JOURNAL OF MAGNETIC RESONANCE **134,** 176–179 (1998)
ARTICLE NO. MN981477

Automated Quantification of Human Brain Metabolites by Artificial Neural Network Analysis from *in Vivo* Single-Voxel ^{1}H NMR Spectra

Jouni Kaartinen,* Šarka Mierisová,† Joni M. E. Oja,‡ Jukka-Pekka Usenius,‡ Risto A. Kauppinen,‡ and Yrjö Hiltunen*,[1]

*Pehr Brahe Laboratory, The Raahe Institute of Computer Engineering, Raahe, Finland; †Faculty of Chemical Technology, Department of Physical Chemistry, Slovak University of Technology, Bratislava, Slovakia; and ‡NMR Research Group, A. I. Virtanen Institute, University of Kuopio, Kuopio, Finland

Received November 25, 1997; revised April 1, 1998

A real-time automated way of quantifying metabolites from *in vivo* NMR spectra using an artificial neural network (ANN) analysis is presented. The spectral training and test sets for ANN containing peaks at the chemical shift ranges resembling long echo time proton NMR spectra from human brain were simulated. The performance of the ANN constructed was compared with an established lineshape fitting (LF) analysis using both simulated and experimental spectral data as inputs. The correspondence between the ANN and LF analyses showed correlation coefficients of order of 0.915–0.997 for spectra with large variations in both signal-to-noise and peak areas. Water suppressed ^{1}H NMR spectra from 24 healthy subjects were collected and choline-containing compounds (Cho), total creatine (Cr), and N-acetyl aspartate (NAA) were quantified with both methods. The ANN quantified these spectra with an accuracy similar to LF analysis (correlation coefficients of 0.915–0.951). These results show that LF and ANN are equally good quantifiers; however, the ANN analyses are more easily automated than LF analyses.

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

INTRODUCTION

Recent technical improvements in localized *in vivo* ^{1}H NMR spectroscopy have made it a valuable tool for studying neurochemistry noninvasively as well as for acting as a potential modality in clinical neuroradiology. If clinical NMR spectroscopy is expected to develop into a widely accepted method for medical diagnosis, data analysis must be reproducible and should occur close to real time when a patient is still in the magnet.

Several procedures have been applied in biomedical NMR spectroscopy to quantify areas of metabolite resonances from frequency domain spectra. With methods such as the lineshape fitting analysis (*1–4*) and the linear combination of model spectra (*5*), noisy ^{1}H NMR spectra or overlapping peaks in spectra can be assessed. However, the reliable use of the lineshape fitting analysis methods needs spectroscopic expertise, so that a fully automated analyzer by these methods is rather difficult to develop.

It has recently been shown that ANN analysis offers some important advantages for biomedical NMR data analysis (*6–11*). These include instant and fully automated modes of action. However, utilization of the ANN analysis is limited, since there is a need for a training data set, which properly describes the different aspects of a particular application. As far as quantification is concerned, this would require peak areas and consequent concentrations of the NMR spectroscopy detectable metabolites to be collected at a largely varying range.

In this study we used a computer simulated spectroscopic data set for training a neural network. This trained ANN was utilized to quantify metabolite peak areas of *in vivo* long echo time ^{1}H NMR spectra of a human brain. Our results indicate that the ANN-based method can be regarded as an automated means of quantifying ^{1}H NMR spectra also in cases where acceptable experimental data sets are not available.

RESULTS

The training process with the simulated ^{1}H NMR data led to complete learning of the neural network. The correspondence between the ANN and the actual output values for the metabolite signals showed correlation coefficients of the order of 0.99–1.00.

The performance of the trained ANN was tested using 100 randomly simulated spectra. These spectra, such as the training set spectra, represented a wide range of different signal-to-noise ratios (3.0–50.3) and half linewidths of the peaks (2.2–6.0 Hz). The correlation coefficients between the ANN and the actual output values were over 0.98 for all metabolite signals (Table 1). The correlation coefficients between the LF and the actual output values were also excellent, which means that the correspondence between the ANN and LF analyses was also very high (Table 1).

[1] To whom correspondence should be addressed.

1090-7807/98 $25.00

TABLE 1
The Results of the Test and Experimental Data of Artificial Neural Network (ANN) and Lineshape Fitting (LF) Analysis

Data set	Cho r^a/rms^b	Cr r^a/rms^b	NAA r^a/rms^b
Test set (ANN and real)	0.991/0.019	0.984/0.020	0.991/0.020
Test set (LF and real)	0.992/0.017	0.982/0.022	0.991/0.020
Test set (ANN and LF)	0.996/0.020	0.994/0.013	0.997/0.011
Experimental set (ANN and LF)	0.947/0.020	0.915/0.010	0.951/0.021

[a] r is the correlation coefficient.
[b] rms is the root mean square error.

The usefulness of the trained ANN was further evaluated by applying it to Cho, Cr, and NAA resonances in a set of 24 experimental *in vivo* ^{1}H NMR spectra, which were also quantified by the LF analysis. The correlation coefficients between these two methods for Cho, Cr, and NAA were 0.947, 0.915, and 0.951, respectively (Table 1 and Fig. 1).

DISCUSSION

The lineshape fitting method in the frequency domain is commonly used to quantify NMR spectra (*1–4*). Our results show that the correspondence between the ANN and LF values was very high, which indicates that the LF and ANN methods are equally good in spectral quantification. Slightly smaller correlation coefficients in the case of the experimental spectra compared to the simulated test spectra may result from the fact that the peaks of experimental spectra do not have ideal Lorentzian lineshapes, which causes some errors to both LF and ANN results. Using several kinds of lineshapes in training of the ANN seems to make it possible to devise a model-independent method for analyzing ^{1}H NMR spectra.

The training of an ANN is the most critical phase in designing an ANN analyzer, because ANN needs a wide training set of good quality to be able to work well. The quality of the spectra should be good, accurate concentrations of metabolites of each spectrum should be known in the training set, and there should be enough variation in the concentration for every metabolite. In the case of biomedical *in vivo* applications it is difficult to find examples good enough for the training set. This problem is often the most significant one in the use of ANN to quantify NMR spectroscopic data. An alternative method, as shown here, would be to use simulated data in training an ANN with the ability to quantify experimental proton spectra. It is also possible in the training process to use both experimental and simulated data together to obtain a more versatile ANN analyzer.

There are several methods for absolute quantification of metabolite concentrations: (i) internal standard, (ii) external standard, and (iii) external standard displacement methods. In the present work we did not use endogenous water (*11, 12*) as a concentration reference, in which case we could not determine absolute metabolite concentrations. However, because the percentual areas of metabolite peaks (Cho, Cr, and NAA) were accurately quantified by ANN, one would expect to derive absolute concentrations from these figures with adequate referencing.

The ANN quantification results agree well with those obtained from an independent LF analysis. This shows that computer simulated NMR spectroscopic data can be used to overcome the practical problem of obtaining an acceptable experimental training data set. Our study indicates that it is possible to construct a fully automated real-time quantifying analyzer for *in vivo* ^{1}H NMR spectra using ANN with simulated training data.

EXPERIMENTAL

In Vivo ^{1}H NMR Spectroscopy Measurements

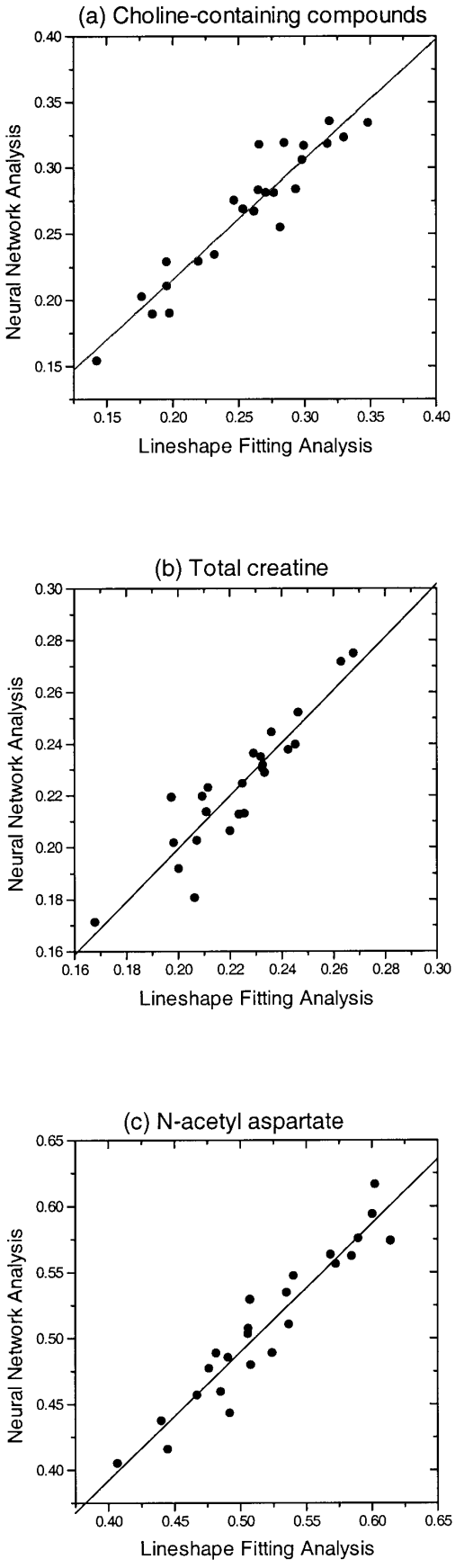
The single-voxel double-spin echo PRESS sequence [sweepwidth of 1 kHz, 1024 complex data points, echo time (TE) of 270 ms, repetition time (TR) of 1500 ms, the volume of interest (VOI) of 8 ml, and 256 scans] was used for *in vivo* ^{1}H NMR spectroscopy data acquisitions from the parietal cortex of 24 healthy subjects with a Siemens Magneton SP63 1.5 T scanner with a quadrature head coil. Water suppression was accomplished by three Gaussian-shaped CHESS pulses with bandwidths of 60 Hz. Eight scans were acquired without water suppression for compensation of eddy currents (*13*).

Simulation of the Spectra

Long echo-time *in vivo* NMR spectra (TE = 270 ms) were simulated using Lorentzian lineshapes (see Eq. [1]) and varying frequencies of peaks (P_L), intensities, and half linewidths of the three detectable metabolites (Cho, Cr, and NAA). The half linewidths were the same for signals at the resonance frequencies of Cho, Cr, and NAA. Simulated noise similar to the noise in experimental data was also added to mimic signal-to-noise ratios in the genuine *in vivo* spectra (see Eq. [1]). This resulted in a complete training set with highly variable metabolite concentrations and signal-to-noise ratios. The following equation was used to calculate the frequency domain spectrum to form an input (T_{input}) for the ANN:

$$T_{input}(v) = \sum_{i=1}^{3} \frac{a_i^2 I_i}{a_i^2 + 4(v - v_i)^2} + N = L(v) + N. \quad [1]$$

Here a_i is the half linewidth, I_i the signal intensity, v_i the resonance frequency, $L(v)$ the Lorentzian lineshape, and N the simulated noise. The peak areas for the resonances, i.e., the actual outputs (T_{output}), were calculated using the equation



$$T_{output} = \sum_{i=1}^{3} \frac{\pi}{2} a_i I_i. \qquad [2]$$

The performance of the trained ANN was further evaluated with an additionally simulated test set. Variations of the parameters in training and test sets used in the computer simulations are shown in Table 2.

Data Processing and Lineshape Fitting Analysis

In order to remove the residual water signal from the experimental spectra, the Hankel Lanczos singular value decomposition (HLSVD) (*14, 15*) method was applied. This was not necessary in the case of simulated data since they were simulated without the residual water signal. Then, the data were analyzed in the frequency domain by means of the FITPLA^C program (*2, 16*) concerning two different kinds of prior knowledge: (a) The linewidths of all three fitted peaks were allowed to be estimated without any constraints, and then (b) the linewidths were constrained to be the same. The second approach has provided more accurate and precise results especially for spectra with very low signal-to-noise ratios. Therefore, only these results were used for an assessment of the performance and accuracy of the trained neural network.

ANN Analysis

In this study the simulated spectra were used to train the network to calculate the spectral points (input) from the metabolite signal percentual areas of peaks (output) corresponding to the human brain tissue metabolites Cho, Cr, and NAA. Both simulated and experimental spectra were processed in the same way in the frequency domain before the ANN analysis. First, the average of the noise of the spectrum was set to 0. Then the intensity of each frequency point of each spectrum was divided by the sum intensity of the area studied (17.58–243.16 Hz). The topology of the feed-forward fully connected three-layer Perceptron was an input layer of 21 neurons, two hidden layers of 22 neurons, and an output layer of 3 neurons. This network topology was evaluated using a trial-and-error process. Logistic transfer functions were used and all inputs and outputs were scaled between 0 and 1. The training process of the ANN was performed using a TurboProp algorithm, which iteratively adjusted the connection weights of the neural network to give a desired output. The training process was accomplished using a 90 MHz Pentium PC applying a commercial NeuroShell 2 program (*17*).

FIG. 1. Correlations in the experimental data set between (a) the choline-containing compounds, (b) total creatine, and (c) the *N*-acetyl aspartate concentrations estimated by the lineshape fitting and the artificial neural network analysis. The values on the axis are relative areas of metabolite peaks.

TABLE 2
Training, Test, and Experimental Set Parameter Variations

Data set	a^a	n^b	$P_L{}^c$	$S/N_{Cho}{}^d$	S/N_{Cr}	S/N_{NAA}
Training set	2.2–6.0	1000	±0.98	1.3–45.4	0.5–55.3	4.4–54.8
Test set	2.2–6.0	100	±0.98	3.0–40.7	3.6–32.5	6.0–50.3
Experimental set	2.4–5.8	24	≈±0.98	5.4–17.1	3.9–13.6	11.8–27.5

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

ACKNOWLEDGMENTS

The authors thank Dr. Mika Ala-Korpela (Wihuri Research Institute, Helsinki, Finland) for helpful discussions. R.A.K. thanks the Academy of Finland, Y.H. the Instrumentarium Foundation, and J.M.E.O. the North Savo Cultural Fund of the Finnish Cultural Foundation, the Magnus Ehrnrooth Foundation, and the Saastamoinen Foundation for financial support.

REFERENCES

1. M. Ala-Korpela, Y. Hiltunen, J. Jokisaari, S. Eskelinen, K. Kiviniitty, M. Savolainen, and Y. A. Kesäniemi, A comparative study of ^{1}H NMR lineshape fitting analyses and biochemical lipid analyses of the lipoprotein fractions VLDL, LDL and HDL, and total human blood plasma, *NMR Biomed.* **6,** 225–233 (1993).
2. M. Ala-Korpela, A. Korhonen, J. Keisala, S. Hörkkö, P. Korpi, L. P. Ingman, J. Jokisaari, M. J. Savolainen, and Y. A. Kesäniemi, ^{1}H NMR based quantitation of human lipoproteins and their lipid contents directly from plasma, *J. Lipid Res.* **35,** 2292–2304 (1994).
3. J.-P. R. Usenius, R. A. Kauppinen, P. A. Vainio, J. A. Hernesniemi, M. P. Vapalahti, L. A. Paljärvi, and S. J. Soimakallio, Quantitative metabolite patterns of human brain tumours: Detection by ^{1}H NMR spectroscopy in vivo and in vitro, *J. Comput. Assist. Tomogr.* **18,** 705–713 (1994).
4. M. Ala-Korpela, J.-P. R. 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 by means of time-domain data analysis, *Magma* **3,** 129–136 (1995).
5. S. W. Provencher, Estimation of metabolite concentrations from localized in vivo proton NMR spectra, *Magn. Reson. Med.* **30,** 672–679 (1993).
6. J. R. Long, V. G. Gregoriou, and P. J. Gemperline, Spectroscopic calibration and quantification using artificial neural networks, *Anal. Chem.* **62,** 1791–1797 (1990).
7. M. Bos, A. Bos, and W. E. van der Linden, Data processing by neural networks in quantitative chemical analysis, *Analyst* **118,** 323–328 (1993).
8. R. de Beer, F. Michaelis, D. van Ormondt, B. P. O. van Tongeren, P. R. Luyten, and H. van Vroonhoven, Reduced lipid contamination in in vivo ^{1}H MRSI using time-domain fitting and neural network classification, *Magn. Reson. Imag.* **11,** 1019–1026 (1993).
9. Y. Hiltunen, E. Heiniemi and M. Ala-Korpela, Lipoprotein-lipid quantification by neural-network analysis of ^{1}H NMR data from human blood plasma, *J. Magn. Reson. B* **106** 191–194 (1995).
10. M. Ala-Korpela, Y. Hiltunen, and J. D. Bell, Quantification of biomedical NMR data using artificial neural network analysis: Lipoprotein lipid profiles from ^{1}H NMR data of human plasma. *NMR Biomed.* **8,** 235–244 (1995).
11. J.-P. R. Usenius, S. Tuohimetsä, P. A. Vainio, M. Ala-Korpela, Y. Hiltunen, and R. A. Kauppinen, Automatic classification of human brain tumours by neural network analysis using in vivo ^{1}H magnetic resonance spectroscopic metabolite phenotypes. *NeuroReport* **7,** 1597–1600 (1996).
12. P. Christiansen, O. Henriksen, M. Stubgaard, P. Gideon, and H. B. W. Larsson, In vivo quantification of brain metabolites by ^{1}H-MRS using water as an internal standard, *Magn. Reson. Imag.* **11,** 107–118 (1993).
13. U. Klose, *In vivo* proton spectroscopy in presence of eddy currents, *Magn. Reson. Med.* **14,** 26–30 (1990).
14. 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).
15. 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).
16. A. van den Boogaart, M. Ala-Korpela, J. Jokisaari, and J. R. Griffiths, Time and frequency domain analysis of NMR data compared: An application to 1D ^{1}H spectra of lipoproteins, *Magn. Reson. Med.* **31,** 347–358 (1994).
17. NeuroShell 2 (Release 3.0), Ward System Group, Inc. (1996).