Journal of Magnetic Resonance **154,** 1–5 (2002)
doi:10.1006/jmre.2001.2457, available online at http://www.idealibrary.com on IDEAL®

Quantification of Human Brain Metabolites from *in Vivo* ^{1}H NMR Magnitude Spectra Using Automated Artificial Neural Network Analysis

Yrjö Hiltunen,* Jouni Kaartinen,† Juhani Pulkkinen,‡ Anna-Maija Häkkinen,§ Nina Lundbom,¶ and Risto A. Kauppinen‡

*Pehr Brahe Software Laboratory, The Raahe Institute of Computer Engineering, Finland; †VTT Electronics, Finland; ‡NMR Research Group, A. I. Virtanen Institute, University of Kuopio, Finland; §Department of Oncology, Helsinki University Hospital, Helsinki, Finland; and ¶Department of Radiology, Helsinki University Hospital, Helsinki, Finland

Received December 14, 2000; revised October 5, 2001

Long echo time (TE = 270 ms) *in vivo* proton NMR spectra resembling human brain metabolite patterns were simulated for lineshape fitting (LF) and quantitative artificial neural network (ANN) analyses. A set of experimental *in vivo* ^{1}H NMR spectra were first analyzed by the LF method to match the signal-to-noise ratios and linewidths of simulated spectra to those in the experimental data. The performance of constructed ANNs was compared for the peak area determinations of choline-containing compounds (Cho), total creatine (Cr), and N-acetyl aspartate (NAA) signals using both manually phase-corrected and magnitude spectra as inputs. The peak area data from ANN and LF analyses for simulated spectra yielded high correlation coefficients demonstrating that the peak areas quantified with ANN gave similar results as LF analysis. Thus, a fully automated ANN method based on magnitude spectra has demonstrated potential for quantification of *in vivo* metabolites from long echo time spectroscopic imaging.

Key Words: **artificial neural network; ^{1}H nuclear magnetic resonance spectroscopy; brain metabolites; quantification; simulated spectra.**

INTRODUCTION

More than 15 low-molecular weight cerebral metabolites have been assigned and quantified *in vivo* by proton magnetic resonance spectroscopy (^{1}H NMR) (*1*) proving its unique value as a noninvasive tool for neurochemistry (*2*, *3*). The majority of this neurochemical information is obtainable by means of the short echo time (TE) ^{1}H NMR. However, the simplified long TE (>100 ms) brain proton spectra have been demonstrated to possess a wealth of potentially useful information for clinical neuroscience (*4*, *5*). If clinical ^{1}H NMR spectroscopy is expected to evolve to a more widely applicable method for medical diagnosis, reproducible data-analysis should occur close to real-time. This is not trivial, since a trend from single-volume ^{1}H NMR to multi-volume magnetic resonance spectroscopic imaging (^{1}H MRSI), producing several hundred spectra from each study, is underway. Automated quantification for ^{1}H MRSI without a need for excessive spectroscopic expertise would facilitate the use of spectroscopic data in clinical decision making (*6–8*).

NMR spectroscopy data are commonly quantified both in time and frequency domains. For clinical applications, it is a convention to use the frequency domain display. Spectroscopy data analysis is currently laborious due to the need of supervision requiring steps such as phase correction. Manual phase correction is time-consuming and also difficult to accomplish in low signal-to-noise ratio spectra often encountered in *in vivo* NMR spectroscopy. Phase correction is a necessary step prior to peak area determination, since lineshape fitting (LF) and artificial neural networks (ANN) analysis of phase corrected NMR spectra has been shown to result in a compromised accuracy in quantification of biological data sets (e.g., *9*). With regard to the clinical applicability, the use of various fitting methods requires spectroscopic expertise, thus making fully automated analysis by LF methods cumbersome and impractical.

We have previously shown that both *in vitro* (*9*) and *in vivo* (*10*) NMR spectroscopy data can be quantified by ANN. In the present work, we have extended the automation procedure by studying the use of magnitude spectra as inputs for ANN quantification to overcome the phase correction problem. It is demonstrated that the new method estimates concentrations of human brain metabolites from long TE ^{1}H MRSI with a similar accuracy as LF analysis of manually phase-corrected spectra.

EXPERIMENTAL

In Vivo ^{1}H NMR Spectroscopy Measurements

^{1}H MRSI data consisting of 90 spectroscopic data sets from 38 glioma patients (oligoastrocytoma, anaplastic astrocytoma, oligodendroglioma, or glioblastoma multiforme) were included in the study. ^{1}H MRSI measurements were performed in a 1.5 T MR scanner (Magnetom Vision, Siemens, Erlangen, Germany) using a standard CP head coil. For MRSI localization, 2D FLASH images in 3 orientations were acquired. A double spin echo sequence with 16 × 16 phase encoding steps was used

1090-7807/02 $35.00

with TR = 2600 ms and TE = 270 ms with 2 acquisitions. Volume preselection of $80/100 \times 80/100 \times 15$ mm^3 and voxel size 1.5 cm^3, was performed including the tumor area. Spectral data was collected either from one or two 15 mm thick slices covering the tumor volume as judged from localizer MR images. From this MRSI data 451 spectra, with a large variation in concentrations for individual metabolites, were chosen for the subsequent quantitative metabolite analysis.

Simulation of the Spectra

A total of 2000 long echo-time *in vivo* NMR spectra were simulated using Lorentzian lineshapes and varying frequencies, intensities, as well as half linewidths of the three detectable metabolites (Cho, Cr, and NAA). The complex spectrum can be written in the form of

$$F(v) = \mathrm{Re}(v) + i\ \mathrm{Im}(v) = |F(v)|\, e^{i\phi(v)}, \qquad [1]$$

where

$$|F(v)| = [\mathrm{Re}^2(v) + \mathrm{Im}^2(v)]^{1/2} \qquad [2]$$

and

$$\phi(v) = \tan^{-1}[\mathrm{Im}(v)/\mathrm{Re}(v)]. \qquad [3]$$

$|F(v)|$ is the amplitude spectral density (called the magnitude spectra in this paper) and $\phi(v)$ the phase angle. The following equations were used to calculate the real (Re) and imaginary part (Im) of the simulated spectra

$$\mathrm{Re}(v) = \sum_{i=1}^{3} ABS_i(v) = \sum_{i=1}^{3} \frac{a_i^2 I_i}{a_i^2 + 4(v - v_i)^2}, \qquad [4]$$

$$\mathrm{Im}(v) = \sum_{i=1}^{3} DISP_i(v) = \sum_{i=1}^{3} \frac{2a_i I_i (v - v_i)}{a_i^2 + 4(v - v_i)^2}, \qquad [5]$$

in which a_i is the half linewidth, I_i is the signal intensity, and v_i is the resonance frequency. $ABS_i(v)$ stands for absorption and $DISP_i(v)$ stands for dispersion part of the Lorentzian i. However, the real and imaginary parts both commonly contain absorption and dispersion components, even after the phase correction. Noise from experimental spectra was added to mimic the genuine *in vivo* spectra. The half linewidths in each spectrum were the same for signals at the resonance frequencies of Cho, Cr, and NAA. Variations of parameters used in calculations of simulated spectra are shown in Table 1. The values are taken from the LF analyses of experimental spectra. The simulated spectra agree with the experimental spectra except that the correct peak areas are known. Equations [4] and [2] were used to form the phase and magnitude spectra data sets for the ANNs, respectively.

TABLE 1
Parameter Variations of the Data Sets

Data set	a^a	n^b	P_L	S/N_{Cho}	S/N_{Cr}	S/N_{NAA}
Experimental set	4.4–18	451	≈±1	1.6–33.6	1.4–22.0	1.2–38.3
Training set	4.4–18	1000	≈±1	1.4–50.5	0.9–38.9	0.6–81.0
Test set	4.4–18	500	≈±1	1.4–50.5	0.9–38.9	0.6–81.0
Independent test set	4.4–18	500	≈±1	1.6–37.4	1.3–25.1	2.4–53.8

[a] a is the half linewidth (in Hz).
[b] n is the number of spectra.
[c] P_L is the displacement of location of the peaks (in Hz).
[d] S/N is the signal-to-noise ratio (intensity of signal/standard deviation of noise).

The peak areas for the resonances, i.e., the actual outputs (A_i), were calculated using

$$A_i = \frac{\pi}{2} a_i I_i. \qquad [6]$$

Data Processing and Lineshape Fitting Analysis

In order to remove the residual water signal from the experimental human brain ^{1}H spectra, the Hankel Lanczos singular value decomposition (HLSVD) (*11, 12*) method was applied. Both real and imaginary parts were used for conversion of spectroscopic data into the magnitude mode. The real parts of the experimental and simulated spectra were analyzed in the frequency domain by means of an automated analysis program (MRSTools Ltd., Kuopio, Finland) under a Matlab-software platform (Mathworks, Natick, MA). The mathematical lineshape model for the three detectable brain metabolites were

$$L(v) = \sum_{i=1}^{3} \frac{a_i^2 I_i}{a_i^2 + 4(v - v_i)^2} \cos(\phi_i) + \sum_{i=1}^{3} \frac{2a_i I_i (v - v_i)}{a_i^2 + 4(v - v_i)^2} \sin(\phi_i) + c_0 + c_1 v, \qquad [7]$$

in which $c_0 + c_1 v$ takes into account the background variation. The linewidths were constrained to be equal, because this approach has been shown to be more accurate especially for spectra with very low signal-to-noise ratio (*10*).

ANN Analysis

The simulated spectra were used in neural network training to calculate the spectral points (input) from absolute metabolite peak areas (output) corresponding to metabolites Cho, Cr, and NAA. The simulated spectra were divided into three subsets. The first subset (1000 of the spectra) was the training set, which was used for computing the gradient and updating the network weights and biases. The second subset (500 of the spectra) was the test set. The error of the test set was monitored during the training process. When the error of the test set began to increase

for a specified number of iterations, the training was terminated and the weights and biases at the minimum of test error were returned. The third subset (500 of the spectra) was the independent test set. The error of this independent test set was not used during the training, but rather to test the model. The experimental spectra have been used only as an independent test set.

A feed-forward, fully connected neural network was used applying the back propagation algorithm in the training. The 130 data points formed the input layer of the network. In the hidden layer, 20 neurons were used and the output layer consisted of 3 neurons. Logistic transfer functions were used. This network topology was evaluated using a trial-and-error process. The training process was accomplished using a 1000 MHz Pentium PC applying a commercial NeuroShell 2 program (NeuroShell 2, Release 4.0, Ward System Group Inc., 1998).

RESULTS AND DISCUSSION

The experimental *in vivo* ^{1}H NMR spectra ($n = 451$) were analyzed by the LF method so that the spectral qualities as depicted by signal-to-noise ratio and linewidths matched those in the simulated spectra. The 500 simulated spectra of the independent test set were also quantified by the LF method showing correlation coefficients with the actual output values of 0.936, 0.905, and 0.968 for Cho, Cr, and NAA, respectively (Table 2 and Fig. 1).

The training processes with the simulated spectral data and actual values led to complete learning of the neural networks. The correlation coefficients between the ANN and actual output values for the independent test set were almost identical with those obtained by LF methods (Table 2). Using the magnitude spectra as inputs, the correlation coefficients between the ANN and the actual output values were also high (Table 2 and Fig. 2).

For evaluating the usefulness of the method in cases of non-ideal data, the 451 experimental magnitude spectra were analyzed using the ANN method. The correlation coefficients between ANN and LF results were 0.881, 0.823, and 0.925 for Cho, Cr, and NAA, respectively. Although these results show less correlation than the simulated results (Table 2), they seem to confirm that the ANN method is not very sensitive to baseline and lineshape distortions. They are also consistent with our

TABLE 2

The Summary of Results of the Metabolite Peak Areas for the Simulated Spectra of the Independent Test Set Using Lineshape Fitting (LF) and Artificial Neural Network (ANN) Analysis

		Cho	Cr	NAA
Method	Spectra	r^a/rms^b	r^a/rms^b	r^a/rms^b
LF	phase spectra	0.936/1315	0.905/1187	0.968/1372
ANN	phase spectra	0.959/1060	0.942/880	0.980/1093
ANN	Magnitude spectra	0.934/1288	0.924/984	0.969/1373

[a] r is the correlation coefficient.

[b] rms is the root mean square error.

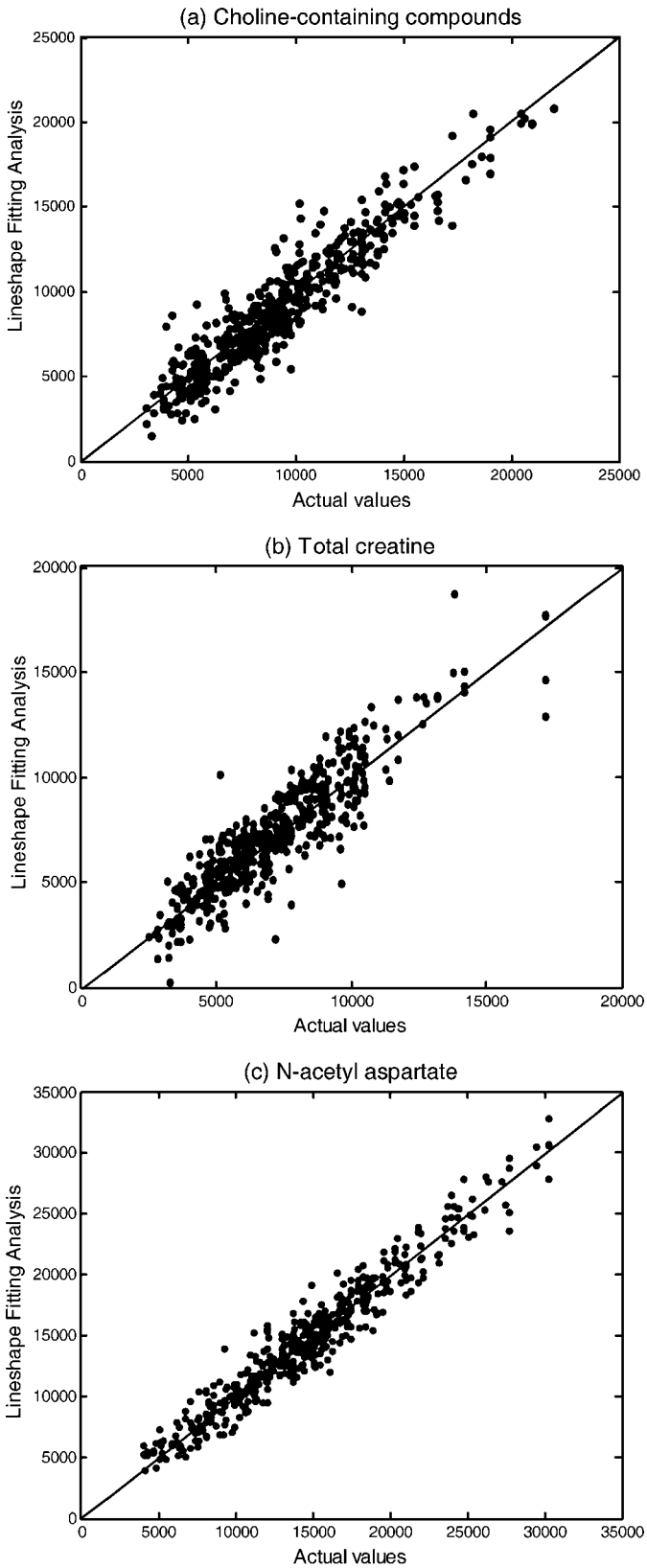

FIG. 1. Correlation between the actual and peak areas obtained by lineshape fitting for (a) choline-containing compounds, (b) total creatine, and (c) N-acetyl aspartate in the simulated data set. The diagonal lines indicate the identity line.

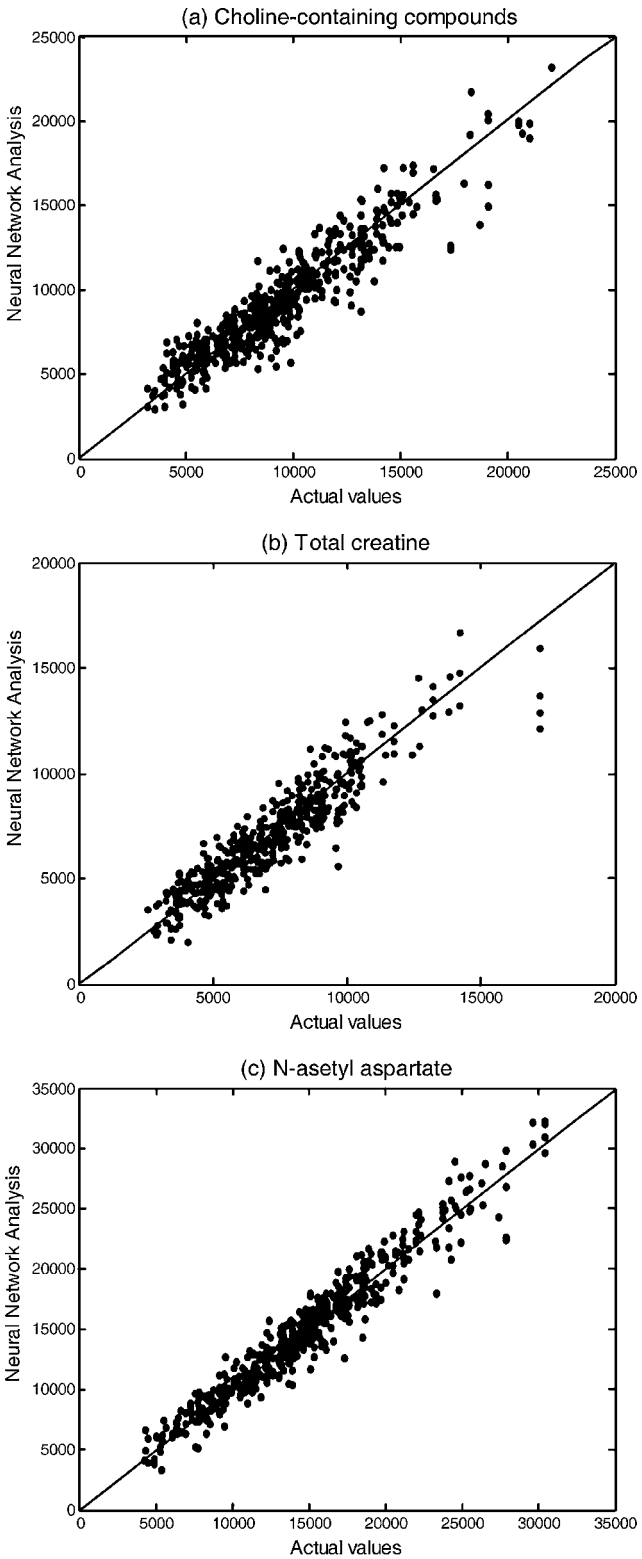

FIG. 2. Correlation between the actual and peak areas obtained by ANN (using magnitude spectra) for (a) choline-containing compounds, (b) total creatine, and (c) N-acetyl aspartate in the simulated data set. The diagonal lines indicate the identity line.

previous study (*10*) that the simulated spectra can be used in the training of ANNs, which is applicable for quantification in cases when sufficiently large experimental data sets are not available.

The lineshape fitting method in the frequency domain is commonly used to quantify NMR spectra (*1, 13–15*). In the sophisticated LF-methods, linear combinations of model spectra of metabolite solutions are exploited for fitting the short TE brain spectra (*1, 13*). This procedure greatly improves quantification of partially overlapping signals. Recently, techniques based on time-domain fitting were proposed for the use of biomedical NMR spectroscopy (*16, 17*). The time domain technique proposed by Vanhamme *et al.* (*16*) is aimed at overcoming the problems posed by low signal-to-noise ratio in *in vivo* spectra by incorporating a model function fitting. An interesting time domain analysis algorithm based on minimum rank normalization was proposed for spectral quantification for samples with few visible signals (*17*). However, this method normally uses the rank of the Hankel matrix constructed from the difference of two time domain signals, and thus it has limited applicability in clinical settings. The present results show that quantification of long TE *in vivo* magnitude brain spectra with an automated ANN method is feasible. The correspondence between the actual and LF values of the simulated spectra is good, although not perfect even though the peaks have ideal Lorentzian lineshapes and correct phases. This suggests that in cases of low signal-to-noise ratio, which may occur in clinical NMR spectroscopy, accurate quantification may turn out to be difficult. Practically, the peaks in the experimental spectra show nonideal lineshapes, and the real and imaginary parts contain both common absorption and dispersion components making the fitting procedure too inaccurate. It should be noted that the variation of metabolite concentrations in human brain tumors is relatively large (*14*). Therefore, we believe that the errors imposed by the data analysis method proposed here does not prevent the use of ^{1}H NMR data in aiding the decision process.

In the case of the magnitude spectra, phase correction is not needed since these spectra do not have a phase. This is expected to affect lineshapes and also intensity ratios of the spectra. However, ANN can easily be trained for alterations of these variables, and therefore, results comparable to those using the exactly phase-corrected spectra can be obtained. This makes it possible to avoid a difficult and time-consuming preprocessing step in the quantification and classification of brain tumor spectra.

CONCLUSIONS

The ANN quantification results from human brain ^{1}H MRSI agree well with those obtained from a LF analysis demonstrating that magnitude spectra as inputs can be used to overcome the practical problem of fast and reliable phase correction. Our study suggests that it is feasible to construct a fully automated real-time quantifying analyzer for long TE *in vivo* ^{1}H NMR spectra using ANN with magnitude spectra.

ACKNOWLEDGMENTS

This study was supported by the Academy of Finland, Helsinki University Central Hospital, and the Magnus Ehrnrooth Foundation.

REFERENCES

1. J. Pfeuffer, I. Tkac, S. W. Provencher, and R. Gruetter, Toward an *in vivo* neurochemical profile: Quantification of 18 metabolites in short-echo-time ^{1}H NMR spectra of the rat brain, *J. Magn. Reson.* **141,** 104–120 (1999).
2. R. G. Shulman, D. L. Rothman, and A. M. Blamire, NMR studies of human brain function, *Trends Biochem. Sci.* **19,** 522–526 (1994).
3. J. M. Hakumäki, M. Ala-Korpela, and R. A. Kauppinen, ^{1}H nuclear magnetic resonance spectroscopy: Noninvasive neurochemistry for basic research and clinical applications, *Curr. Top. Neurochem.* **1,** 59–85 (1997).
4. B. Ross and T. Michaelis, Clinical applications of magnetic resonance spectroscopy, *Magn. Reson. Q.* **10,** 191–247 (1994).
5. P. B. Barker, J. D. Glickson, and R. N. Bryan, *In vivo* magnetic resonance spectroscopy of human brain tumors, *Top. Magn. Reson. Imaging* **5,** 32–45 (1993).
6. M. C. Preul, Z. Caramanos, D. L. Collins, J. G. Villemure, R. Leblanc, A. Olivier, R. Pokrupa, and D. L. Arnold, Accurate, noninvasive diagnosis of human brain tumors by using proton magnetic resonance spectroscopy, *Nat. Med.* **2,** 323–325 (1996).
7. J.-P. Usenius, S. Tuohimetsä, P. Vainio, M. Ala-Korpela, Y. Hiltunen, and R. A. Kauppinen, Automated classification of human brain tumours by neural network using *in vivo* ^{1}H magnetic resonance spectroscopic metabolite phenotypes, *NeuroReport* **7,** 1597–1600 (1996).
8. A. R. Tate, J. R. Griffiths, I. Martinez-Perez, A. Moreno, I. Barba, M. E. Cabanas, D. Watson, J. Alonso, F. Bartumeus, F. Isamat, I. Ferrer, F. Villa, E. Ferrer, A. Capdevila, and C. Arus, Towards a method for automated classification of ^{1}H MRS spectra from brain tumours, *NMR Biomed.* **11,** 177–91 (1998).
9. Y. Hiltunen, E. Heininiemi, and M. Ala-Korpela, Lipoprotein-lipid quantification by neural-network analysis of ^{1}H NMR data from human blood plasma, *J. Magn. Reson. Ser. B* **106,** 191–194 (1995).
10. J. Kaartinen, S. Mierisova, J. M. Oja, J.-P. Usenius, R. A. Kauppinen, and Y. Hiltunen, Automated quantification of human brain metabolites by artificial neural network analysis from *in vivo* single-voxel ^{1}H NMR sectra, *J. Magn. Reson.* **134,** 176–179 (1998).
11. W. W. F. Pijnappel, A. van den Boogaart, R. de Beer, and D. van Ormondt, SVD-based quantification of magnetic resonance signals, *J. Magn. Reson.* **97,** 122–134 (1992).
12. A. van Den Boogaart, D. van Ormondt, W. W. F. Pijnappel, R. de Beer, and M. Ala-Korpela, Removal of the water resonance from ^{1}H magnetic resonance spectra, *in* "Mathematics in Signal Processing III" (J. G. McWhirter, Ed.), pp. 175–195, Clarendon Press, Oxford (1994).
13. S. W. Provencher, Estimation of metabolite concentrations from localized *in vivo* proton NMR spectra, *Magn. Reson. Med.* **30,** 672–679 (1994).
14. J. P. R. Usenius, R. A. Kauppinen, P. A. Vainio, J. A. Hernesniemi, M. P. Vapalahti, L. A. Päljärvi, and S. J. Soimakallio, Quantitative metabolite patterns of human brain tumors: Detection by H-1 NMR spectroscopy *in vivo* and *in vitro*, *J. Comput. Assist. Tomogr.* **18,** 705–713 (1994).
15. M. Ala-Korpela, J.-P. Usenius, J. Keisala, A. van den Boogaart, P. A. Vainio, J. Jokisaari, S. Soimakallio, and R. A. Kauppinen, Quantification of metabolites from single-voxel *in vivo* ^{1}H NMR data of normal human brain means of time-domain analysis, *MAGMA* **3,** 129–136 (1995).
16. L. Vanhamme, T. Sundin, P. Van Hecke, S. Van Huffel, and R. Pintelon, Frequency-selective quantification of biomedical magnetic resonance spectroscopy data, *J. Magn. Reson.* **143,** 1–16 (2000).
17. R. Romano, M. T. Santini, and P. L. Indovina, A time-domain algorithm for NMR spectral normalization, *J. Magn. Reson.* **146,** 89–99 (2000).