated from the combined index are significantly different (38 versus 16 months, $P<0.05$) based on 41 cases in the study [23]. This association is precisely what this project was designed to discover. Specifically, we proposed initially

Table 1
Comparison of pre- and posttreatment PME/NTP ratios in CR, NR and PR responses to treatment assessed by ^{31}P MR spectroscopy

PME/NTP	CR ($n=14$)	PR ($n=13$)	NR ($n=16$)
Pretreatment[a,c]	1.47±0.11	1.88±0.15	2.08±0.18
Posttreatment[b,c]	0.47±0.11	1.30±0.22	2.77±0.72

Means ± standard error of the mean (SEM).

[a] Pre treatment PR versus CR, $P<0.05$, CR versus NR, $P<0.01$, NR versus PR $P>0.1$.

[b] Post-treatment PR versus CR, $P<0.01$, CR versus NR, $P<0.01$, NR versus PR $P>0.08$.

[c] Pre-treatment versus posttreatment CR $P<0.001$, PR $P<0.05$, NR $P>0.1$.

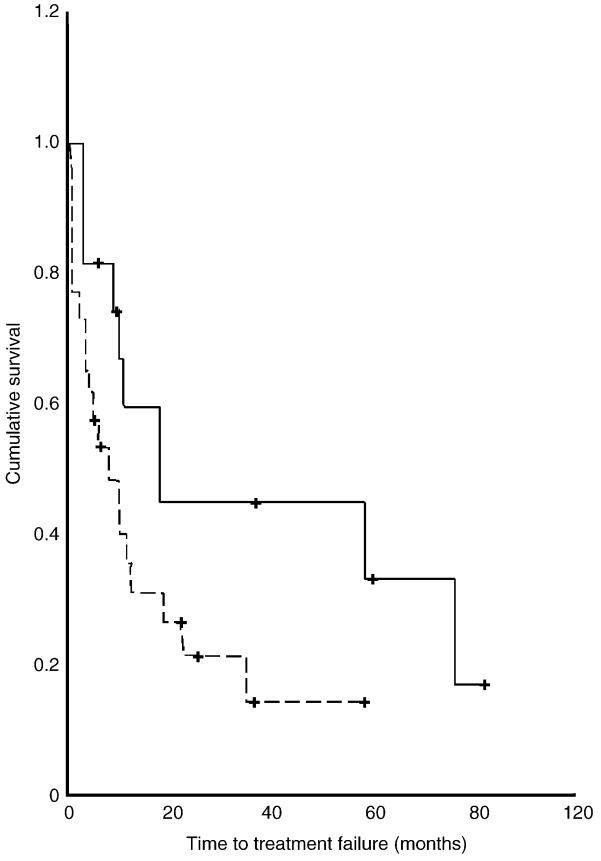

Fig. 6. Kaplan–Meier survival curves of time-to-treatment failure using the standard International Prognostic Index for non-Hodgkin's lymphoma (NHL)'s integrated with the PME/NTP levels obtained from pretreatment spectra to create a combined index. The solid line predicts the outcome from Complete Responders (CR) whereas the dotted line predicts the outcome from the Non-Complete Responders (i.e. Partial Responders (PR) plus Non-Responders (NR)). For further details refer to Arias-Mendoza and colleagues [23].

to develop well-defined technical procedures, suitable for multi-institutional use, to acquire high-quality ^{31}P NMR spectral information from human tumours and, subsequently, to use these techniques to determine if this information contains early indicators of response to treatment—which indeed it does.

To date in this collaborative study, we have successfully scanned, pretreatment, more than 91 patients with NHL of which over 47 have had at least one posttreatment scan and most have had two. We have shown that in spite of the problems of different versions of MR hardware and software, and differences in clinical treatment, not only between hospitals but between countries, multi-institutional clinical trials of MRS methods are feasible. Non-invasive ^{31}P MR scanning suggests that early changes in PME/NTP levels obtained are significant predictors of long-term clinical response and time-to-treatment failure in NHL, clearly important clinical results. At present, these trials are difficult and can only be performed in centres with MR expertise, but with further manufacturer input and improvements in software, clinical trials using MRS should become routinely possible.

Acknowledgements

The group is supported by NCI grants CA62554–CA62561 to the above individual institutions.

Appendix

Members of The Multi-Institutional Group on MRS Application to Cancer are located at the following Institutions: Columbia University, USA; Fox Chase Cancer Center, USA; Duke University, USA; Memorial Sloan-Kettering Cancer Center, USA; The Royal Marsden Hospital, UK ; St. George's Hospital Medical School, UK ; University of California at San Francisco, USA; University Hospital Nijmegen, The Netherlands; University of Pennsylvania, USA; and Wayne State University, USA.

References

1. Howe FA. Magnetic resonance spectroscopy *in vivo*. In *MRI: Principles and Clinical Applications*, Stamford, CT, USA, Appleton and Lange, 1998.
2. Tate AR, Griffiths JR, Martinez-Perez I, *et al*. Towards a method for automated classification of 1H MRS spectra from brain tumours. *NMR Biomed* 1998, **11**, 177–191.
3. Preul MC, Caramanos Z, Collins DL, *et al*. Accurate, non-invasive diagnosis of human brain tumors by using proton magnetic resonance spectroscopy. *Nature Med* 1996, **2**, 323–325.
4. Negendank W. Studies of human tumors by MRS: a review. *NMR Biomed* 1992, **5**, 303–324.
5. Negendank WG, Padavic-Shaller KA, Li C-W, *et al*. Metabolic characterisation of Non-Hodgkin's lymphomas *in vivo* with the use of proton-decoupled phosphorus magnetic resonance spectroscopy. *Cancer Res* 1995, **55**, 3286–3294.
6. Findlay MP, Leach MO. *In vivo* monitoring of fluoropyrimidine metabolites: magnetic resonance spectroscopy in the evaluation of 5-fluorouracil. 1994. *Anti-Cancer Drugs* 1994, **5**, 260–280.
7. Griffiths JR, McIntyre DJO, Howe FA, *et al*. Issues of normal tissue toxicity in patients and animal studies: effect of carbogen ($95\%O_2/5\%CO_2$) breathing in rats after 5FU treatment. *Acta Oncol*. 2001, **40**, 609–614.
8. Leach MO, Arnold D, Brown TR, *et al*. International workshop on standardization in clinical MRS: proceedings and recommendations. *Technol Healthcare* 1994, **2**, 217–234.
9. Negendank WG, Sauter R, Brown TR, *et al*. Proton magnetic resonance spectroscopy in patients with glial tumors: a multicenter study. *J Neurosurg* 1996, **84**, 449–458.
10. Sijens PE, Knopp MV, Brunetti A, *et al*. ^{1}H MR spectroscopy in patients with metastatic brain tumors: a multicentre study. *Magn Reson Med* 1995, **33**, 818–826.
11. Webb PG, Sailasuta N, Kohler SJ, Raidy T, Moats RA, Hurd RE. Automated single-voxel proton MRS: technical development and multisite verification. *Magn Reson Med* 1994, **31**, 365–373.

12. Murphy-Boesch J, Jiang H, Stoyanova R, *et al*. Quantification of phosphorus metabolites from chemical shift imaging spectra with corrections for point spread effects and B-1 inhomogeneity. *Magn Reson Med* 1998, **39**, 429–438.
13. Klomp DWJ, Collins DJ, van den Boogert HJ, *et al*. Radio-frequency probe for ^{1}H decoupled ^{31}P MRS of the head and neck region. *Magn Reson Imaging* 2001, **19**, 755–759.
14. Garwood M, Ke Y. Symmetric pulses to induce arbitrary flip angles with compensation for RF inhomogeneity and resonance offsets. *J Magn Reson* 1991, **94**, 511–525.
15. Bottomley PA, Ouwerkerk R. An improved implementation of low-angle adiabatic (Bir-4) excitation pulses. *J Magn Reson A* 1993, **103**, 242–243.
16. Schwarz AJ, Rijpkema M, Collins DJ, *et al*. SAR and tissue heating with a clinical 31P MRS protocol using surface coils, adiabatic pulses, and proton-decoupling. *Magn Reson Med* 2000, **44**, 692–700.
17. Ordidge RJ, Connelly A, Lohman JAB. Image selected in vivo spectroscopy (ISIS). A new technique for spatially selective NMR spectroscopy. *J Magn Reson* 1986, **66**, 283–294.
18. Brown TR, Kincaid BM, Ugurbil K. NMR chemical shift imaging in three dimensions. *Proc Natl Acad Sci USA* 1982, **79**, 3523–3526.
19. Doyle VL, Howe FA, Griffiths JR. The effect of respiratory motion on CSI localised MRS. *Phys Med Biol* 2000, **45**, 1–12.
20. Schwarz AJ, Leach MO. Implications of respiratory motion for the quantification of 2D MR spectroscopic imaging data in the abdomen. *Phys Med Biol* 2000, **45**, 2105–2116.
21. Arias-Mendoza F, Brown TR, Charles HC, *et al*. Methodological standardization for a multi-institutional in vivo trial of localised 31P MR spectroscopy in human cancer research. *Proc Int Soc Magn Reson Med* 1999, **7**, 1585.
22. Van der Veen JWC, De Beer R, Luyten PR, van Ormondt D. Accurate quantification of *in vivo* ^{31}P NMR signals using the variable projection method and prior knowledge. *Magn Reson Med* 1989, **6**, 92–98.
23. Arias-Mendoza F, Zakian K, Stubbs M, *et al*. Investigation of the predictive vaue of pre-treatment tumor content of phosphoethanolamine and phosphocholine measured by in vivo 31 P MR spectroscopy in non-Hodgkins lymphoma in a multi institutional setting. *Proc Intl Soc Mag Reson Med* 2001, **9**, 274.