Magnetic Resonance in Medicine 56:803–810 (2006)

Optimal Acquisition Schemes for In Vivo Quantitative Magnetization Transfer MRI

Mara Cercignani[1*] and Daniel C. Alexander[2]

This paper uses the theory of Cramer-Rao lower bounds (CRLB) to obtain optimal acquisition schemes for in vivo quantitative magnetization transfer (MT) imaging, although the method is generally applicable to any multiparametric MRI technique. Quantitative MT fits a two-pool model to data collected at different sampling points or settings of amplitude and offset frequency in the MT saturation pulses. Here we use simple objective functions based on the CRLB to optimize sampling strategies for multiple parameters simultaneously, and use simulated annealing to minimize these objective functions with respect to the sampling configuration. Experiments compare optimal schemes derived for quantitative MT in the human white matter (WM) at 1.5T with previously published schemes using both synthetic and human-brain data. Results show large reductions in error of the fitted parameters with the new schemes, which greatly increases the clinical potential of in vivo quantitative MT. Since the sampling-scheme optimization requires specific settings of the MT parameters, we also show that the optimum schemes are robust to these settings within the range of MT parameters observed in the brain. Magn Reson Med 56:803–810, 2006. © 2006 Wiley-Liss, Inc.

Key words: magnetization transfer; Cramer-Rao lower bound; optimal sampling; quantitative MRI; brain

Quantitative MRI techniques (1) map parameters of physical processes across a sample to measure physical properties and highlight spatial differences of materials and tissue. Typically, the methods fit a model of the dependence of the MR signal on the physical process to a number of MRI measurements obtained at different settings of the acquisition pulse sequence, which is sensitized to the physical process of interest. Examples include T_1 and T_2 imaging methods, which map the longitudinal and transverse relaxation constants by fitting exponential decay models in each voxel (e.g., Refs. 2 and 3); diffusion MRI, which typically maps the apparent diffusion coefficient (ADC) (4) or diffusion tensor (5); perfusion imaging, which maps parameters that characterize the hemodynamics (6); and quantitative magnetization transfer (MT) (7), which maps indices of the molecular chemical environment.

In this paper we are concerned primarily with quantitative MT, which is a contrast mechanism based on cross-relaxation and chemical exchange between protons in free water ("liquid" protons) and those bound to macromolecules (8). Protons bound to macromolecules have very short transverse relaxation time and are invisible to MRI. MT-weighted MRI selectively saturates the macromolecular protons by exposing the sample to radiofrequency (RF) energy several kilohertz off resonance from the Larmor frequency to which protons in free water are less sensitive. The exchange mechanism transfers some of this preferential saturation to the liquid spins, where it contributes to the MR signal, thus making some of the properties of macromolecular protons accessible. These properties are interesting biological markers. In particular, evidence suggests that molecules associated with myelin dominate this exchange process in white matter (WM) (9–10). Indices derived from MT-weighted MRI thus reflect the degree of myelination in WM and can highlight demyelination from WM diseases, such as multiple sclerosis (11–13).

Henkelman et al. (7) developed a simple two-pool model of the MT phenomenon in gels exposed to continuous wave irradiation, which was later extended to in vivo pulsed-MT applications by Sled and Pike (14), Yarnyck (15), and Ramani et al. (16). The two-pool model characterizes each pool by its spin density M_0, its longitudinal relaxation rate R_1, and its transverse relaxation time T_2, and assumes that the pools exchange magnetization with rate R. All of these parameters are potentially interesting to measure, and they can be estimated by fitting the model to images acquired with MT pulses of various combinations of the amplitude ω_1 and the offset frequency Δ. Indices of myelination, such as the macromolecular proton fraction f (16) (the macromolecular proton to total proton ratio) and the relative size of the macromolecular pool F (17) (the macromolecular proton to liquid proton ratio), are simple to obtain from the fitted model.

For all quantitative techniques based on model fitting, the precision and accuracy of the parameter estimates depend on the choice of the sampling points (in MT, these are the settings of ω_1 and Δ in the MT-weighted images). In most published MT protocols, the set of sampling points is selected empirically. Sled and Pike (14,17) did not perform a formal analysis, but explored different sampling schemes with 60 sample points. They reported little benefit in using more than two values of ω_1 each with a range of Δ. The choice of points is particularly critical for in vivo imaging where time limits the number of images we can acquire. Ramani et al. (16) acquired 10 images, using three unique settings for ω_1 (two with three distinct settings of Δ, and one with four). Cercignani et al. (18), following Ramani et al. (16), acquired 10 images and, following Sled and Pike, used two settings of ω_1, each with five unique Δ. The central aim of this paper is to explore optimal configurations of sample points for in vivo quantitative MT, where the number of acquisitions is limited by scan time constraints.

[1]NMR Research Unit, Department of Neuroinflammation, Institute of Neurology, London, UK.

[2]Centre for Medical Image Computing, Department of Computer Science, University College London, London, UK.

*Correspondence to: Mara Cercignani, Dept. of Neuroinflammation, 6th Floor, Queen Square House, Institute of Neurology, Queen Square, London WC1N 3BG, UK. E-mail: m.cercignani@ion.ucl.ac.uk

Received 15 March 2006; revised 16 May 2006; accepted 5 June 2006.

DOI 10.1002/mrm.21003

Published online 10 August 2006 in Wiley InterScience (www.interscience.wiley.com).

The issue of determining optimal sampling strategies arises in all quantitative MRI and other model-fitting techniques. Several approaches are available. For simple models in which we can express the parameter of interest directly as a function of measurements with known variance, error propagation provides an estimate of the parameter variance, which we can optimize with respect to the sample points. For example, error propagation provides optimal *b*-values in the basic ADC measurement (19). Numerical simulations also provide insight into the dependence of parameter estimates on sample points, and aid in searching for optimal configurations (e.g., Ref. 20). For more complex models, such as quantitative MT models, the theory of Cramer-Rao lower bounds (CRLB) provides a useful analogue to error propagation. The CRLB (21) sets a lower bound on the variance of any unbiased estimate. In other words, the variance of any parameter estimate obtained by fitting a model to noisy measurements cannot be lower than the CRLB. Although the CRLB is not an estimate of the parameter variance directly, we can use it in a similar way to obtain good sampling configurations by minimizing it with respect to the sample points. In quantitative MRI, this method has been used to optimize sampling configurations for measuring T_2 relaxation time (22), diffusion coefficients (23), and the decomposition of water and fat signals in Dixon imaging (24). In a preliminary study Samson et al. (25) used the CRLB to optimize MT acquisition. They optimized sampling to measure a single parameter of the MT model. They used numerical simulations to show that an optimized scheme with six points can estimate the macromolecular proton fraction as precisely as the 10-point scheme used by Ramani et al. (16). However, their method sets all the other parameters to fixed values, which is unrealistic for heterogeneous samples like the human brain.

All of the above applications of the CRLB optimize the sampling configuration to estimate a single parameter. However, in MT several fitted parameters are of interest. Here we construct simple objective functions based on the CRLB to optimize sampling strategies for multiple parameters simultaneously, and use simulated annealing to search for sampling configurations that minimize these objective functions. We use this approach to obtain optimal acquisition schemes for in vivo quantitative MT imaging, although the method is generally applicable to any multiparametric MRI technique and beyond. Experiments compare optimal schemes derived for quantitative MT in the human WM at 1.5T with previously published schemes (18), using both synthetic and human-brain data. The results show large reductions in error of the fitted parameters (factors of ~3 in standard deviation (SD)) with the new scheme, which greatly increases the clinical potential of in vivo quantitative MT.

THEORY

MT Model

The relative size of the macromolecular pool (17) is

$$F = \frac{M_0{}^B}{M_0{}^A}, \qquad [1]$$

where A labels the liquid pool of spins in free water, and B labels the macromolecular pool of spins so that $M_0{}^A$ and $M_0{}^B$ are the fully relaxed values of magnetization in the two pools, respectively.

We can rewrite Henkelman et al.'s (7) two-pool model in terms of F to obtain

$$S(\omega_1,\Delta) = \frac{S_0\left(R_B\left[\frac{RM_0{}^AF}{R_A}\right] + R_{RFB}(\omega_1,\Delta) + R_B + RM_0{}^A\right)}{\left[\frac{RM_0{}^AF}{R_A}\right](R_B + R_{RFB}(\omega_1,\Delta)) + \left(1 + \left[\frac{\omega_1}{2\pi\Delta}\right]^2\left[\frac{1}{R_AT_2^A}\right]\right)(R_{RFB}(\omega_1,\Delta) + R_B + RM_0{}^A)}, \qquad [2]$$

where $S(\omega_1, \Delta)$ is the MT-weighted MR signal with amplitude ω_1 and offset frequency Δ; S_0 is the signal with no MT weighting; R is the MT exchange rate between the two pools; R_A (= $1/T_1{}^A$) and R_B are the longitudinal relaxation rates of the free and macromolecular pools, respectively; and T_2^A is the transverse relaxation time of the liquid pool. The rate R_{RFB} of loss of longitudinal magnetization owing to the irradiation by the MT pulse depends on ω_1 and Δ, and on the transverse relaxation time of the semisolid pool $T_2{}^B$. Li et al. (26) showed that in brain tissue R_{RFB} is closely modeled by a super-Lorentzian:

$$R_{RFB}(\omega_1,\Delta) = \omega_1^2\sqrt{2\pi} \times \left[T_2^B\int_0^1 \frac{1}{|3u^2-1|}\exp\left(-2\left(\frac{2\pi\Delta T_2^B}{3u^2-1}\right)^2\right)du\right] \qquad [3]$$

and we adopt this model here.

For in vivo applications, when pulsed-MT saturation replaces a continuous RF wave, we use the continuous wave power equivalent (CWPE) for ω_1, which is the precession frequency of a continuous wave with the same mean square amplitude (16). The model in Eq. [2] then contains six independent parameters: $RM_0{}^A$, S_0, R_B, F/R_A,$1/R_AT_2{}^A$, and $T_2{}^B$ (16), which we estimate by fitting the model to six or more measurements with independent ω_1 and Δ. The dependence of $S(\omega_1, \Delta)$ on R_B is weak, which makes fitting this parameter unstable. Since the estimates of the other parameters are largely insensitive to its value, R_B is usually kept fixed at 1 s^{-1} (7,16,17). This reduces the number of fitted parameters to 5. To determine F from the fitted F/R_A, we must estimate R_A. The observed longitudinal relaxation rate of the sample, R_{Aobs} (= $1/T_{1obs}$), can be measured independently, and R_A can be estimated, rearranging Eq. [23] in Ref. 7, from:

$$R_A = R_{Aobs} - \frac{RM_0{}^A F(R_B - R_{Aobs})}{R_B - R_{Aobs} + RM_0{}^A}. \quad [4]$$

CRLB

The theory behind the CRLB is extensively covered elsewhere (e.g., Ref. 21), and only the concepts required for the present application will be reviewed here. Consider a model S with Q unknown parameters $p_1, \ldots, p_Q$ fitted to $N \geq Q$ measurements $A(\mathbf{x}_1), \ldots, A(\mathbf{x}_N)$ acquired with different known settings $\mathbf{x}_1, \ldots, \mathbf{x}_N$. The Cramer-Rao inequality states that $\boldsymbol{\Xi} \geq \mathbf{J}^{-1}$, i.e., that the difference $\boldsymbol{\Xi}\mathbf{J}^{-1}$ is a positive semi-definite matrix, where $\boldsymbol{\Xi}$ is the covariance matrix of an unbiased estimate of $p_1, \ldots, p_Q$, $\mathbf{J}$ is the Fisher information matrix, which has ij^{th} element

$$J_{ij} = \frac{1}{\sigma^2}\sum_{n=1}^{N} \frac{\partial S(p_1,\cdots,p_Q;\mathbf{x}_n)}{\partial p_i} \frac{\partial S(p_1,\cdots,p_Q;\mathbf{x}_n)}{\partial p_j}, \quad [5]$$

and σ is the SD of the noise, which we assume is independent of $\mathbf{x}$. Specifically, the variance of any unbiased estimator of p_i cannot be smaller than the i^{th} diagonal element $[\mathbf{J}^{-1}]_{ii}$ of $\mathbf{J}^{-1}$.

To find a set of sample points $\mathbf{x}_1, \ldots, \mathbf{x}_N$ that minimize the variance $\sigma_i{}^2$ of a particular p_i, the standard approach in the absence of a direct estimate of $\sigma_i{}^2$ is to minimize the lower bound $[\mathbf{J}^{-1}]_{ii}$ instead with respect to $\mathbf{x}_1, \ldots, \mathbf{x}_N$. To maximize the precision of multiple parameter estimates, we must find sample points that minimize several diagonal elements of $\mathbf{J}^{-1}$ simultaneously. To ensure that the relative precision of each parameter is similar, we weight the variance of each parameter by the inverse square of the parameter's value (27) and aim to minimize $\sum_{i=1}^{Q} p_i^{-2}\sigma_i^2$. However, since we do not know the variances, we substitute the CRLB for each parameter to obtain the objective function

$$V = \sum_{i=1}^{Q} p_i^{-2}[\mathbf{J}^{-1}]_{ii} \quad [6]$$

for a typical set of p_i.

We emphasize here that Eq. [6] requires the choice of a particular set of values for $p_1, \ldots, p_Q$. Often, however, we wish to map the parameters over a heterogeneous sample, such as the human brain, in which they assume a range of values. Thus we use an alternative to V that considers a range of settings of the p_i:

$$V_{\max} = \max_k \left\{ \sum_{i=1}^{Q} p_{ik}^{-2}[\mathbf{J}_k^{-1}]_{ii} \right\}, \quad [7]$$

where k ranges over K combinations of the p_i, p_{ik} is the value of p_i in combination k and $\mathbf{J}_k$ is the Fisher information matrix for combination k. Minimizing $V_{\max}$ maximizes the precision of the parameters for the worst case in the K combinations of settings; it is analogous to the standard G-optimality criterion in optimal experiment design (27).

Application to MT

In quantitative MT, the acquisition variables that define the sampling points are ω_1 and Δ, so $\mathbf{x}_n = (\omega_{1n}, \Delta_n)$. Our application of interest is clinical quantitative MT imaging, in which acquisition time is limited. We set $N = 10$, as in previously published acquisition protocols (16,18). In each voxel we fix R_B at 1 s^{-1} and estimate the remaining parameters by fitting the model in Eq. [2] to the measurements using a Levenberg-Marquardt algorithm, as in Ref. 18. Since the second term on the right-hand side of Eq. [4] is very close to zero, F/R_A is almost linear in F, and we choose to optimize the sampling directly for the fitted parameters even though F is the parameter of interest rather than F/R_A. In fact, we include only four terms ($Q = 4$) in the sum in Eqs. [6] and [7], which correspond to the parameters $RM_0{}^A$, F/R_A, $T_2{}^B$, and $1/R_A T_2{}^A$. We exclude the terms for R_B, which we fix to 1 s^{-1}, and S_0, which is of little interest.

MATERIALS AND METHODS

Optimization

To optimize the sampling points for quantitative MT, we must search for the sampling configurations, $\mathbf{x}_1, \ldots, \mathbf{x}_N$, that minimize V and $V_{\max}$ from Eqs. [6] and [7]. The global minima of these cost functions are hard to find because of local minima. Thus, we use a minimization technique that combines simulated annealing with the downhill simplex method (28). The routine makes 100 evaluations of the cost function at each temperature T, using expansions, contractions, and reflections of the simplex to reach a local minimum. The cost-function value is randomly perturbed by a quantity proportional to $\log(T)$ to allow occasional uphill transitions, which increases the probability of finding the global minimum. The temperature T, decreases according to the annealing schedule

$$T(n+1) = (1-\varepsilon)T(n), \quad [8]$$

where $1 \gg \varepsilon > 0$, until it reaches a prespecified final temperature. The initial temperature is 100, above which we find empirically that perturbations are too large to be useful, and the final temperature is 0.001, below which perturbations make no difference within the accuracy of the scanner's settings of ω_1 and Δ.

To further reduce the effects of local minima in the objective function, we repeat the optimization from a number of different starting points. In each optimization, we use six unique starting points and select the final configuration with the minimum value of the cost function. This procedure does not guarantee that the global-minimum configuration will be found, but in practice it gives good results that we do not expect to improve on dramatically with more starting points or slower annealing schedule. For larger numbers of points ($N > 10$), more careful optimization may be required. From Eq. [2] we can see that $S(\omega_1, \Delta) = S(\pm\omega_1, \pm\Delta)$; therefore, two final configurations are considered to be identical if their points differ only in sign.

Off-resonance irradiation inevitably produces some direct saturation of the liquid pool (7), which rapidly in-

Table 1
Standard and Optimized MT Acquisition Schemes*

1) Standard (see Ref. 18) $V = 7.0$; $V_{max}^{hypercube} = 48.1$; $V^{max15\ ROIs} = 12.4$		2) Optimized for a single set $V = 2.2$; $V_{max}^{hypercube} = 11.8$; $V^{max15\ ROIs} = 3.5$		3) Optimized for 15 ROIs $V = 2.3$; $V_{max}^{hypercube} = 12.3$; $V^{max15\ ROIs} = 3.5$		4) Optimized for the hypercube $V = 3.1$; $V_{max}^{hypercube} = 11.3$; $V^{max15\ ROIs} = 4.9$	
ω_1	Δ	ω_1	Δ	ω_1	Δ	ω_1	Δ
222	0.4	37	42	12	14.84	217	0.15
222	1	265	0.15	255	0.15	310	0.26
222	3	128	100	365	100	400	1.13
222	7.5	397	1.38	398	1.28	444	1.74
222	20	405	1.48	465	2.05	495	2.34
885	0.4	405	1.48	483	2.36	522	82.26
885	1	950	0.97	950	0.88	950	1.28
885	3	950	0.98	950	1.02	950	1.30
885	7.5	950	0.98	950	1.04	950	13.62
885	20	950	11.6	950	11.15	950	56.79

*V and V_{max} are the objective functions used for the optimization and are defined by Eqs. [6] and [7] in the text. The superscripts on V_{max} indicate the set of parameter combinations. ω_1 is measured in rad s^{-1}; Δ is measured in kHz.

creases as the offset frequency of the MT pulse is reduced, and eventually suppresses any MR signal when Δ becomes small compared to the MT pulse bandwidth. Values of Δ close to zero thus result in dramatic signal loss, particularly in the presence of static field inhomogeneities, where the actual offset frequency can deviate slightly from its nominal value and vary across the sample. Moreover, safety issues associated with the deposition of RF energy into the patient, as measured by the specific absorption rate (SAR), impose some restrictions on the maximum value of ω_1. We constrain $\omega_1 < 950$ rad s^{-1} and $|\Delta| > 0.15$ kHz to enforce these practical limitations on the combinations of ω_1 and Δ. The value of 0.15 kHz is determined empirically by selecting the minimum offset frequency that does not result in macroscopic areas of signal loss in the image at the median value of ω_1 ($\omega_1 \approx 450$ rad s^{-1}). The most appropriate choice may vary depending on the MT pulse bandwidth and the static field strength.

We compute and test three optimal configurations (schemes 2–4), and compare them with a standard configuration from Ref. 18, which we label scheme 1. Scheme 2 minimizes V for a set of typical MT parameters measured in WM at 1.5T. Specifically, we set $RM_0{}^A = 25$ s^{-1}, $F/R_A = 0.08$ s^{-1}, $T_2{}^B = 10$ μs, $1/R_A T_2{}^A = 10$, $R_B = 1$ s^{-1}, and $S_0 = 1000$ (arbitrary), which are typical measured values in Refs. 13 and 18. Schemes 3 and 4 minimize V_{max} for different combinations of settings. For scheme 3 we use the values measured in $K = 15$ regions of interest (ROIs) in gray matter (GM) and WM by Cercignani et al. (18). For scheme 4 we define a range of values for each fitted parameter that are typical for the human brain at 1.5T, and take the $K = 33$ combinations from the center and the 32 vertices of the "hypercube" spanning all the ranges. We estimate the ranges from the measurements in Ref. 18. Specifically, $RM_0{}^A \in [25, 35]$ s^{-1}, $F/R_A \in [0.05, 0.1]$ s^{-1}, $T_2{}^B \in [7, 15]$ μs, $1/R_A T_2{}^A \in [10, 20]$, $S_0 \in [700, 1300]$, and $R_B = 1$ s^{-1}. In the minimization of V, we set $\varepsilon = 0.0001$, but for V_{max} we increase ε to 0.001, since computation of the cost function is more expensive. With $\varepsilon = 0.0001$, the number of temperatures between the starting and final temperatures is approximately 115,000, while with $\varepsilon = 0.001$ it is approximately 11,500. An Intel Pentium 4 3.2GHz machines computes V about 50 times per second (thus V_{max} $50/K$ times per second), so each optimization typically takes a few days.

EXPERIMENTS AND RESULTS

This section outlines some experiments that were conducted to test the new acquisition protocols computed from the methods outlined above. We had two central hypotheses: 1) little improvement would come from optimizing V_{max} for many combinations of parameter settings over the simpler optimization of V for a single well-chosen set, and 2) all of the optimum point sets would produce more stable maps of the MT parameters compared to standard acquisition schemes from the literature.

Optimal Configurations

Table 1 compares the three optimized sampling schemes with the standard 10-point scheme (scheme 1) from Ref. 18. Table 1 shows the values of V and V_{max} for every scheme in order to provide a theoretical comparison of their performances. The minimum V_{max} is 11.3 for the hypercube (scheme 4) and 3.5 for the ROIs (scheme 3). For the optimum sampling scheme returned for a single setting (scheme 2), V_{max} is 11.8 for the hypercube and 3.5 for the ROIs, supporting our first hypothesis that the optimum sampling would be fairly insensitive to the exact p_i used in the optimization. In particular, the sets of points in schemes 2 and 3 are very similar. With the standard scheme (scheme 1), V_{max} is 48.1 for the hypercube and 12.4 for the ROIs, which suggests large improvements in precision with the optimum configuration compared to the standard one. Table 1 shows similar reductions for V when scheme 1 ($V = 7.0$) is compared with scheme 2 ($V = 2.2$), scheme 3 ($V = 2.3$), and scheme 4 ($V = 3.1$).

Simulations

We use Monte Carlo simulations at various noise levels to measure the error in the fitted parameters from the differ-

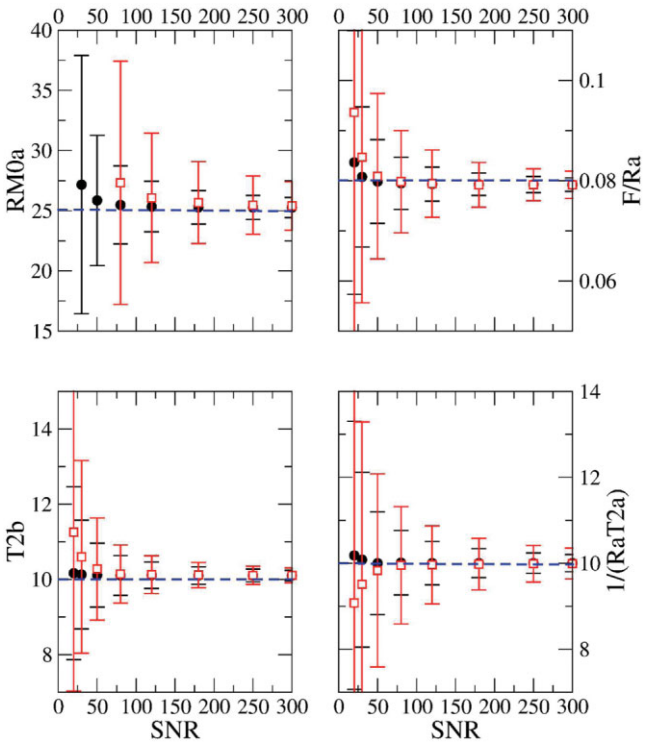

FIG. 1. Plot of mean MT parameters from 10000 Monte Carlo simulations against SNR in the unweighted image for the standard scheme 1 (red) and optimized scheme 2 (black). The error bars show the SD. The dashed blue line shows the value of parameter used to synthesize the data. No symbol is present for $RM_0{}^A$ from scheme 1 at SNR $\leq$ 50 because the mean value was too large to be in the scale of the figure.

ent acquisition schemes. First, four (one for each scheme) noise-free synthetic sets of 10 MT-weighted signal intensities are generated by substituting the MT parameters used to obtain scheme 2 into Eq. [2]. We add complex noise with zero-mean Gaussian real and imaginary parts, and take the modulus to obtain a noisy data sample. The SD of the Gaussian noise is S_0/Σ, where Σ is the desired SNR in the unweighted image, which we vary over the interval [20, 300]. For each level of noise we generate 10000 sets of noisy independent samples and fit Eq. [2] to each set.

For all four parameters ($RM_0{}^A$, F/R_A, $T_2{}^B$, and $1/R_AT_2{}^A$), the estimates obtained using schemes 2–4 are more precise than those obtained with scheme 1, with SDs smaller by a factor of at least 1.5. Accuracy also increases, particularly at low SNR ($\leq$50). Schemes 2 and 3 produce similar performance, while scheme 4 produces slightly larger SDs for the estimates of $T_2{}^B$ and $1/R_AT_2{}^A$.

Figure 1 compares the estimates of the MT parameters from schemes 1 and 2 and confirms an improvement in both precision and accuracy for all four parameters at all SNR levels when the optimized scheme is used. We omit the results for schemes 3 and 4 from Fig. 1 for clarity, since they are almost indistinguishable from the scheme 2 results.

To confirm the near-equivalence in performance of the three optimized schemes, we repeat the simulation experiments using the worst-case combinations of parameters from both the hypercube and 15-region sets of combinations. The set of MT parameter values that gave the worst V among the 15 ROIs was {$RM_0{}^A$ = 33.2 s^{-1}; F/R_A = 0.07 s^{-1}; $T_2{}^B$ = 12.3μs; $1/R_AT_2{}^A$ = 11.4; and S_0 = 804}. Using these values for the simulations, schemes 2 and 3 perform similarly in terms of precision (as expected given their similarity), with slightly lower variance from scheme 3 at very low SNR ($\leq$50) for $T_2{}^B$. The estimates obtained using scheme 4 are more precise than those obtained using scheme 1, but less precise than those obtained using schemes 2 and 3.

The set of MT parameter values that gave the worst V among the 33 vertices-and-center of the hypercube was {$RM_0{}^A$ = 35.0 s^{-1}; F/R_A = 0.05 s^{-1}; $T_2{}^B$ = 7.0 μs; $1/R_AT_2{}^A$ = 20.0; and S_0 = 700}. Schemes 2 and 3 performed similarly, with slightly larger variances than scheme 4 for all four parameter estimates. The SD of $RM_0{}^A$ estimated using scheme 1 was larger than that obtained with any other scheme at all SNR levels, while all four schemes behaved similarly for estimates of F/R_A, $T_2{}^B$, and $1/R_AT_2{}^A$ at SNR > 120.

MRI Acquisition and Analysis

In the experiment described in this section we assess the precision of the in vivo MT parametric maps and aim to confirm the conclusions from the simulation experiments. We estimate and compare the SDs of each MT parameter from different acquisition schemes in several ROIs in the human brain. Acquiring enough separate data sets for good estimates of the SD in each voxel requires a prohibitively long scan time. Therefore, we collect a small number of data sets for each scheme and use the bootstrap method (29) to produce a large number of data sets by resampling with replacement on a voxel-by-voxel basis. To further reduce the scan time, we restrict the in vivo analysis to schemes 1 and 2, since the Monte Carlo simulations showed that schemes 3 and 4 provide little extra advantage.

One healthy subject (male, 35 years old) was scanned twice on a 1.5 T system (SIGNA Horizon Echospeed; General Electric, Milwaukee, WI, USA) using a 3D MT-weighted fast spoiled-gradient recalled-echo (SPGR) sequence (18) (TR/TE = 28/5.1 ms, flip angle = 5°, Gaussian MT pulses, duration = 14.6 ms, SD = 2.98 ms, bandwidth = 125 Hz, 28 reconstructed slices). During each session, three complete data sets were obtained using either scheme 1 or scheme 2. In addition to the MT data, two four-shot spin-echo echo-planar imaging (EPI) sequences (TR/TE = 15000/13.6 ms, flip angle = 60° and 120°, matrix = 64 × 64), and two 3D SPGRs (TR = 16 ms, TE = 4 and 8.54 ms, flip angle = 25°) were collected for B_1 and B_0 mapping, respectively. The total scan time for each session was just under an hour. The body coil was used for signal transmission and the eight-channel head coil was used for reception, with the exception of the B_0 mapping sequence, in which the body coil was used for reception as well as transmission. This avoids the scanner's automatic combined reconstruction when using multichannel receiver coils and therefore preserves the complex data. After image coregistration (30), B_1 maps are obtained from the

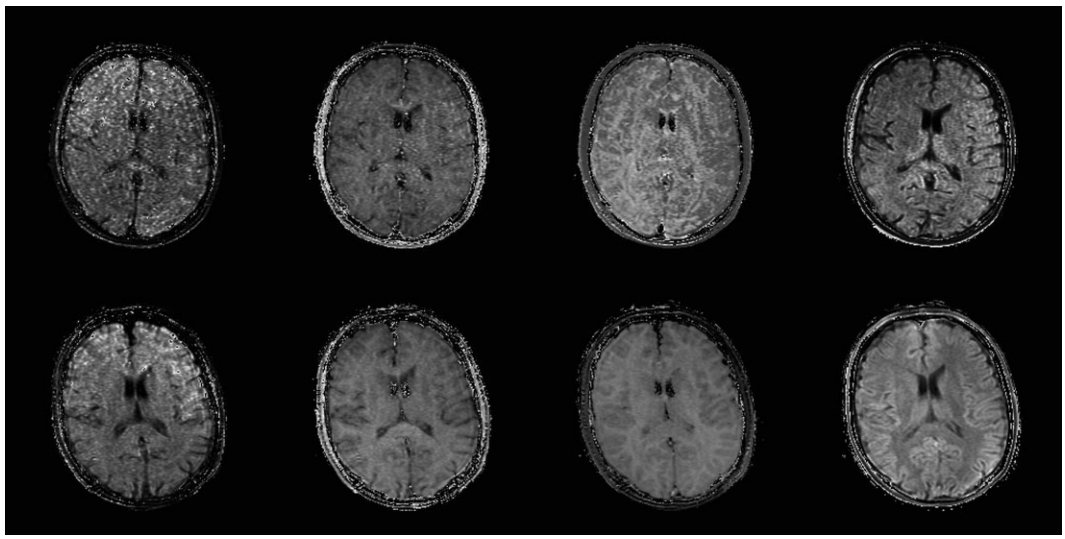

FIG. 2. MT parametric maps obtained with the standard (top) and the optimized (bottom) sampling schemes. From left to right: RM_0^A, F/R_A, T_2^B, and $1/R_AT_2^A$. Spatial homogeneity and WM/GM contrast increase in all maps, and in particular the RM_0^A map, for the optimized scheme.

multishot EPI data with the use of the double-angle method (31) and the resulting maps are smoothed using a third-order polynomial fitting. Maps of the B_0 deviation from the value in the center of the head are calculated from the phase images obtained from the SPGR data, according to Cusack and Papadakis (32).

To compare the in vivo results with the numerical simulations, we measure the SNR of the MT-SPGR data in WM. The SNR is difficult to measure because it is not uniform in images obtained with multichannel coils. However, a good estimate can be obtained using two images with the same weighting, as described by Firbank et al. (33). Using this method, we estimate the following SNR values in four WM regions on the least saturated image: 86 in the anterior periventricular WM, 123 in the posterior periventricular WM, 112 in the corona radia, and 70 in the frontal WM. We fit the MT model in Eq [2], as in Ref. 18, to the first repeat from each acquisition scheme to obtain an initial qualitative comparison of the two schemes. During the fitting, B_1 and B_0 maps are used to correct for any deviation from the nominal values of ω_1 and Δ (14,17). Figure 2 shows the MT parametric maps from the first repeat from each scheme. Maps from the optimized scheme show clear improvements in spatial homogeneity and GM/WM contrast.

Next, with the three repeats we generate 1000 MT raw data sets for each scheme by bootstrapping. We fit the model to each of the 1000 bootstrapped samples, providing 1000 estimates of RM_0^A, F/R_A, T_2^B, and $1/R_AT_2^A$. We compute the coefficient of variance (COV, defined as the SD divided by the mean across the 1000 bootstrapped samples) of the MT parameters in each voxel. Table 2 compares the average COVs (expressed in percentage units) obtained for two GM (putamen and thalamus) and three WM (genu of corpus callosum, anterior periventricular, and posterior periventricular) ROIs outlined by hand on the least saturated image of each set. The COV for the optimized scheme 2 is consistently less than half that for the standard scheme 1, and reduces by a factor of up to 5 in some regions. This shows dramatic improvements in parameter estimation and confirms the qualitative observations from Fig. 2.

DISCUSSION

We have presented a method based on the theory of CRLB for optimizing the sample positions for any multiparametric quantitative MRI technique. We used this method to optimize in vivo quantitative MT acquisition, which collects a series of MT-weighted images and estimates several

Table 2
Mean Parameter COV (Percentage Units) Across 1000 Bootstrapped Samples for Schemes 1 (Standard) and 2 (Optimized)*

Scheme	RM_0^A		F/R_A		T_2^B		$1/(R_AT_2^A)$	
Scheme	1	2	1	2	1	2	1	2
Putamen	31	11	16	6	18	5	15	3
Thalamus	23	8	10	4	10	2	10	3
Corpus callosum	28	11	12	5	8	4	10	4
Anterior periv. WM	27	13	12	7	10	4	11	4
Posterior periv. WM	21	9	10	5	9	3	10	5

*For bilateral structures, the value is the average of left and right. periv. = periventricular; WM = white matter.

parameters by fitting a nonlinear model to the measurements. The uncertainty associated with the estimate of RM_0^A is particularly large (18,34), and some studies based on Ramani et al.'s (16) model have failed to estimate this parameter entirely (13). The Monte Carlo simulations and in vivo results showed that selecting appropriate sampling schemes greatly enhances both the precision and accuracy of MT parametric maps. In particular, the optimized acquisitions reduce the error in parameter estimates by factors of around 2 or 3 and allow us to estimate RM_0^A with accuracy and precision similar to those of the other parameters. The sampling-scheme optimization requires specific settings of the MT parameters. The results in the previous section show that the optimum schemes are robust to these settings within the range of MT parameters observed in the brain. Specifically, we propose two types of objective function. First, we use Eq. [6] to select the optimum sampling points for measuring a single set of MT parameters. Second, we use Eq. [7] to account for the variability of MT parameters within the brain (18,35). The optimizations for the single set and the 15 ROIs return very similar schemes (schemes 2 and 3), with almost identical values of V and $V_{\max}$, which suggests that the optimum sampling depends only weakly on the exact parameter settings in the optimization, at least over the range of combinations observed in healthy brain tissue. Monte Carlo simulations confirmed that schemes 2 and 3 provide very similar precision of parameter estimates across a range of settings typical of the human brain. Scheme 4, which is optimized for the hypercube, provides slightly lower precision for the typical combinations, probably because even though the hypercube covers the range of possible parameter values, the vertices are not necessarily realistic combinations. Table 1 shows that although scheme 3 (optimized for ROIs) is close to optimal for the hypercube ($V_{max} = 12.3$; optimal $V_{max} = 11.3$), scheme 4 (optimized for hypercube) is not so close to optimal for the ROIs ($V_{max} = 4.9$; optimal $V_{max} = 3.5$). This suggests that although some of the ROI combinations have high variance with the optimum configuration for the hypercube, we can reduce that variance significantly (by optimizing the ROI combinations directly) with only a slight cost at the parameter combinations in the hypercube. Since the hypercube contains the combinations in each ROI (though not at the vertices where we sample it), denser sampling of the hypercube is likely to reduce the ROIs metric while low variance is retained in all portions of the hypercube. However, the ROIs optimization produces a sample set with a near-optimal hypercube metric, which suggests that optimizing over portions of the hypercube that are not close to any of the ROI combinations provides little benefit. Moreover, computation time is linear in the number of parameter combinations K, so denser sampling of the hypercube rapidly becomes impractical. Since the ROI combinations are directly measured in the brain (18), we expect that the configuration from the ROI optimization is more useful than the hypercube in practice. One possible refinement would be to weight combinations by the frequency of the observations in the brain.

We cannot be certain that the optimization will find the global minimum configurations. Minimization of Eqs. [6] and [7] is complicated by the presence of local minima. To reduce the effects of local minima, we use six unique starting points and a method based on simulated annealing, which is particularly suitable for this type of problem. The improvement in the quality of the parametric maps is evident in Fig. 2, and scheme 2 clearly provides substantial advantages compared to scheme 1. We expect little further practical improvement from more careful optimization that gets closer to the global minimum, although we may require this for larger N.

The three optimal schemes (2–4) in Table 1 include six to seven unique points, while some positions are sampled twice. They all include four points at high power (of which three are unique), and one measurement that is as close to the Larmor frequency as possible with the $\Delta \geq 0.15$kHz constraint. Lifting the constraint on Δ moves that point even closer to the Larmor frequency, which rapidly brings the signal down due to direct saturation of the liquid pool. At offset frequency lower than 0.15 kHz, Δ approaches the magnitude of the B_0 field inhomogeneities, which causes large artifacts and areas of low SNR in the image. Furthermore, although the MT model in Eq. [2] does account for direct effect (7), it assumes an infinitesimally narrow pulse bandwidth. As the offset frequency approaches zero, the finite pulse width becomes more significant.

Furthermore, all of the optimized schemes include at least one point with either large Δ or small ω_1, which yields measurements that are virtually identical to those with no MT weighting (i.e., S_0). This suggests that a good estimate of S_0 is necessary to fit the remaining parameters accurately.

These results are in contrast to the experimental observations of Sled and Pike (14), who employed regular sampling at constant logarithmic frequency steps using only two values of ω_1. In the present work we restrict our analysis to 10-point acquisition schemes, whereas Sled and Pike used a different model (14,36) and a larger number of points. It seems likely from our experiments that, for $N > 10$, the configuration of ω_1 and Δ that minimizes the variance of MT parameters is substantially different from the two-ω_1 scheme they suggest. We confine investigations of this hypothesis to future work.

Another area for further work is to investigate the dependence of V and $V_{\max}$ on N, to establish whether there is cutoff above which the benefit gained by increasing the number of points is small compared to the cost in terms of increased scan time. Using an approach similar to ours (although it was limited to the measurement of a single MT parameter while keeping the rest fixed), Samson et al. (25) attempted to identify optimal combinations of ω_1 and Δ with small N, and demonstrated that N can be reduced to 3 without compromising the SD of the estimated T_2^B, or to 6 without compromising the SD of the estimated f.

Equation [2] differs from that proposed by Ramani et al. (16) in that we use $F = M_0^B/M_0^A$ instead of $f = M_0^B/(M_0^A + M_0^B)$. We restated the model to keep a near-linear relationship between the fitted parameter, F/R_A, and the parameter of interest, F. In practice, since $f \ll 1$ in the brain, we do not expect to observe large differences in the optimum sets of points returned for either choice of parameter. The model we use here is only an approximation when using pulsed MT. Several alternative MT models have been proposed (15,35,36), and the optimization procedure we in-

troduce here can be applied equally to these more complex models. This also we confine to further work.

In this study we optimized quantitative MT sampling schemes for healthy brain tissue, and we assume that changes in MT parameters in pathology will be small enough that the sampling will remain close to optimal. In particular, the parameter F is likely to change most in pathology. However, change in F is generally accompanied by a change in R_A that keeps F/R_A (which is the important parameter for sampling optimization) fairly stable. Nevertheless, alternative approaches are possible. An alternative optimization might choose a configuration that maximizes the contrast between healthy and damaged tissue. However, a complication with this approach in a disease like multiple sclerosis is that a range of abnormalities can occur in, for example, chronic and inflammatory lesions or normal-appearing WM. The optimum schemes we propose provide a substantial improvement for the quantification of MT parameters in both healthy and pathologic brain compared to previous acquisitions, and sampling optimization for contrasting specific parameter combinations is another area for future work.

In some applications, F may be the only parameter of interest. In such cases, Eq. [6] reduces to a classic CRLB optimization problem. One last area for further work is to quantify the improvement in the precision of a single parameter (for example F, provided by the use of a scheme optimized for the estimation of F only) compared to those optimized for all MT parameters at the same time.

REFERENCES

1. Tofts P. Quantitative MRI of the brain. New York: Wiley; 2003.
2. Crawley AP, Henkelman RM. A comparison of one-shot and recovery methods in T1 imaging. Magn Reson Med 1988;7:23–34.
3. Whittall KP, MacKay AL, Graeb DA, Nugent RA, Li DK, Paty DW. In vivo measurement of T2 distributions and water contents in normal human brain. Magn Reson Med 1997;37:34–43.
4. Le Bihan D, Breton E, Lallemand D, Grenier P, Cabanis E, Laval-Jeanet M . MR imaging of intra-voxel incoherent motions: application to diffusion and perfusion in neurologic disorders. Radiology 1986;161: 401–407.
5. Basser PJ, Mattiello J, LeBihan D. Estimation of the effective self-diffusion tensor from the NMR spin echo. J Magn Reson B 1994;103: 247–54.
6. Boxerman JL, Hamberg LM, Rosen BR, Weisskoff RM. MR contrast due to intravascular magnetic susceptibility perturbations. Magn Reson Med 1995;34:555–66.
7. Henkelman RM, Huang X, Xiang QS, Stanisz GJ, Swanson SD, Bronskill MJ. Quantitative interpretation of magnetization transfer. Magn Reson Med 1993;29:759–66.
8. Wolff SD, Balaban RS. Magnetization transfer contrast (MTC) and tissue water proton relaxation in vivo. Magn Reson Med 1989;10: 135–144.
9. Kucharczyk W, Macdonald PM, Stanisz GJ, Henkelman RM. Relaxivity and magnetization transfer of white matter lipids at MR imaging: importance of cerebrosides and pH. Radiology 1994;192:521–529.
10. Koenig SH, Brown 3rd RD, Spiller M, Lundbom N. Relaxometry of brain: why white matter appears bright in MRI. Magn Reson Med 1990;14: 482–495.
11. Dousset V, Grossman RI, Ramer KN, Schnall MD, Young LH, Gonzalez-Scarano F, Lavi E, Cohen JA. Experimental allergic encephalomyelitis and multiple sclerosis: lesion characterization with magnetization transfer imaging. Radiology 1992;182:483–491.
12. van Buchem MA, McGowan JC, Kolson DL, Polansky M, Grossman RI. Quantitative volumetric magnetization transfer analysis in multiple sclerosis: estimation of macroscopic and microscopic disease burden. Magn Reson Med 1996;36:632–636.
13. Davies GR, Tozer DJ, Cercignani M, Ramani A, Dalton CM, Thompson AJ, Barker GJ, Tofts PS, Miller DH. Estimation of the macromolecular proton fraction and bound pool T2 in multiple sclerosis. Mult Scler 2004;10:607–613.
14. Sled JG, Pike GB. Quantitative imaging of magnetization transfer exchange and relaxation properties in vivo using MRI. Magn Reson Med 2001;46:923–931.
15. Yarnykh VL. Pulsed Z-spectroscopic imaging of cross-relaxation parameters in tissues for human MRI: theory and clinical applications. Magn Reson Med 2002;47:929–939.
16. Ramani A, Dalton C, Miller DH, Tofts PS, Barker GJ. Precise estimate of fundamental in vivo MT parameters in human brain in clinically feasible times. Magn Reson Imaging 2002;20:721–731.
17. Sled JG, Pike GB. Quantitative interpretation of magnetization transfer in spoiled gradient echo MRI sequences. J Magn Reson 2000;145:24–36.
18. Cercignani M, Symms MR, Schmierer K, Boulby PA, Tozer DJ, Ron M, Tofts PS, Barker GJ. Three-dimensional quantitative magnetisation transfer imaging of the human brain. NeuroImage 2005;27:436–441.
19. Xing D, Papadakis NG, Huang CL, Lee VM, Carpenter TA, Hall LD. Optimised diffusion-weighting for measurement of apparent diffusion coefficient (ADC) in human brain. Magn Reson Imaging 1997;15:771–784.
20. Alexander DC, Barker GJ. Optimal imaging parameters for fiber-orientation estimation in diffusion MRI. Neuroimage 2005;27:357–367.
21. Kay SM. Fundamentals of statistical signal processing: estimation theory. Englewood Cliffs, NJ: Prentice Hall; 1993.
22. Jones JA. Optimal sampling strategies for the measurement of relaxation times in proteins. J Magn Reson 1997;126:283–286.
23. Brihuega-Moreno O, Heese FP, Hall, LD. Optimization of diffusion measurements using Cramer-Rao lower bound theory and its application to articular cartilage. Magn Reson Med 2003;50:1069–1076.
24. Pineda AR, Reeder SB, Wen Z, Pelc NJ. Cramer-Rao bounds for three-point decomposition of water and fat. Magn Reson Med 2005;54:625–635.
25. Samson RS, Symms MR, Cercignani M, Tozer DJ, Tofts PS. Optimisation of quantitative magnetization transfer (QMT) sequence acquisition parameters. In: Proceedings of the 14th Annual Meeting of ISMRM, 2006, Seattle, WA, USA. p 2493.
26. Li JG, Graham SJ, Henkelman RM. A flexible magnetization transfer line shape derived from tissue experimental data. Magn Reson Med 1997;37:866–871.
27. Atkinson AC, Donev AN. Optimum experimental designs. Oxford: Oxford University Press; 1992.
28. Press WH, Teukolsky SA, Vettering WT, Flannery BP. Chapter 10: Minimization or maximization of functions. In: Numerical recipes in C. Cambridge: Cambridge University Press; 1992. p 444–455.
29. Efron B. Bootstrap methods: another look at the jackknife. Ann Stat 1979;7:1–16.
30. Symms MR, Boulby PA, Barker GJ. Cross-modality registration: validation and application in magnetisation transfer imaging. In: Proceedings of the 11th Annual Meeting of ISMRM, Toronto, Canada, 2003. p 1050.
31. Stollberger R, Wach P. Imaging of the active B_1 field in vivo. Magn Reson Med 1996;35:246–51.
32. Cusack R, Papadakis N. New robust 3-D phase unwrapping algorithms: application to magnetic field mapping and undistorting echo-planar images. Neuroimage 2002;16:754–764.
33. Firbank MJ, Coulthard A, Harrison RM, Williams ED. A comparison of two methods for measuring the signal to noise ratio on MR images. Phys Med Biol 1999;44:N261–N264.
34. Graham SJ, Henkelman RM. Understanding pulsed magnetization transfer. J Magn Reson Imaging 1997;7:903–912.
35. Sled JG, Levesque I, Santos AC, Francis SJ, Narayanan S, Brass SD, Arnold DL, Pike GB. Regional variations in normal brain shown by quantitative magnetization transfer imaging. Magn Reson Med 2004;51: 299–303.
36. Pike GB. Pulsed magnetization transfer contrast in gradient echo imaging: a two-pool analytic description of signal response. Magn Reson Med 1996;36:95–103.