NMR IN BIOMEDICINE
NMR Biomed. 2008; **21**: 33–41
Published online 8 March 2007 in Wiley InterScience
(www.interscience.wiley.com) DOI:10.1002/nbm.1151

Automated tissue segmentation and blind recovery of ^{1}H MRS imaging spectral patterns of normal and diseased human brain

Shuyan Du,[1] Xiangling Mao,[2] Paul Sajda[1] and Dikoma C. Shungu[2]*

[1]Department of Biomedical Engineering, Columbia University, New York, NY, USA
[2]Department of Radiology, Weill Medical College of Cornell University, New York, NY, USA

Received 2 August 2006; Revised 30 November 2006; Accepted 29 December 2006

ABSTRACT: Constrained non-negative matrix factorization (cNMF) with iterative data selection is described and demonstrated as a data analysis method for fast and automatic recovery of biochemically meaningful and diagnostically specific spectral patterns of the human brain from ^{1}H MRS imaging (^{1}H MRSI) data. To achieve this goal, cNMF decomposes *in vivo* multidimensional ^{1}H MRSI data into two non-negative matrices representing (a) the underlying tissue-specific spectral patterns and (b) the spatial distribution of the corresponding metabolite concentrations. Central to the proposed approach is automatic iterative data selection which uses prior knowledge about the spatial distribution of the spectra to remove voxels that are due to artifacts and undesired metabolites/tissues such as the strong lipid and water components. The automatic recovery of diagnostic spectral patterns is demonstrated for long-*TE* ^{1}H MRSI data on normal human brain, multiple sclerosis, and serial brain tumor. The results show the ability of cNMF with iterative data selection to automatically and simultaneously recover tissue-specific spectral patterns and achieve segmentation of normal and diseased human brain tissue, concomitant with simplification of information content. These features of cNMF, which permit rapid recovery, reduction and interpretation of the complex diagnostic information content of large multi-dimensional spectroscopic imaging data sets, have the potential to enhance the clinical utility of *in vivo* ^{1}H MRSI. Copyright © 2007 John Wiley & Sons, Ltd.

KEYWORDS: magnetic resonance spectroscopic imaging (MRSI); iterative data selection; non-negative matrix factorization (NMF); tissue segmentation; neurological disorders; human brain

INTRODUCTION

In vivo single-voxel ^{1}H MRS (1,2) and its multi-voxel metabolic imaging counterpart, ^{1}H MRSI (3,4), have emerged in recent years as powerful non-invasive techniques for use as adjuncts to conventional structural MRI in the diagnostic evaluation of a variety of neurological disorders (5,6), including brain cancers, epilepsy, stroke, mitochondrial disorders, as well as numerous adult and pediatric white matter diseases. Whereas the acquisition and processing of single-voxel MRS data are now fully automated on virtually every commercial clinical MR instrument, the level of complexity of the vast majority of multi-dimensional ^{1}H MRSI studies is such that the acquisition, processing, quantitative analysis, interpretation and reporting of such data still require substantial technical expertise, thus limiting its wider clinical use and availability. A further impediment to the widespread clinical implementation of ^{1}H MRSI is the recent availability of high-field MR instruments with phased-array coils and fast multi-planar spectroscopic imaging techniques which allow the acquisition of higher spatial resolution data in clinically acceptable scan time. Although these substantial gains in data acquisition speed and spatial resolution are welcome advances, the resulting large spectroscopic imaging data sets are rapidly leading to an unintended or undesired side effect: the extraction of meaningful spectral information and interpretation of each MRSI data set has become challenging for most available conventional peak-tracking and quantitation techniques (7,8). For such large multi-planar ^{1}H MRSI data sets, availability of a data analysis method that can quickly recover the important spectral patterns within a data set and then present these to clinicians in a way that allows immediate, *visual*, and reliable assessment of the diagnostic information would greatly contribute to the practical utility and widespread clinical use of this increasingly powerful but relatively complex technology.

**Correspondence to:* D. C. Shungu, Department of Radiology, Citigroup Biomedical Imaging Center, Weill Medical College of Cornell University, 516 E 72nd Street, New York, NY 10021, USA.
E-mail: dcs7001@med.cornell.edu

Abbreviations used: cNMF, constrained non-negative matrix factorization; Lac, lactate; MRSI, MRS imaging; MS, multiple sclerosis; NAA, N-acetyl-l-aspartate; NMF, non-negative matrix factorization; tCho, total choline; tCr, total creatine.

Recently, we introduced the constrained non-negative matrix factorization (cNMF) approach (9,10), a data analysis method that allows fast recovery of biochemically meaningful spectral patterns contained in a spectroscopic imaging data set. cNMF decomposes the recorded MRSI data into two non-negative matrices representing (a) the underlying tissue-specific spectral patterns and (b) the spatial distribution of their concentrations. This dual non-negative matrix representation of MRSI data will be shown to be a form of automated 'tissue segmentation' that minimizes the partial volume averaging of spectral intensities from different tissues contributed to each voxel. Both the basic cNMF algorithm and the iterative data selection method have been previously described (9,10). For completeness, however, we will begin this report with a very brief overview of the method. We will then demonstrate the potential broad clinical utility of cNMF for recovering biochemically meaningful spectral patterns of the normal brain, as well as those of a number of brain disorders. Each case is selected to illustrate a specific characteristic of the method, and overall to demonstrate the utility of cNMF with iterative data selection in the analysis of *in vivo* human brain ^{1}H MRSI data.

EXPERIMENTAL

Brief background and overview of cNMF

In MRSI, each tissue type can be viewed as having a characteristic spectral profile related to its biochemical composition. The observed spectra are thus simply a combination of different metabolite resonances, the amplitudes of which are proportional to the concentrations of the metabolites in each voxel. As a result the observed MRSI data, **X**, can be expressed as

$$\mathbf{X} = \mathbf{AS} + \mathbf{N} \tag{1}$$

where the columns in **A** represent the concentration of metabolites in the constituent tissue within each voxel and the rows in **S** represent the corresponding spectral resonances. **N** represents additive noise. The concentration matrix **A** has M columns (one for each spectral pattern) and N rows (one for each voxel). **X** and **S** have L columns (one for each resonance).

As one can interpret **A** as concentration, the matrix can be assumed to be non-negative. In addition, as the constituent spectra **S** represent amplitudes of resonances, in theory the smallest resonance amplitude is zero, corresponding to the absence of resonance at a given frequency (where we ignore cases of negative peaks such as in J-modulation). The factorization of eqn (1) is therefore constrained by:

$$\mathbf{A} \geq \mathbf{0},\ \mathbf{S} \geq \mathbf{0} \tag{2}$$

It should, however, be noted that the spatial inhomogeneity of the static magnetic field may introduce unknown phase shifts, which must be separately estimated for every voxel, as errors in estimating these phase shifts, as well as measurement noise, may lead to violations of the positivity constraint of the observed spectra by causing **X** to have negative values.

In recent years there has been a great deal of effort to simultaneously exploit the statistical structure of MRSI data sets to solve eqn (1) as a blind source separation problem (9–14). For example, Ochs *et al*. (13) formulated the decomposition as a problem in Bayesian inference, which was called Bayesian spectral decomposition, whereas in (14), the well-established independent component analysis approach (11) was used. Our contribution has been to develop cNMF (9,10), which was based on the NMF algorithm of Lee and Seung (15,16). As an extension of NMF, cNMF enables non-negative factorization even for noisy observations, which may result in observed spectra having negative values, by including a positivity constraint, forcing negative values in the recovered spectral patterns (sources) and concentration distributions to be approximately zero, as these negative values are thought to be due to noise interference and not physically realistic. As cNMF includes non-negativity constraints on the matrix factorization, it recovers biochemically meaningful spectral patterns, particularly for tissue types that are significantly different [e.g. brain versus muscle in ^{31}P MRSI (9)] by decomposing the observed spectral data into two non-negative matrices [as indicated in eqn (2)].

A second problem with recovering biochemically meaningful spectral patterns in brain ^{1}H MRSI is interference from large-amplitude and broad spectral resonances arising from tissue water and pericranial lipids. This is problematic because of the large dynamic range between the lipid and water resonance amplitudes and the subtle differences in brain tissue types that one seeks to recover using cNMF. Simply increasing the number of source components does not result in robust source recovery, and may lead to results with only diluted patterns, and not well-separated spectral patterns with high specificity.

Computationally, the cNMF algorithm is fast and efficient (10). The runtime for a complete implementation of cNMF with active data selection algorithm was typically less than 100 s (CPU: 2.4 GHz Pentium) for each of the clinical data sets presented in this study, with ~50% of this computation time being used in the active data selection of the algorithm. The algorithmic and computational details of NMF (15,16) and cNMF (9,10) have been described previously and compared with similar blind source separation methods, such as Bayesian spectral decomposition (13) and non-negative sparse coding (17). We thus provide here in Tables 1 and 2 only summaries of the basic cNMF algorithm and its extension

DOI: 10.1002/nbm

Table 1. Procedure for cNMF without data selection

1	Initialize: choose dimensions of **A** and **S** (i.e. M) and initialize with non-negative values (e.g. random **A** and constrained least-squares for **S**)
2	Update **A**
3	Force negative values of **A** to be approximately zero
4	Update **S**
5	Force negative values of **S** to be approximately zero
6	Iterate (back to 2) until convergence

Table 2. Procedure for cNMF with data selection

1	Set $\hat{\mathbf{X}} = \mathbf{X}$ (use all voxels)
2	Apply cNMF to the voxels $\hat{\mathbf{X}}$
3	Analyze peak resonance bands of the spectra and prior information of the spatial distribution to select a target source m
4	Construct a spatial mask **t**: thresholding the concentration matrix to select voxels which have a significant concentration of the target spectrum for $A_{i,m},\ i=1,\ldots,N$, $t_i = \begin{cases} 1 \ if \ A_{i,m} > threshold \\ 0 \ if \ A_{i,m} \le threshold \end{cases}$, and apply morphology analysis to improve the spatial mask **t** by removing small objects, smoothing edges, filling 'holes', etc.
5	Voxels pass through the mask[1]: $[nzr, nzc] = find(\mathbf{X}^*\ repmat(\mathbf{t}, 1, L));$ $\hat{\mathbf{X}} = \mathbf{X}(nzr(1 : \sum \mathbf{t}), :)$
6	Continue to 2 if additional specificity is required, else stop.

[1]Note in item 5 which is in Matlab notation, t is a column vector, the length of which equals the number of observation samples in the spectroscopic image.

with iterative data selection for automatic artifact removal. In addition, a brief geometric interpretation of the algorithm is provided in Fig. 1.

In vivo human brain ^{1}H MRSI data

All the presented data were acquired with institutional review board approval and informed consent, on a 1.5 T GE Signal Horizon 5.x or 3.0 T LX MR system, using the multi-slice MRSI sequence of Duyn *et al*. (18). After acquisition of localizer MRI series, four-section ^{1}H MRSI scans were performed, with 15 mm section thickness, 3.5 mm gaps, *TE*/*TR* 280/2300 ms, field of view 240 mm, 32×32 phase-encoding steps, with circular k-space sampling and full echo acquisition. These sequence parameters were identical for the two field strengths except that 512 sample points and a spectral width of 2500 Hz were used at 3.0 T, whereas 256 sample points and a spectral width of 1000 Hz were used at 1.5 T. Before analysis of the recorded MRSI data with cNMF, the following pre-processing steps were performed using spectroscopic imaging data analysis software developed in-house by two of the investigators (X.M., D.C.S.): the raw data were separated into individual slices, zero-filled once along the spatial domains to a 64×64 data matrix, and filtered with a shifted Gauss–Lorentz window along the time domain and a Hamming window and Fermi filter along spatial domains. These pre-conditioned time-domain data were then processed by standard three-dimensional fast Fourier transformation, followed by automatic correction of the resulting spectra for phase shifts, as well as for susceptibility-induced chemical-shift variations caused by static magnetic field inhomogeneity across the brain.

RESULTS AND DISCUSSION

The potential clinical utility of cNMF in recovering biochemically and physiologically meaningful spectral patterns of the normal and diseased human brains will now be demonstrated for (a) normal brain, (b) multiple sclerosis (MS), and (c) serial brain tumor scans. Each case was selected retrospectively to illustrate one or more properties of cNMF, as well as its broad utility. We should point out that: (a) the variations in pixel intensities in the spatial distribution of cNMF-recovered images presented below correspond only to the variations in the spatial concentration/abundance of a specific 'spectrum' recovered by cNMF, i.e. from spatial location to spatial location; (b) relative amplitudes of the resonances [e.g. relative concentrations of *N*-acetyl-L-aspartate (NAA), total choline (tCho), and total creatine (tCr)] in a cNMF-recovered 'spectrum' remain the same, and only the amplitude (concentration) of that entire pattern may increase or decrease. This is in contrast with conventional metabolite images in which the amplitudes or concentrations of individual resonances within a spectrum can vary from voxel to voxel across the brain. For this reason, we have used the term 'spectral pattern' rather than 'spectrum' to designate cNMF-recovered spectral information.

Normal human brain

^{1}H MRSI data for normal human brain provide normative spectral patterns against which a variety of suspected brain metabolite changes can be assessed. This subsection presents the results of using cNMF to automatically extract the ^{1}H MRSI spectral patterns from a normal brain data set acquired at 3.0 T. The extracted spectral patterns are shown in Fig. 2(a–c). In the absence of significant metabolite concentration differences across the investigated normal brain slice, cNMF has segmented the underlying brain tissue on the basis of relative local

DOI: 10.1002/nbm

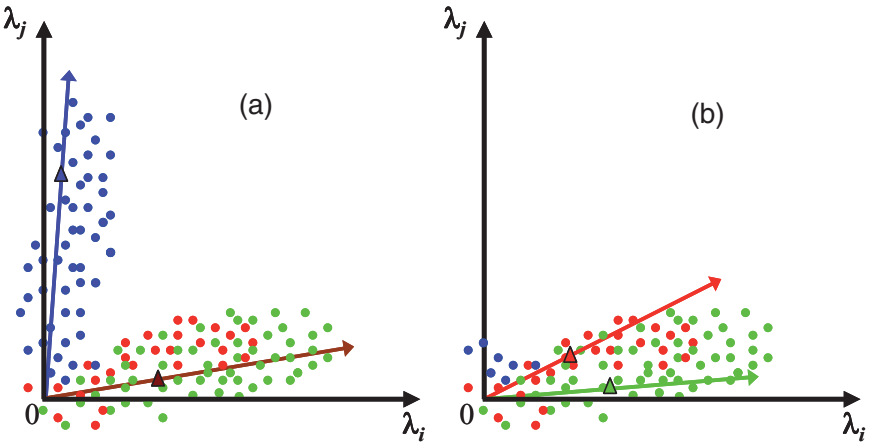

Figure 1. Geometric interpretation of cNMF with data selection. The axes in (a) and (b) are two selected resonances. In (a) the blue points are 'mostly lipid' voxels, the red points are 'mostly tumor' voxels, and the green points are 'mostly normal brain' voxels. The blue triangle is the recovered source spectrum or spectral pattern corresponding to 'lipids' and the blue arrow is its span. The brown triangle represents a mixture of all the other recoverable spectral patterns, which in this case, would correspond to those lying within the brain, with the brown arrow being their span. Owing to the strong lipid resonance interference, differences between these other recoverable spectral patterns, i.e. those representing normal brain and tumor, cannot be separated. This is because in the data space, normal brain and tumor spectral patterns are located close to one another in relation to the strong lipid resonances. In (b), after removal of the strong lipid contribution by active data selection, cNMF can be applied again to recover the 'tumor' and 'normal brain' spectral patterns with greater specificity (the red and green triangles, with the red and green arrows representing their spans, respectively).

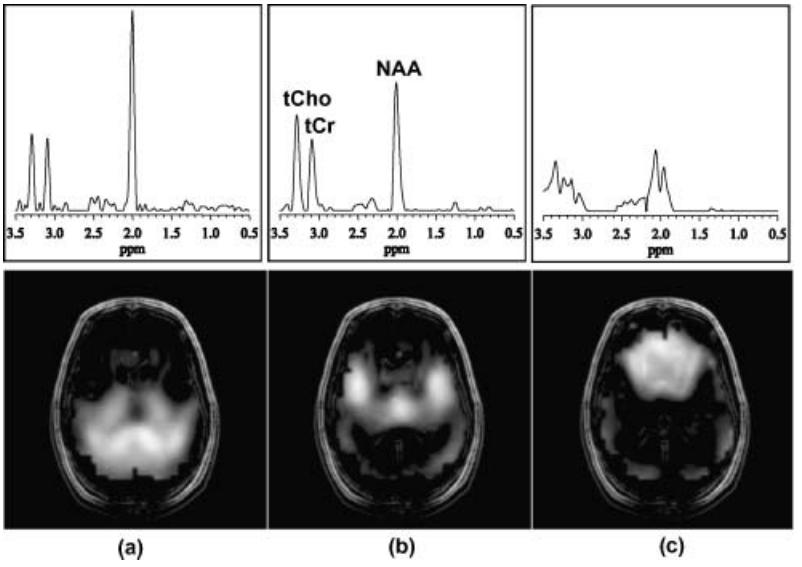

Figure 2. Results of using cNMF to analyze a normal human brain. In the absence of large spectral differences across the brain, cNMF has recovered spectral patterns that reflect relative local magnetic susceptibility effects and associated field inhomogeneity. There is a clear posterior-to-anterior gradient in magnetic field homogeneity, with the highest homogeneity occurring in the occipital lobe region (a), followed by a slightly poorer homogeneity in the middle of the brain (b), and a completely degraded field homogeneity in the frontal area (c) because of the high tissue/air magnetic susceptibility effects associated with the underlying nasal cavity and frontal sinuses.

DOI: 10.1002/nbm

magnetic susceptibility effects and associated field inhomogeneity. In this particular normal human brain, there is clearly a posterior-to-anterior gradient in magnetic field homogeneity. This is consistent with experimental observations that the magnetic field homogeneity in the occipital lobe region (Fig. 2a) is usually higher or easier to optimize than that for the frontal area (Fig. 2c), which is generally quite poor and difficult to optimize because of the high tissue/air magnetic susceptibility effects associated with the underlying nasal cavity and frontal sinuses. The spectral pattern for the region between the latter two regions exhibits intermediate homogeneity and can be judged to be still useable, whereas those for the frontal area would have to be excluded from any meaningful analysis. Although brain metabolite differences have been reported between gray and white matter (19), we should point out that, in this normal brain (Fig. 2), these gray matter/white matter differences are relatively small compared with the gross spatial heterogeneity across the brain caused by static magnetic field inhomogeneity, and have thus been masked. These results demonstrate a useful feature of cNMF: it can be used as a pre-processor for conventional peak-fitting methods for quickly defining regions of interest that should be included or rejected from further analysis on the basis of spectral quality, thereby saving time.

Multiple sclerosis

MS is a disorder of the central nervous system in which neural impulse conduction is impaired because of axonal myelin loss or damage. The damaged areas form lesions or plaques, which can usually be detected with high sensitivity by contrast-enhanced MRI (20,21). However, the specificity of MRI for MS is somewhat lower, as there are often plaque-like lesions, especially in older patients, which may not be due to MS. In addition, although acute MS lesions can be contrast-enhanced on MRI, chronic lesions usually cannot. Therefore, a definitive diagnosis of MS often requires additional evaluations, including assessment of clinical symptoms and signs, evoked potential measurements, and analysis of cerebrospinal fluid and blood. ^{1}H MRSI has been shown to be sensitive to some MS changes (22,23) and, therefore, could serve as an additional non-invasive confirmatory test which can be administered in conjunction with MRI. Here, we present the results of automated cNMF analysis of a 1.5 T multi-slice ^{1}H MRSI study of a case of MS that consisted of both acute and chronic lesions. The study shows the ability of cNMF to segment MS lesions to uncover spectral changes that are not specifically detected or are completely missed by conventional MRSI data analysis techniques.

Figures 3 and 4 show results of the automated analysis of two slices in the brain of a patient with confirmed acute

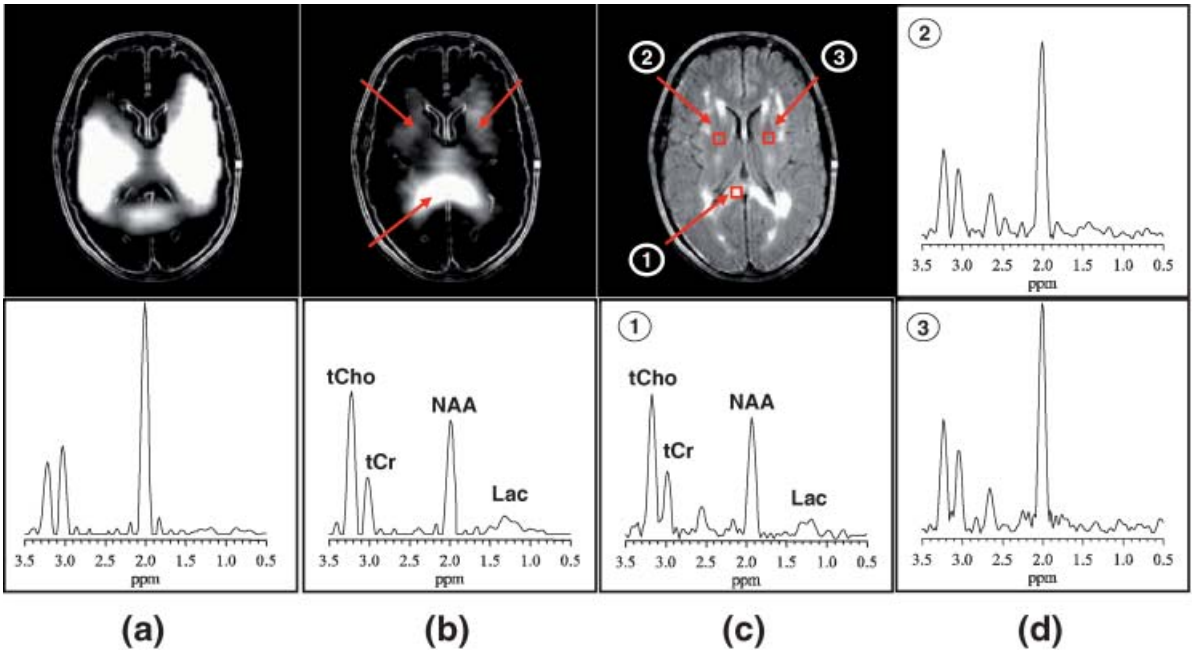

Figure 3. cNMF analysis of an acute MS case. The recovered spectral patterns (bottom row) and corresponding spatial concentration distributions (top row) are associated with two types of tissue: (a) normal brain, characterized by normal tCho, tCr and NAA concentrations, and (b) chronic MS lesion characterized by raised tCho and Lac, and decreased NAA resonances. Conventional ^{1}H MRSI spectra from voxels (red squares on FLAIR image) (c) in the lesion of the splenium of the corpus callosum and (d) bilaterally in normal-appearing tissue are shown for comparison. Note that the cNMF-recovered spectral pattern (b, bottom) is nearly identical with the conventional spectrum (c, bottom) from the voxel (①) in the corpus callosum lesion, but completely different from conventional spectra in voxels (d; ② and ③) for the normal-appearing brain tissue, which exhibit nearly normal relative metabolite levels. This shows the ability of cNMF to recover potentially diagnostic information that can be missed by conventional MRSI data analysis methods.

DOI: 10.1002/nbm

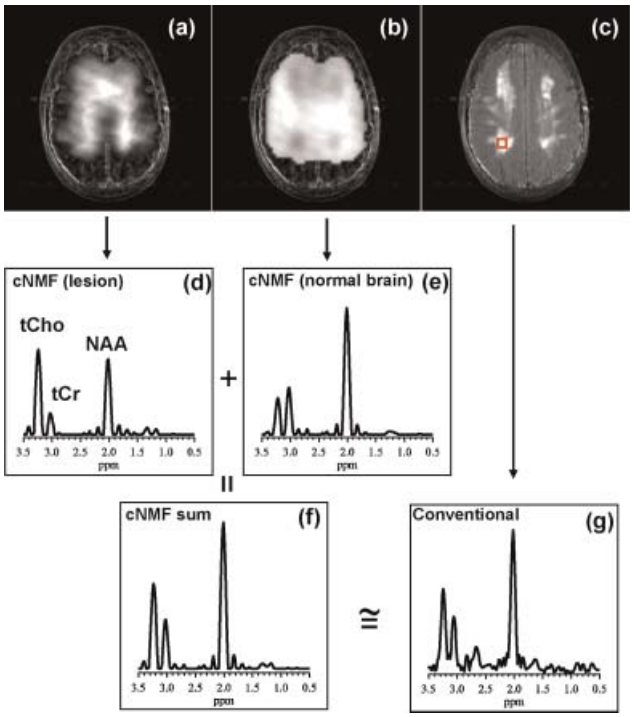

Figure 4. cNMF analysis of a chronic MS lesion. Shown are the spatial metabolite concentration distribution (a) and spectral pattern (d) for the MS lesion recovered by cNMF. The cNMF-recovered spectral pattern is markedly abnormal compared with a conventional spectrum (g) obtained from a voxel located in the center of one of the chronic lesions [box on FLAIR image in (c)]. Note that the spatial distribution of the cNMF-recovered abnormal spectral pattern (a) encompasses and extends beyond the regions of abnormal signal intensity on the FLAIR image (c), suggesting that the tissue 'segmentation' property of cNMF is able to more accurately delineate the extent of the chronic lesions. In addition to the MS lesion, cNMF has also recovered the spectral pattern (e) and spatial concentration distribution (b) for normal brain tissue, which can be clearly seen to overlap with that of the MS lesions (a). This overlap, in fact, explains the relatively slight spectral changes in the conventional spectrum (g) compared with the cNMF-recovered spectral pattern (d): the conventional spectrum (g) is the result of the partial volume averaging of the overlapping normal brain (e) and MS lesion (d) spectral patterns. This can be seen in (f) where a summation of the MS lesion and normal brain spectral patterns (d + e) yields a composite cNMF spectral pattern (f) that is nearly identical with the conventional spectrum (g). This figure is available in colour online at www.interscience.wiley.com/journal/nbm.

(Fig. 3) and chronic (Fig. 4) MS lesions, with comparisons with standard spectra from voxels in the center of each of the two types of lesion, as well as with those from normal-appearing brain tissue. cNMF analysis of the slice containing the acute MS plaques has recovered the spatial distribution, relative concentration, and spectral patterns associated with two tissue types: normal brain (Fig. 3a) and the MS lesion (Fig. 3b). The extracted MS spectral pattern showing (Fig. 3b, bottom) raised tCho and lactate (Lac) and decreased NAA is consistent with previously published spectra of acute MS lesions (22). For direct comparison, a conventional ^{1}H MRSI spectrum extracted from a voxel in the splenium of the corpus callosum lesion (Fig. 3c, bottom; voxel location ① on the FLAIR image in Fig. 3c, top) is plotted adjacent to the cNMF-recovered spectral pattern (Fig. 3b, bottom). The two spectra are virtually identical, indicating that both cNMF and the conventional MRSI data analysis method have recovered a spectrum that accurately reflects the underlying tissue type. This is not surprising because of the relatively large size of this acute MS lesion in the corpus callosum, as significant partial volume effects, which would have altered the appearance of the conventional spectrum, are not expected. On the other hand, whereas the cNMF-recovered spectral pattern intensity 'map' (Fig. 3b, top) shows the presence of a significant MS lesion pattern even in regions that appear normal on the FLAIR image (indicated by red arrows in Fig. 3b, top, and labeled as voxels ② and ③ in Fig. 3c, top), spectra obtained conventionally from these two regions near the anterior horn of the lateral ventricle appear completely normal, with a robust NAA peak and no lactate peak (Fig. 3d, ② and ③). cNMF is, therefore, able to recover abnormal spectral patterns from normal-appearing tissues, which would otherwise be missed by conventional MRSI data analysis methods. Thus, in cNMF, the true extent of a lesion can be more accurately and reliably visualized, as all tissue types not conforming to the lesion pattern are automatically removed, leaving only the concentration and spatial distribution of the lesion. This demonstration of the built-in *tissue segmentation* property of the method is dramatically illustrated and explained for the analysis of the chronic MS lesion described below.

The results of using cNMF to analyze the ^{1}H MRSI data of a slice containing chronic MS lesions from the same patient as in Fig. 3 are shown in Fig. 4. The cNMF-recovered spectral pattern (Fig. 4d) has been compared with a conventional spectrum (Fig. 4g) obtained from a voxel located in the center of one of the chronic lesions (see voxel location in Fig. 4c). Although the FLAIR MR image (Fig. 4c) shows the lesions to be extensively distributed throughout this brain slice, the conventional spectrum from the lesion (Fig. 4g) is only slightly abnormal, with a moderate increase in tCho, an almost insignificant increase in Lac and decrease in NAA. On the other hand, the cNMF spectral pattern recovered for the lesions is markedly abnormal, showing a very large decrease in NAA and increase in tCho, with a slight increase in Lac (Fig. 4d), *similar to the pattern shown for the large acute lesion* (Fig. 3b, bottom). Note that, as for the acute lesion, the spatial distribution of this abnormal spectral pattern for the chronic lesion (Fig. 4a) encompasses and then extends beyond the regions of abnormal signal intensity in the FLAIR image (Fig. 4c), further supporting our conclusion that the tissue segmentation property of cNMF more accurately delineates the extent of chronic lesions than conventional MRSI analysis.

DOI: 10.1002/nbm

In addition to recovering the distribution, concentration and spectral pattern associated with the chronic MS lesions, cNMF has also recovered another type of tissue, the spectral pattern of which (Fig. 4e) can be seen to be that of normal brain tissue, and the spatial distribution of which (Fig. 4b) overlaps with that of the MS lesions (Fig. 4a). The overlap of the spatial distributions of these two tissue types, brain and MS lesions, can, in fact, explain the relatively slight spectral changes seen in the conventional spectrum of the lesions (Fig. 4g): partial volume averaging of the spectral contributions from the MS lesions and normal brain has led to a conventional 'mild lesion' spectrum (Fig. 4g) which is a linear combination of the spectral patterns from the two tissue types, as shown in Fig. 4f. Note the great similarity between the composite cNMF spectral pattern in Fig. 4f (i.e., $4d + 4a \approx 4f$) and the conventional spectrum (Fig. 4g). This masking of potentially diagnostic changes by the partial volume effect in the conventional spectrum (Fig. 4g) is clearly avoided by the intrinsic tissue segmentation property of cNMF (Fig. 4d). These internally self-consistent results are both a demonstration and a validation of the ability of cNMF to identify some types of lesion better than conventional MRSI data analysis methods. Without the built-in tissue segmentation feature of cNMF, the true spectral pattern and spatial extent of the chronic MS lesion would have been missed. ^{1}H MRSI with cNMF analysis could thus be a powerful adjunct to MRI for clinical evaluation of suspected MS lesions.

Serial monitoring of brain tumor response to therapy

Although standard structural MRI is now firmly established as an effective non-invasive technique for monitoring the therapeutic response of brain tumors, it has limited sensitivity in distinguishing recurrent tumors from radiation-induced necrosis or post-resection changes such as edema. MRI, moreover, can often underestimate the true size of a tumor if the latter extends beyond the boundaries of the regions of contrast enhancement. In recent years, ^{1}H MRSI has been shown to have the potential to overcome these limitations of MRI in the evaluation of brain tumors (24,25). The results of this section demonstrate the usefulness of the cNMF technique in achieving blind recovery of the spatial distribution of tumor metabolite concentrations and spectral patterns that could be used to monitor the therapeutic response of brain tumor on the basis of serial ^{1}H MRSI studies.

Figure 5 shows spectral patterns and the associated spatial distributions recovered from three serial ^{1}H MRSI studies of a patient who had been diagnosed with an oligodendroglioma, and was undergoing a chemotherapy regimen. In Fig. 5, Baseline column, the cNMF-recovered data consist of a spectral pattern from normal brain tissue (Baseline, b1) and another pattern from a large occipital lobe mass (Baseline, b2), the spatial extent of which can be clearly visualized in panels a1 and a2, respectively, because the surrounding tissue has been 'neatly segmented away' by cNMF. The spectral pattern of the tumor shows the characteristic tumor pattern of increased tCho, decreased NAA, and increased Lac. After 20 months of chemotherapy, cNMF is applied to new data (shown in Fig. 5, Month 20) and recovers spectral patterns of normal brain tissue (Month 20, b3) and lesion (Month 20, b4), as well as their corresponding spatial distributions (Month 20, a3 and a4). However, whereas the normal brain pattern remained unchanged, there was a marked change in both the spectral pattern and spatial distribution for the lesion. A notable normalization of the tCho peak and decrease in Lac can be seen to have occurred as a result of chemotherapy. The distribution of this spectral pattern became more diffuse and less concentrated. These results are consistent with a marked improvement in the metabolic status of the tumor.

The cNMF results of the third study of the series, acquired 24 months after the baseline scan, are shown in Fig. 5, Month 24. The recovered lesion spectral pattern (Fig. 5, Month 24, b6) and spatial distribution (Fig. 5, Month 24, a6) can now be seen to have changed again compared with those for month 20. Whereas the lesion spectral pattern has reverted to that of a tumor (robust increase in tCho and decrease in NAA), the spatial distribution or extent of the lesion has dramatically decreased and is represented by a relatively small area in the left anterior horn of the lateral ventricle. Note also how cNMF has achieved tissue segmentation to permit unobstructed visualization of the location and extent of the lesion.

In summary, the presented cNMF-recovered serial brain tumor data demonstrate that all of the properties of cNMF (blind recovery of spectral pattern, concentration, and spatial distribution with automatic tissue segmentation to enhance visualization and minimize volume averaging) can be combined to achieve rapid and meaningful assessment of the response of brain tumors to therapy.

CONCLUDING REMARKS

Computer-aided analysis of MRSI data is becoming more important for realizing the broad clinical utility of the modality. The increased spatial and spectral resolution has resulted in very large data sets with a wealth of potentially diagnostic information and biomarkers of human disease. Moreover, there has been increased interest in using MRI-based tissue segmentation to correct MRSI data for the potential confounding effect of partial volume averaging [e.g. (19)]. In this study, we have described a method, cNMF with iterative data selection, that utilizes prior information about physical

DOI: 10.1002/nbm

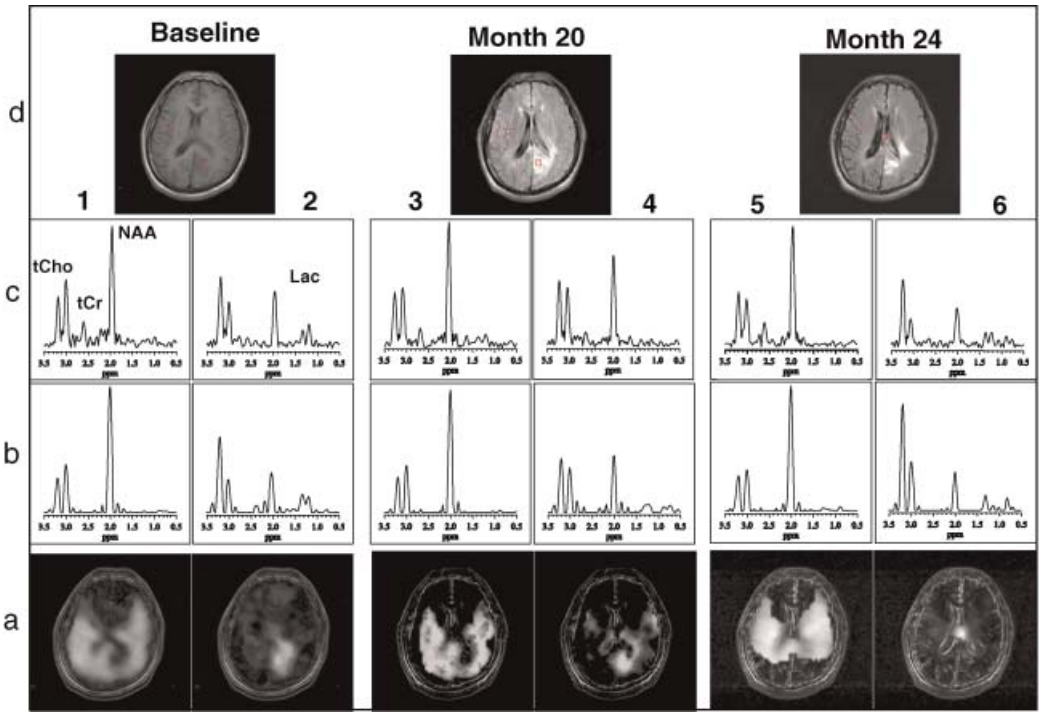

Figure 5. cNMF analysis of three serial ^{1}H MRSI studies of an oligodendroglioma for a patient undergoing chemotherapy treatment. Results are shown for MRSI data recorded at baseline, and at 20 and 24 months during the course of chemotherapy. (a) cNMF-recovered spectral pattern intensity distributions or maps; (b) the cNMF-recovered spectral patterns; (c) the conventional MR spectra from voxels depicted in the anatomic images shown in (d). There are columns of data, which are identified numerically for clarity. See the text for a complete description of the cNMF-recovered spectral patterns and spatial distributions. This figure is available in colour online at www.interscience.wiley.com/journal/nbm.

constraints and properties of spectral mixing, as well as properties of brain anatomy, to automatically and simultaneously achieve tissue segmentation and recovery of potentially diagnostic spectral patterns of the human brain in long-*TE in vivo* ^{1}H MRSI data. Unlike MRI-based tissue segmentation of MRSI data, which does not correct the appearance of the spectra for distortions due to partial volume averaging, and requires availability of high-resolution/high-contrast volumetric MR images, cNMF automatically achieves tissue segmentation during spectral pattern recovery. This yields spectral patterns that reflect the true characteristics of the underlying tissue, and, therefore, would be visually more diagnostic and suitable for rapid clinical assessment and interpretation of the underlying pathology (e.g., the chronic MS lesion case in Fig. 4). The data-selection procedure provides a means for unmasking spectra that might not otherwise be consistently recoverable because of interference from the strong signals of residual lipids and water. The recovered biochemically meaningful spectral patterns can ultimately be used as patterns for classification and image labeling, to construct nosologic images (26), in which each voxel is pseudo-colored according to the estimated histopathological class. The presented sample applications have demonstrated how the complex diagnostic information contained in thousands of spectroscopic imaging data can be simplified by cNMF into only a few meaningful spectral patterns and associated spatial distributions which can be more conveniently presented to a trained diagnostician allowing rapid interpretation and assessment of the underlying pathology. This feature of cNMF, along with its computational efficiency, makes it ideally suited for use in clinical or diagnostic applications of MRSI.

In this study, we have limited our demonstration of cNMF to long-*TE* MRSI data and have not considered short-*TE* spectra, which are generally more complex because of a significant spectral overlap, a prominent baseline arising from macromolecules and lipids, and a multiplicity of spectral resonances. However, one aspect of our current research focuses on recovering potentially diagnostic spectral patterns for a variety of brain disorders in short-*TE* spectra, as such data would provide diagnostic information that is not present at longer *TE*s. It should be noted that cNMF with data selection holds even greater promise for analysis of short-*TE* spectra, because the amplitudes of resonances in such spectra would typically be positive and, thus, would generally be more compatible with the non-negativity constraints of the method.

DOI: 10.1002/nbm

The algorithm described here is very general and has been shown to be useful in assessing a variety of neurological disorders. In all cases considered, the fundamental mixing (eqn 1) is across space. However, the algorithms can be equally well applied to mixtures across time. For example, we have recently shown how the method is able to unmix 600 MHz ^{1}H NMR spectra of biofluids, such as blood and urine, to characterize the metabolic status of animals and identify biomarkers and time course of toxicity (27).

Acknowledgements

This research was supported by an NSF CAREER Award (BES-0133804, to P.S.), the DoD Multidisciplinary University Research Initiative (MURI) program administered by the Office of Naval Research (N00014-01-0625, to P.S.), and Cornell University Faculty Development Funds (to D.C.S.).

REFERENCES

1. Bottomley PA. Spatial localization in NMR spectroscopy *in vivo*. *Ann. NY Acad. Sci.* 1987; **508**: 333–348.
2. Frahm J, Bruhn H, Gyngell ML, Merboldt KD, Hänicke W, Sauter R. Localized high resolution proton NMR spectroscopy using stimulated echoes: initial applications to human brain *in vivo*. *Magn. Reson. Med.* 1989; **9**: 79–93.
3. Brown TR, Kincaid BM, Ugurbil K. NMR chemical shift imaging in three dimensions. *Proc. Natl. Acad. Sci. USA.* 1982; **79**: 3523–3526.
4. Maudsley AA, Hilal SK, Perman WH, Simon HE. Spatially resolved high resolution spectroscopy by four-dimensional NMR. *J. Magn. Reson.* 1983; **51**: 147–152.
5. Rudkin TM, Arnold DL. Proton magnetic resonance spectroscopy for the diagnosis and management of cerebral disorders. *Arch. Neurol.* 1999; **56**: 919–926.
6. Burtscher IM, Holtås S. Proton, MR spectroscopy in clinical routine. *J. Magn. Reson. Imaging.* 2001; **13**: 560–567.
7. Mierisova S, Ala-Korpela M. MR spectroscopy quantitation: a review of frequency domain methods. *NMR Biomed.* 2001; **14**: 247–259.
8. Vanhamme L, Sundin T, Hecke PV, Huffel SV. MR spectroscopy quantitation: a review of time-domain methods. *NMR Biomed.* 2001; **14**: 233–246.
9. Sajda P, Du S, Brown TR, Parra LC, Stoyanova R. Recovery of constituent spectra in 3D chemical shift imaging using non-negative matrix factorization. In *Proc. 4th Int. Symp. ICA and BSS.* 2003; 71–76.
10. Sajda P, Du S, Brown TR, Stoyanova R, Shungu DC, Mao X, Parra LC. Non-negative matrix factorization for rapid recovery of constituent spectra in magnetic resonance chemical shift imaging of the brain. *IEEE Trans. Med. Imag.* 2004; **23**: 1453–1465.
11. Hyvärinen A, Karhunen J, Oja E. *Independent component analysis*. Wiley-Interscience: New York, 2001.
12. Cardoso JF. Blind signal separation: statistical principles. In *Proc. IEEE, Special Issue on Blind Identification and Estimation*1998; **9**: 2009–2025.
13. Ochs MF, Stoyanova RS, Arias-Mendoza F, Brown TR. A new method for spectral decomposition using a bilinear Bayesian approach. *J. Magn. Reson.* 1999; **137**: 161–176.
14. Ladroue C, Howe FA, Griffiths JR, Tate AR. Independent component analysis for automated decomposition of *in vivo* magnetic resonance spectra. *Magn. Reson. Med.* 2003; **50**: 697–703.
15. Lee DD, Seung HS. Learning the parts of objects by non-negative matrix factorization. *Nature* 1999; **401**: 788–791.
16. Lee DD, Seung HS. Algorithms for non-negative matrix factorization. *Adv. Neural Info. Proc. Syst.* 2001; **13**: 556–562.
17. Hoyer PO. *Non-negative sparse coding*. Neural Networks for Signal Processing XII. *Proc. IEEE workshop on Neural Networks for Signal Processing*. Martigny: Switzerland 2002; 557–565.
18. Duyn JH, Gillen J, Sobering G, van Zijl PCM, Moonen CT. Multisection, MR spectroscopic imaging of the brain. *Radiology* 1993; **188**: 277–282.
19. Hetherington HP, Pan JW, Mason GF, Adams D, Vaughn JT, Twieg DB, Pohost GM. Quantitative ^{1}H spectroscopic imaging of human brain at 4.1 T using image segmentation. *Magn. Reson. Med.* 1996; **36**: 21–29.
20. Wallace CJ, Seland TP, Fong TC. Multiple sclerosis: the impact of MR imaging. *AJR Am. J. Roentgenol.* 1992; **158**: 849–857.
21. Miller DH, Grossman RI, Reingold SC, McFarland HF. The role of magnetic resonance techniques in understanding and managing multiple sclerosis. *Brain* 1998; **121**: 3–24.
22. Pan JW, Coyle PK, Bashir K, Whitaker JN, Krupp LB, Hetherington HP. Metabolic differences between multiple sclerosis subtypes measured by quantitative MR spectroscopy. *Multiple Sclerosis* 2002; **8**: 200–206.
23. Simone IL, Tortorella C, Federico F. The contribution of ^{1}H magnetic resonance spectroscopy in defining the pathophysiology of multiple sclerosis. *Ital. J. Neurol. Sci.* 1999; **20**: S241–S245.
24. Nelson SJ, Vigneron DB, Dillon WP. Serial evaluation of patients with brain tumors using volume MRI and 3D ^{1}H MRSI. *NMR Biomed.* 1999; **12**: 123–138.
25. Balmaceda C, Critchell D, Mao X, Cheung K, Pannula S, Delapaz RL, Shungu DC. Multisection ^{1}H magnetic resonance spectroscopic imaging assessment of glioma response to chemotherapy. *J. Neuro-Oncology.* 2005; **76**: 185–191.
26. Edelenyi FSD, Rubin C, Esteve F, Grand S, Decorps M, Lefournier V, Bas JFL, Remy C. A new approach for analyzing proton magnetic resonance spectroscopic images of brain tumors: nosologic images. *Nature Med.* 2000; **6**: 1287–1289.
27. Du S, Sajda P, Stoyanova R, Brown TR. Recovery of metabolomic spectral sources using non-negative matrix factorization. *Proc. 27th Annual Int. Conf. EMBC*. 2005; **2**: 1095–1098.

DOI: 10.1002/nbm