RESEARCH ARTICLE

WILEY **NMR** IN BIOMEDICINE

Analysis of metabolic abnormalities in high-grade glioma using MRSI and convex NMF

Nuno Pedrosa de Barros[1] | Raphael Meier[1] | Martin Pletscher[1] | Samuel Stettler[1] | Urspeter Knecht[1] | Mauricio Reyes[2] | Jan Gralla[1] | Roland Wiest[1] | Johannes Slotboom[1]

[1] Support Center for Advanced Neuroimaging (SCAN), Neuroradiology, University Hospital Inselspital, Bern, Switzerland

[2] Institute for Surgical Technology and Biomechanics (ISTB), University of Bern, Bern, Switzerland

Correspondence
Nuno Pedrosa de Barros, Support Center for Advanced Neuroimaging (SCAN), Neuroradiology, University Hospital Inselspital, Bern, Switzerland.
Email: nunopbarros@gmail.com

Clinical use of MRSI is limited by the level of experience required to properly translate MRSI examinations into relevant clinical information. To solve this, several methods have been proposed to automatically recognize a predefined set of reference metabolic patterns. Given the variety of metabolic patterns seen in glioma patients, the decision on the optimal number of patterns that need to be used to describe the data is not trivial. In this paper, we propose a novel framework to (1) separate healthy from abnormal metabolic patterns and (2) retrieve an optimal number of reference patterns describing the most important types of abnormality. Using 41 MRSI examinations (1.5 T, PRESS, T_E 135 ms) from 22 glioma patients, four different patterns describing different types of abnormality were detected: *edema*, *healthy without Glx*, *active tumor* and *necrosis*. The identified patterns were then evaluated on 17 MRSI examinations from nine different glioma patients. The results were compared against BraTumIA, an automatic segmentation method trained to identify different tumor compartments on structural MRI data. Finally, the ability to predict future contrast enhancement using the proposed approach was also evaluated.

KEYWORDS

applications, cancer, head and neck cancer methods and engineering, MRS and MRSI methods, post-acquisition processing, spectroscopic imaging, visualization methods and engineering

1 | INTRODUCTION

MRS provides relevant metabolic information for the assessment of brain tumors, allowing us to distinguish different tumor types and grades,[1-3] distinguish radiation effects (pseudoprogression) from true progression[4,5] and identify regions with high tumor cellularity that are not visible in structural MRI.[6-8] Many publications[9] focus on the translation of one or two MRS features, such as choline (Cho)/NAA (N-acetyl aspartate) and Cho/Cr (creatine), into clinically meaningful information for the tasks described above. Regardless of what can already be achieved with the analysis of individual metabolite ratios, the use of metabolic patterns for the identification of tissue types and diseases has the potential to allow more precise characterization of brain tumors.

Several approaches[10-21] have been suggested for the analysis of metabolic patterns present in brain tumor MRS data. Before such methods can be used to interpret new data, a library of metabolic patterns has to be defined. The decision on the order of the model used to interpret

the data, ie the number of tissue patterns that are included in the reference pattern library, is one of the main challenges of these approaches. In all the above-mentioned methods, the decision on the number of patterns to use is based on assumptions made by the authors. When multiple tumor types/grades are included, often one pattern per tumor type or grade is assumed, as for instance in Reference [10 or 13]. This approach provides interesting results for the identification of the tumor grade and type but is not adequate to describe brain tumors with very distinct tumor compartments, such as GBM. Other methods[11,14] consider only two metabolic patterns, focusing solely on the separation between tumor and healthy brain tissue in each MRSI examination. The use of only two classes is motivated by the robustness of the differentiation between the average healthy brain tissue pattern and the average tumor pattern, which have very distinct features. This approach is suitable for tumor detection and gross tumor volume segmentation, but, similarly to the previous one, fails to describe the tumor heterogeneity.

To combine the robustness of the binary tissue-type separation with a more complete description of the tumor environment, Li et al[16] proposed hierarchical non-negative matrix factorization (hNMF), where at each level the spectra are decomposed in two different sources. In this method, healthy brain tissue is separated from the abnormal voxels in a first step, and in a second step the abnormal voxels are separated into tumor and necrosis. Following this work, Sauwen et al[18] applied hNMF to multi-parametric MRI data, including MRS, of glioma patients. The inclusion of different MR modalities allowed for the identification of multiple distinct patterns (tumor, necrosis, edema, cerebral-spinal fluid (CSF), white matter, gray matter and blood vessels), enabling an interesting description of the tumor heterogeneity in glioma patients. Despite this, the segmentation of each imaging study depended on the input of an experienced neuroradiologist, who needed to identify the tissues present in each examination. In order to simplify the manual input required for the segmentation, the same author proposes in Reference 21 a semi-automated approach where the user is asked to identify "seed" points for each tissue type.

In this paper, we propose a novel framework to (1) distinguish abnormal from healthy metabolic patterns and (2) retrieve an *optimal* number of reference patterns describing the most important types of abnormality present in an MRSI dataset. The method is based on convex non-negative matrix factorization (cNMF)[14,22-25] and starts by identifying a suitable basis set that provides a "good" description of the different metabolic patterns characteristic of healthy brain tissue. This is performed using a dataset collected from healthy volunteers. Then, spectra from brain tumor patients are classified into healthy or abnormal depending on how well they can be reconstructed using healthy sources. The subset of the abnormal spectra is then used to identify the main metabolic abnormalities seen in brain tumor data. The decision on the order of the model used is made as the "best" compromise between solution stability and reconstruction residue.

The proposed method was used to identify the main types of metabolic patterns seen in high-grade gliomas (HGGs). The spatial distribution of the metabolic abnormalities was compared with the tumor segmentation performed by BraTumIA (Brain Tumor Image Analysis),[26,27] which is an automatic brain tumor segmentation method trained to reproduce the manual MRI segmentation performed by experienced neuroradiologists. Besides this, we also include an analysis of two follow-up cases where the MRSI maps are compared with the corresponding structural MRI acquired several months after, showing the ability of MRSI to predict future contrast enhancement in HGG.

2 | METHODS

2.1 | Data

Table 1 shows the details of the glioma patients considered in this study, including genetic information regarding IDH1, 1p/19q loss of heterozygosity (LOH) and MGMT. A total of 58 MRSI examinations (PRESS, CHESS water suppression, T_E 135 ms, T_R 1500 ms) were collected from 31 different patients. Data acquisition was performed using two 1.5 T Siemens scanners (models Aera and Avanto, Siemens, Erlangen, Germany). The MRSI data had an original resolution of 12 × 12, with a voxel size of 13.33 × 13.33 × 15 mm^3, and were interpolated before DICOM storage to a resolution of 32 × 32 and a voxel size of 5 × 5 × 15 mm^3. All spectra from outside the PRESS box, as well as one single outer voxel row of the excited PRESS volumes, were discarded, given that often unreliable quantification results are found on the edges of the PRESS box due to partial excitation. The data was separated into a training dataset, containing 41 MRSI studies from 22 different patients (7624 spectra), and a test dataset, with the remaining 17 cases of 9 other patients (3276 spectra). The two datasets had similar characteristics, as seen in Table 2. Besides spectroscopy, the protocol included acquisition of T_1-weighted (MP-RAGE, T_E 2.67 ms, T_R 1580 ms, T_I 900 ms; 1 × 1 × 1 mm^3), T_1-weighted gadolinium-enhanced (MP-RAGE, T_E 4.57 ms, T_R 2070 ms, T_I 1100 ms; 1 × 1 × 1 mm^3), T_2-weighted (SPACE, T_E 380 ms, T_R 3000 ms; 1 × 1 × 1 mm^3) and T_2-weighted FLAIR MRI (T_E 88 ms, T_R 8000 ms, T_I 2500 ms; 1 × 1 × 3.3 mm^3).

In addition to the data collected from glioma patients, seven MRSI examinations were recorded from three healthy volunteers using the same acquisition parameters. This healthy control dataset contained 1372 spectra after excluding the voxels of the edges and from outside the PRESS box.

The measurements were performed in conformance with local and national ethical regulations, meaning that all patients gave their written informed consent for the data to be used for scientific purposes.

TABLE 1 Details of the HGG patients included in this study. The table shows the number of studies from each patient that were included in the training and test datasets. The age at the time of the first examination is shown in the table. For each patient, information regarding the R132H-IDH1 mutation status was included (no other codons of the IDH1 gene were tested). The LOH 1p/19q and the MGMT promoter methylation was also evaluated for the majority of the patients. The final column shows the number of studies from each patient that were acquired after tumor resection

Dataset	Patient	No. of studies	Gender	Age	Grade	R132H-IDH1	LOH 1p/19q	MGMT	No. resected
Training	1	4	Female	37	IV	Positive	Positive	Positive	2
Training	2	3	Male	65	IV	Negative	Positive	Negative	0
Training	3	2	Male	67	IV	Negative	Positive	Positive	0
Training	4	3	Female	53	IV	Negative	Negative	Negative	1
Training	5	2	Male	67	IV	Negative	Negative	Negative	1
Training	6	2	Male	68	IV	Negative	Negative	Negative	0
Training	7	1	Female	70	IV	Negative	—	—	1
Training	8	1	Female	68	IV	Negative	Positive	Negative	1
Training	9	2	Male	70	IV	Negative	Negative	Negative	1
Training	10	2	Female	70	IV	Negative	Negative	Positive	0
Training	11	1	Male	43	IV	Negative	—	Positive	1
Training	12	2	Male	55	III	Positive	Negative	Positive	0
Training	13	3	Female	67	IV	Negative	Negative	Positive	2
Training	14	1	Female	64	IV	Negative	—	Positive	0
Training	15	1	Male	59	IV	Negative	—	Negative	0
Training	16	1	Male	77	IV	Negative	Negative	Positive	0
Training	17	1	Male	56	IV	Negative	Negative	Positive	0
Training	18	1	Male	79	IV	Negative	Positive	Positive	0
Training	19	4	Female	60	IV	Negative	Positive	Positive	3
Training	20	1	Male	73	IV	Negative	Negative	Negative	0
Training	21	1	Male	46	IV	Negative	Positive	Negative	0
Training	22	2	Male	44	III	Negative	—	—	2
Test	23	4	Male	62	IV	Negative	Positive	Positive	1
Test	24	1	Female	76	IV	Negative	—	—	0
Test	25	3	Female	58	IV	Negative	Negative	Positive	1
Test	26	3	Female	58	IV	Negative	-	Negative	0
Test	27	2	Female	58	III	Positive	Positive	Positive	1
Test	28	1	Male	70	IV	Negative	Negative	Positive	1
Test	29	1	Male	76	IV	Negative	Negative	Positive	1
Test	30	1	Female	70	IV	Negative	Negative	Negative	0
Test	31	1	Male	79	IV	Negative	Negative	Negative	0

TABLE 2 Summary table of the data shown in Table 1 comparing the details of the training and test datasets

Dataset	No. of patients	No. of studies	% male	Ave. age ± SD	% grade III	% IDH pos.	% LOH pos.	% MGMT pos.	% resect.
Training	22	41	53.7	60.1 ± 11.5	9.8	14.6	45.7	57.9	36.6
Test	9	17	41.2	63.8 ± 7.4	11.8	11.8	46.2	68.8	29.4
All data	31	58	50.0	61.2 ± 10.5	10.3	13.8	45.8	61.1	34.5

% male—percentage of studies from male patients; Ave. age ± SD—average age (± standard deviation) per study in each dataset; % Grade III—percentage of studies with Grade III tumors; % (IDH/LOH/MGMT) pos.—percentage of studies with a positive result for each genetic characteristic; % resect.—percentage of studies acquired after tumor resection.

2.2 | MRS data pre-processing pipeline

All spectra were pre-processed using jMRUI's SpectrIm plugin.[28] The pre-processing pipeline consisted of the following six processing steps, presented in the order in which they were executed.

1. *Residual water peak removal using HLSVD.*[29] HLSVD was applied considering eight components between 4.0 and 9.0 ppm.

2. *Automatic quality control.*[30-32] Quality control was performed using a random forest model that was trained to reproduce the judgment of an expert. The criteria for rejection of the training data were the presence of visible signal artifacts and low SNR. The trained model outputs a probability of acceptance by an expert. Therefore, after assessing all spectra with the automatic classifier, all spectra with a probability of acceptance by an expert below 0.5 were removed from the dataset. For more details on automatic quality filtering the reader is referred to References 30-32.

3. *Frequency-shift correction.*[33]

4. *Auto-phasing.*[33]

5. *Quantification.* Quantification was performed using QUEST[34] and a metabolite model that included the following metabolites: Cho, Cr, NAA, glutamate (Glu), glutamine (Gln), lactate (Lac), lipids 1.3 ppm (Lip1.3), lipids 0.9 ppm (Lip0.9), myo-inositol and glycine. The quantification results of the last two metabolites were not included in this paper for two different reasons. Myo-inositol was detected in most of the voxels but, given that a long T_E was used, the quantification results were very noisy and not very reliable. Glycine, on the other hand, was only detected in small amounts in two of the cases (Patients 12 and 27).

6. *Selection profile correction.*[32] Excitation profile imperfections of amplitude modulated RF pulses, which are used in the PRESS sequence, cause non-uniform spatial excitation. Additionally, the scanners' limited gradient and RF power cause spins resonating at different chemical shifts to be excited in a spatially shifted manner, which is the so-called chemical shift displacement artifact. These effects can be corrected using reference phantom MRSI recordings performed on metabolite solutions that use the same acquisition parameters as the in vivo measurements. Before applying the correction, the values for each metabolite in the reference scan are normalized so that the mean of the voxels contained in the excited volume is 1. After normalization, selection profile correction is performed by dividing the quantification results by the corresponding value in the normalized reference scan. Average Cho/NAA maps before and after correction, for both healthy controls and glioma patients, are included in the Supplementary Material (Figure S1) to highlight the importance of this correction.

2.3 | Strategy used to determine healthy and HGG sources

The overall strategy used to retrieve the healthy and HGG sources from the pre-processed and quality-filtered data is depicted in Figure 1. The approach used starts by identifying a set of healthy sources capable of describing the variability of metabolic patterns seen in healthy brain tissue. Then, using non-negative least squares (NNLS),[35] we try to approximate the spectra collected from HGGs as a positive linear combination of the healthy sources and measure the error of the approximation. Spectra with large reconstruction error, ie that cannot be described as a combination of healthy sources, are classified as abnormal. Finally, the additional sources required to describe HGG data are determined using the set of abnormal spectra.

2.4 | cNMF

cNMF[14,22-25] was used to determine the signal sources present in healthy and HGG data. This blind source separation method identifies K patterns, or sources, whose positive linear combination can be used to reconstruct the different data points present in a given input dataset. The number K is user defined. The strategy used to find the optimal K is described in the next sections.

cNMF is a variation of NMF where the sources are determined as a convex combination of the input data points. Compared with conventional NMF, the sources derived from cNMF are often easier to interpret given that they are combinations of input data points, in this case, spectra. Moreover, cNMF becomes a natural choice for the first step of the algorithm where we aim to obtain a basis set that defines the convex set of healthy spectra.

In this paper, we used the NMF MATLAB toolbox.[36] cNMF was initialized using K-means and for each spectroscopic voxel the following features were included in the input data matrix: %NAA, %Cho, %Cr, %Glx, %Lac and %Lip. Here %metabolite represents the ratio between the area of the spectrum of each metabolite, as determined by QUEST, and the areas of all six metabolites considered. %Glx corresponds to the sum of %Gln and %Glu, and %Lip to the sum of %Lip1.3 and %Lip0.9. Glu and Gln were not included separately, given that, due to their overlap, the quantification of these metabolites on a spectrum by spectrum basis is prone to errors, which may affect negatively the stability of the cNMF results.
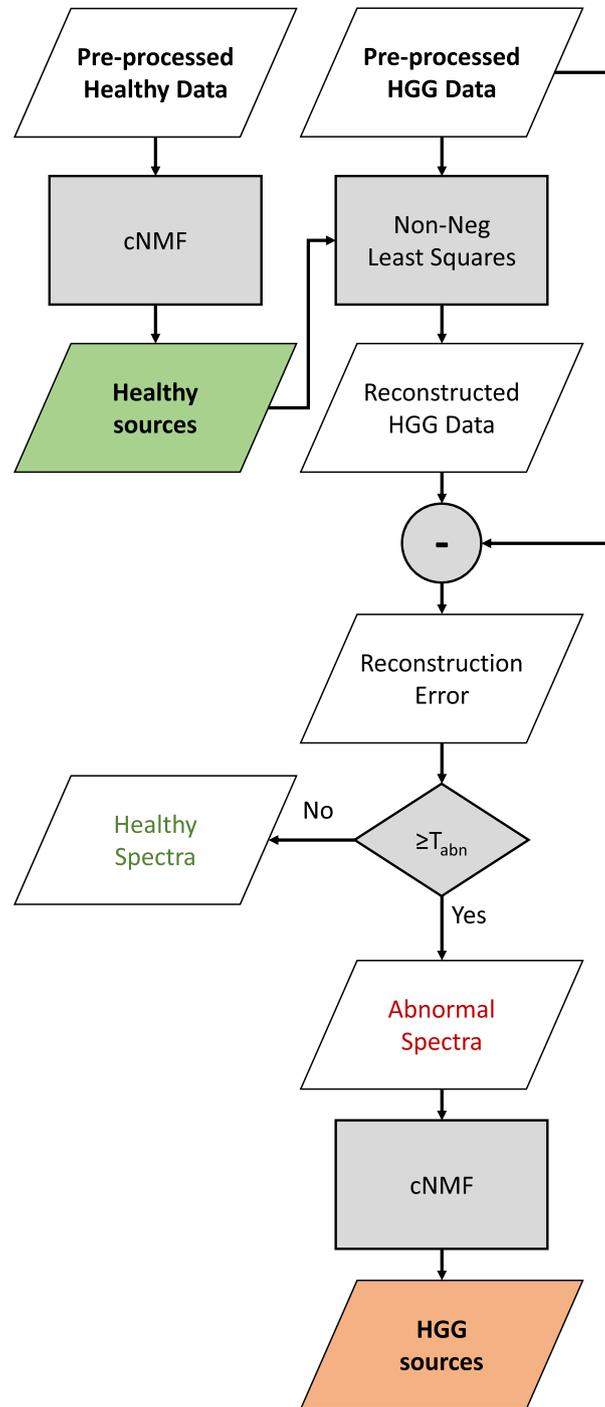
FIGURE 1   Workflow of the approach used to determine the healthy and abnormal sources from the training set. The workflow is executed after data pre-processing and quality control

2.5 | Determining the number of healthy sources

Identification of the set of healthy sources has the main purpose of defining the healthy region of the feature space. In this case, no meaning was assigned to each individual pattern. Therefore, the optimal number of healthy sources, K_H, was chosen simply as the value above which no significant decrease in the reconstruction error of healthy spectra is observed.

To determine K_H, leave-one-out cross-validation (LOOCV) was used, where each time the spectra from one healthy control were excluded from the cNMF's input data and used to measure the reconstruction error. The experiment was repeated 10 times for every value of the number of healthy sources (k_H) between 1 and 15. The reconstruction error as a function of k_H was calculated for each spectrum n of the validation set as

$$R^{n,l}(k_H) = \frac{1}{m}\sum_{i=1}^{m}\left|x_i^n - \sum_{k=1}^{k_H}a_k^n s_{i,k}^l\right|$$ (1)

where m is the number of features, l is a given cross-validation step, x_i^n is the value of feature i of spectrum n, $s_{i,k}^l$ is the value of feature i of the source k from a given solution l and a_k^n are the factors determined by NNLS.

After running the experiment and measuring the different $R^{n,l}(k_H)$ values, the optimal number of healthy sources, K_H, was chosen as the minimum value of k_H whose mean error is statistically not significantly different from the mean error of the maximum number of sources tested, which in this case was 15.

2.6 | Abnormality detection

After fixing the number of healthy sources, K_H, the abnormality threshold (T_{abn}) was defined as the 99th percentile of the measured reconstruction error with K_H sources. As explained above, these reconstruction errors were measured in the different validation sets. Therefore, a spectrum was classified as abnormal if the reconstruction error, using NNLS and the predetermined healthy sources, was greater than or equal to T_{abn}.

2.7 | Determining the number of HGG sources

The HGG sources aim to describe the different types of metabolic abnormality seen in MRSI data of HGG patients. The retrieved set of sources should be simultaneously (1) *representative* of the various types of spectrum seen in HGG and (2) *reproducible*, ie a similar solution should be found if the same method is applied to another dataset with the same characteristics (acquisition, population etc), provided that both datasets are large enough. Therefore, the choice of the optimal number of HGG sources, K_H, was made as a compromise between (a) the *mean reconstruction error* for a given number k_G of HGG sources, $\overline{R}(k_G)$, and (b) the *stability*[37] of the solutions obtained with cNMF for a given k_G.

To evaluate the reconstruction error and the solution stability as a function of k_G, several executions of the pipeline described by Figure 1 were performed, varying k_G and the input data used to determine the HGG sources. Prior to the experiment, the healthy sources were determined by applying cNMF with K_H sources to the entire healthy control dataset. L_1 normalization was applied to all healthy sources.

A total of 100 iterations were performed, where in each iteration 7 of the 22 patients of the training set were selected randomly for validation. The data of the remaining 15 patients was used as the input dataset to determine the sources for different values of k_G between 1 and 6. For each iteration and k_G value, NNLS was used to determine the fractions of the healthy sources that provide the best reconstruction of the input dataset. Then, the reconstruction error was calculated and the abnormal voxels were identified. Next, a set of HGG sources was retrieved by applying cNMF to these abnormal voxels of the input dataset. The resulting sources were normalized using L_1 normalization. Finally, the abnormal voxels of the validation set were reconstructed using the retrieved HGG sources and the corresponding reconstruction error was calculated. The reconstruction error and the obtained HGG sources were recorded for each iteration and value of K_G.

In order to assess the stability of the determined sources as a function of k_G, the following dissimilarity measure was introduced:

$$\overline{D}(k_G) = \frac{1}{mk_G}\binom{L}{2}^{-1}\sum_{k=1}^{k_G}\sum_{c\in C(L,2)}\sum_{i=1}^{m}\left|s_{i,k}^{c(1)} - s_{i,k}^{c(2)}\right|$$ (2)

where m is the number of features, L is the number of times cNMF was executed for each k_G, $C(L,2)$ corresponds to all possible (unordered) pairs of the L solutions, $c(1)$ and $c(2)$ correspond to the two elements of a given pair of solutions and $s_{i,k}^l$ corresponds to the value of the feature i of the source k from a given solution l. The formula assumes that the sources of each solution are ordered so that sources occupying the same index k in two different solutions of k_G sources represent a similar effect and can be directly compared. Therefore, prior to evaluating the dissimilarity between cNMF solutions of the same k_G, the sources from the corresponding solutions were ordered. The ordering was made considering as reference the first solution for each k_G and following an iterative method. In each iteration, the L_1 distance is measured for every possible pair formed by a source of the reference solution and a source of the solution being ordered. Then, the pair with the shortest L_1 distance is selected and its elements excluded from the next iterations. This is perfomed until all sources are paired with a source of the reference solution.

The mean reconstruction error for the abnormal voxels of the validation sets was calculated as

$$\overline{R}(k_G) = \frac{1}{mNL}\sum_{n=1}^{N}\sum_{l=1}^{L}\sum_{i=1}^{m}\left|x_i^n - \sum_{k=1}^{k_G}a_k^n s_{i,k}^l\right|.$$ (3)

Finally, the optimal number of HGG sources K_G was selected as the value that minimizes the energy function,

$$E(k_G) = \overline{D}(k_G) + \lambda\overline{R}(k_G)$$ (4)

ie

$$K_G = \underset{k_G}{\operatorname{argmin}}(E(k_G)). \tag{5}$$

For the sake of simplicity, and given that $\overline{D}(k_G)$ and $\overline{R}(k_G)$ produce values of the same order, λ was set to 1.

2.8 | HGG sources' maps

After determining the optimal numbers of healthy and HGG sources, the pipeline of Figure 1 was applied to the training set and the final sources were determined. Then, considering all (L_1-normalized) sources at once, NNLS was used to reconstruct the spectra of the test set. Finally, the sources' fractions were determined as the ratios between the factors for each source and the sum of all factors.

2.9 | Image segmentation using BraTumIA

For each MRSI examination, the corresponding MR images were segmented into the following compartments: white matter, gray matter, CSF, edema, non-enhancing tumor, enhancing tumor and necrosis. *MRI-visible tumor* was defined as the union of non-enhancing tumor, enhancing tumor and necrosis.

The enhancing tumor is visible on T_{1c}-weighted MR images as a hyperintense region caused by the contrast enhancement due to the blood–brain barrier disruption. Enhancing tumor is usually surrounded by a large region of edema, which appears hyperintense on T_2-weighted images and hypointense on T_1-weighted images. The FLAIR image is utilized to differentiate edema from CSF, which appears also hyperintense in the T_2 sequence but dark in the FLAIR image due to the suppression of the free water signal. Glioblastoma can contain necrotic parts, which do not enhance on T_{1c}-weighted images but appear hyperintense on T_2-weighted images and hypointense on T_1-weighted images. The most difficult compartment to identify is the non-enhancing tumor, which does not enhance in T_1c-weighted images and generally exhibits a lower intensity in T_2-weighted images than necrosis with the corresponding T_1 hypointensity.

The segmentation of the different compartments was performed using BraTumIA,[26,27] an automatic method that was trained to reproduce the manual image segmentation performed by experienced neuroradiologists. This method uses as input the previously introduced MR image sequences: T_1, T_{1c}, T_2 and FLAIR. BraTumIA is a machine learning-based segmentation method, which was trained on an independent dataset of 54 pre- and postoperative MRI examinations of glioblastoma. It uses supervised learning to generalize from manually segmented ground truth data. A more detailed description of the algorithm and training data can be found in the study by Meier et al.[26]

The process of manually segmenting glioblastoma includes first the segmentation of the complete tumor visible as a hyperintense area in the FLAIR image. Second, the gross tumor volume (including necrosis, non-enhancing and enhancing tumor) is delineated using mainly the co-registered T_{1c} and T_2 sequences. This region exhibits a lower intensity in T_2 than the surrounding edema. The mismatch between the two regions is considered as the edema. The enhancing tumor is segmented based on the hyperintense region in the T_{1c} sequence. Necrosis is differentiated from non-enhancing tumor based mainly on information from the T_2-weighted sequences.

The image segmentation results were converted to the same orientation and resolution as MRSI to allow comparison with the HGG sources' maps. During this conversion, the MRI segmentation class chosen for each spectroscopic voxel was obtained by majority vote, considering the image voxels contained in each spectroscopic voxel.

3 | RESULTS

3.1 | Healthy sources

Figure 2 shows violin plots of the reconstruction error measured by LOOCV and using the healthy control dataset. In the same figure are included the *p*-values associated with the differences in reconstruction error between each number of sources and the max k_H value tested, 15. Considering a significance level of 0.01 and Bonferroni correction, the results show that for $k_H \geq 12$ there is no significant decrease in the mean error in comparison with $k_H = 15$. Therefore, in this paper, K_H was set to 12.

Figure 3 shows the 12 sources obtained using cNMF applied to the healthy control dataset. The different healthy sources provide an overview of the range of metabolic patterns seen in healthy brain tissue. The results show one source containing Lac and lipid signals. By analyzing the MRSI data of the healthy controls it was revealed that the origin of this source is related to lipid contamination from the skull as well as quantification errors.
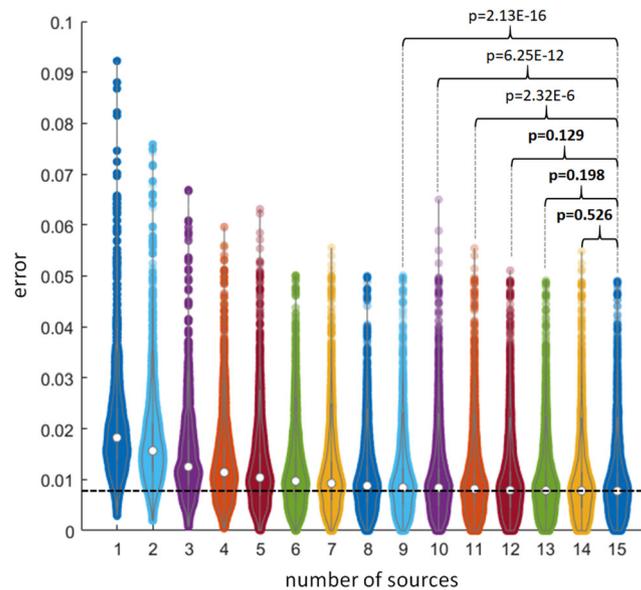
FIGURE 2 Violin plots of the healthy control data reconstruction residue using different numbers of healthy sources. The values were obtained using LOOCV, where each time the data from one patient was used to measure the reconstruction error and the remaining patients were used to determine the healthy sources. For each cross-validation step, the determination of the sources and evaluation of the residue was repeated 10 times. The plot shows the *p*-value (one-sided *t*-test considering unequal variance) associated with the differences between the mean residues of the maximum number of sources tested ($K = 15$) and the other number of sources. For fewer than nine sources the *p*-values were all lower than the ones shown in the figure. Considering a significance level of 0.01 and Bonferroni correction, only for $K < 12$ is the mean residue observed to be significantly higher when compared with $K = 15$
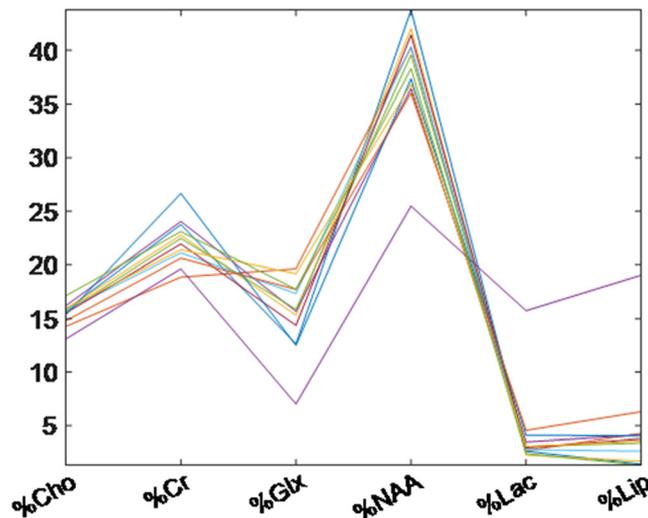
FIGURE 3 Healthy sources retrieved from the healthy control dataset using cNMF with $K = 12$

Figure 4 shows the histogram of the error measured for the reconstruction of the healthy control dataset using 12 healthy sources. The error was measured in the validation sets using LOOCV, as explained in the methods section. In the same figure, the abnormality threshold is shown, defined as the 99th percentile of the reconstruction error.

3.2 | HGG sources

Figure 5 shows a combined plot of (1) the value of the mean dissimilarity between solutions, (2) the mean reconstruction error for a different number of HGG sources (k_G) and (3) the energy function defined as the sum between the previous two quantities. As expected, there is a trend showing a decrease in reconstruction error and a decrease in solution stability as the number of sources increases. Despite this, the plot shows that the
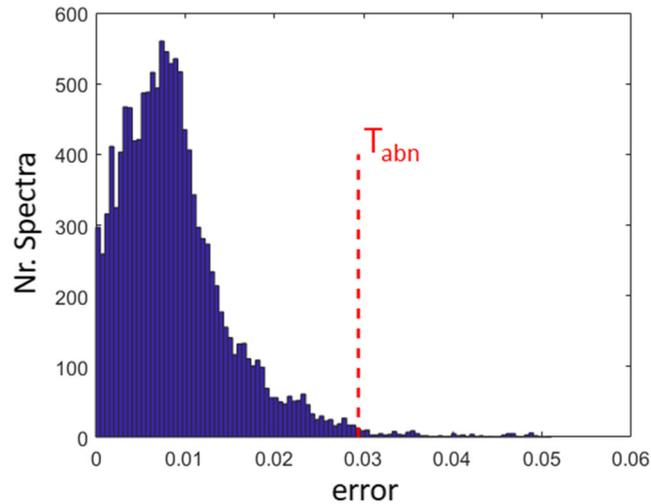
FIGURE 4 Histogram of the healthy control data reconstruction residue obtained using 12 healthy sources. The reconstruction residue was measured as described in the legend of Figure 2 (LOOCV with 10 repetitions for each cross-validation step). The plot shows in red the abnormality threshold, defined as the 99th percentile of the reconstruction residue
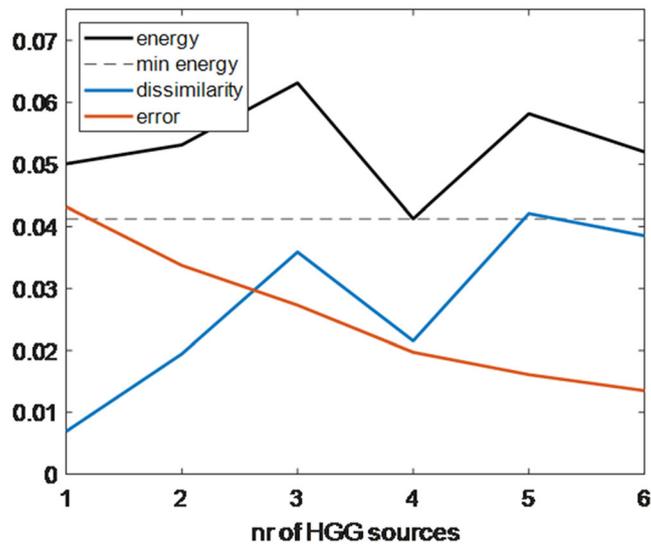
FIGURE 5 Plot of (1) the value of the mean dissimilarity between solutions, (2) the mean reconstruction error for a different number of HGG sources (k_G) and (3) the energy function defined as the sum between the previous two quantities. The different quantities were calculated as described in Section 2 (see expressions 2, 3 and 4). A horizontal line at the level of the minimum value of the energy function obtained was added to assist the analysis of the plot

solutions obtained using four sources are more stable than those obtained with k_G = 3. The curve of the energy function shows that the best compromise between solution stability and reconstruction error is obtained with four sources. Therefore, the chosen number of HGG sources, K_G, was 4.

The differences in stability between the sources obtained with k_G equal to 3 and 4 are clearly visible in Figure 6, where the sources obtained for different values of k_G are shown. In this figure, it can be seen that only for k_G equal to 1, 2 and 4 does cNMF produce stable results on the subset of abnormal spectra. Moreover, the visual comparison of these three stable solutions shows an interesting evolution of the separation of different metabolic alterations seen in HGG as k_G increases. For k_G = 1, the obtained solution corresponds to the average metabolic pattern of the tumor affected region, ie presence of lipids and higher Cho/Cr and Cho/NAA than in normal brain tissue. For two sources, the metabolic alterations associated with tumor necrosis[38,39] are separated from the remaining tumor-related changes. Finally, for k_G = 4 the source associated with necrosis remains almost unchanged, and the effects that were grouped as one source in the solution with two sources are separated into three distinct sources: one showing very high Cho (active tumor[38]), another showing very high levels of Glx (edema[40]) and a third one showing a close-to-healthy pattern but with very low levels of Glx (healthy without Glx).

In order to better understand the source containing low Glx and how it compares with the previously determined healthy sources, we tried to reconstruct it using only healthy sources. As seen in Figure 7, the main difference between this source and its reconstruction using healthy sources
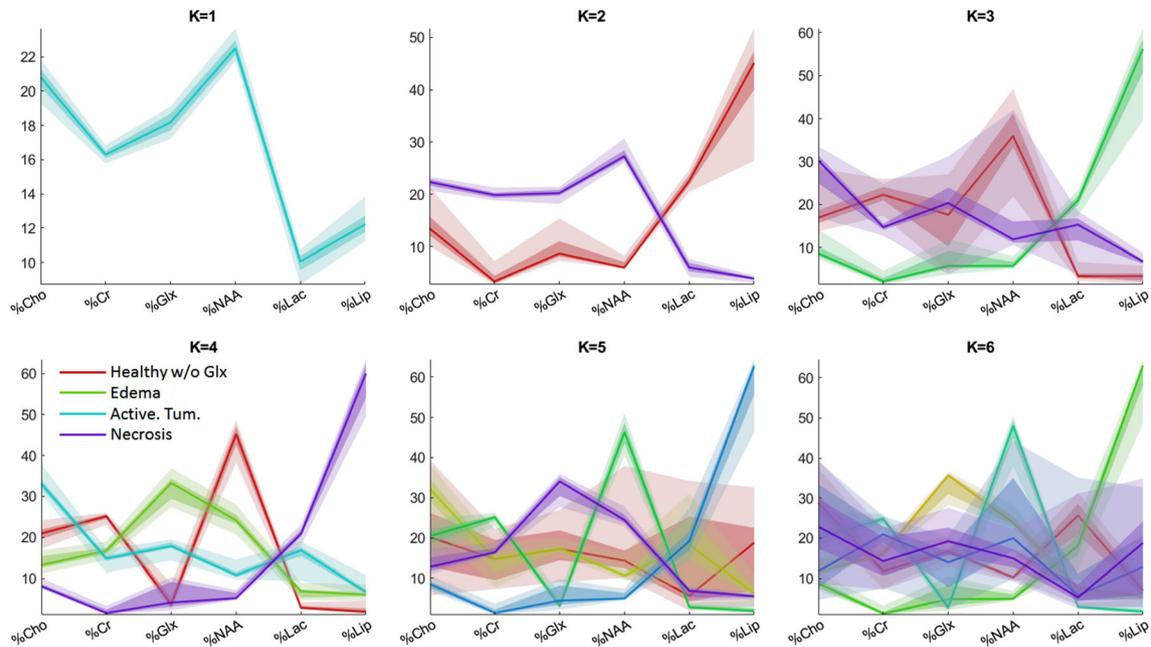
FIGURE 6 Different sources obtained for k_G from 1 to 6. In each plot, the different shades of the same color represent the 5th, 25th, 75th and 95th percentiles of each source. The solid lines represent the median of each source. Before plotting, the sources were ordered as described in Section 2
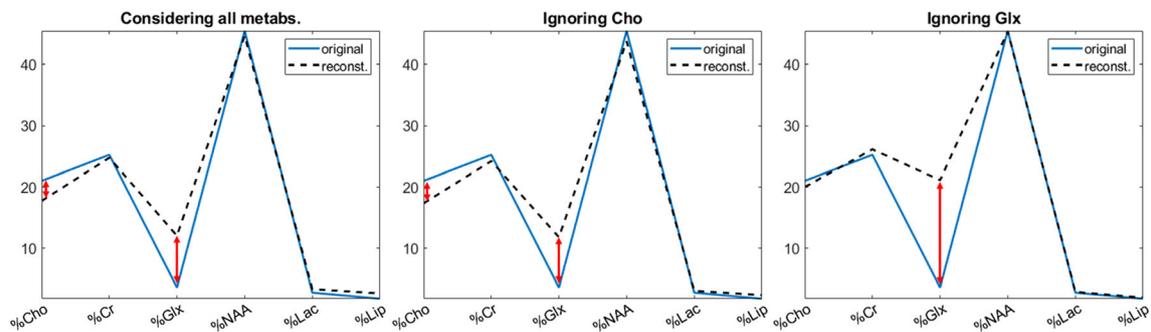
FIGURE 7 Reconstruction of the "healthy without Glx" source using healthy sources and NNLS. To calculate the source factors, different sets of features were provided to NNLS. In the left-hand plot, all sources were used. In the middle one %Cho was ignored, and in the right-hand one %Glx was not considered by the NNLS algorithm used. The differences between the reconstructions and the original "healthy without Glx" source are highlighted using red arrows

are an increase in %Cho and a decrease in %Glx. However, if while determining the healthy source factors with NNLS we ignore each of these two metabolites one at a time, we see that whereas ignoring %Cho does not change the results ignoring %Glx allows us to obtain a low reconstruction error for all the other metabolites. This confirms that the parameter that distinguishes this pattern from the healthy group is %Glx and motivated the name given to this source.

3.3 | Spectra per source

In Figure 8 the spectra associated with each of the obtained sources are shown. The plots were obtained by (1) fitting the test set with all (L_1-normalized) sources using NNLS, (2) calculating the sources' fractions, ie the ratio between the factor of each source and all the factors, and (3) selecting for each group the spectra with fractions higher than 0.75. For the healthy group, the healthy fraction corresponded to the sum of the fractions of all healthy sources. Given that, for each source, there is a small contribution from other sources, and due to the low spatial resolution of the data, the spectra from Figure 8 cannot be regarded as pure spectra from specific tissue types. This explains, for instance, the existence of Cho in the necrosis spectra. Moreover, an abnormally elevated Cho/NAA is seen in the edema source. This may be not only a result of the mentioned factors but also due to neuronal loss in the vicinity of the tumor.
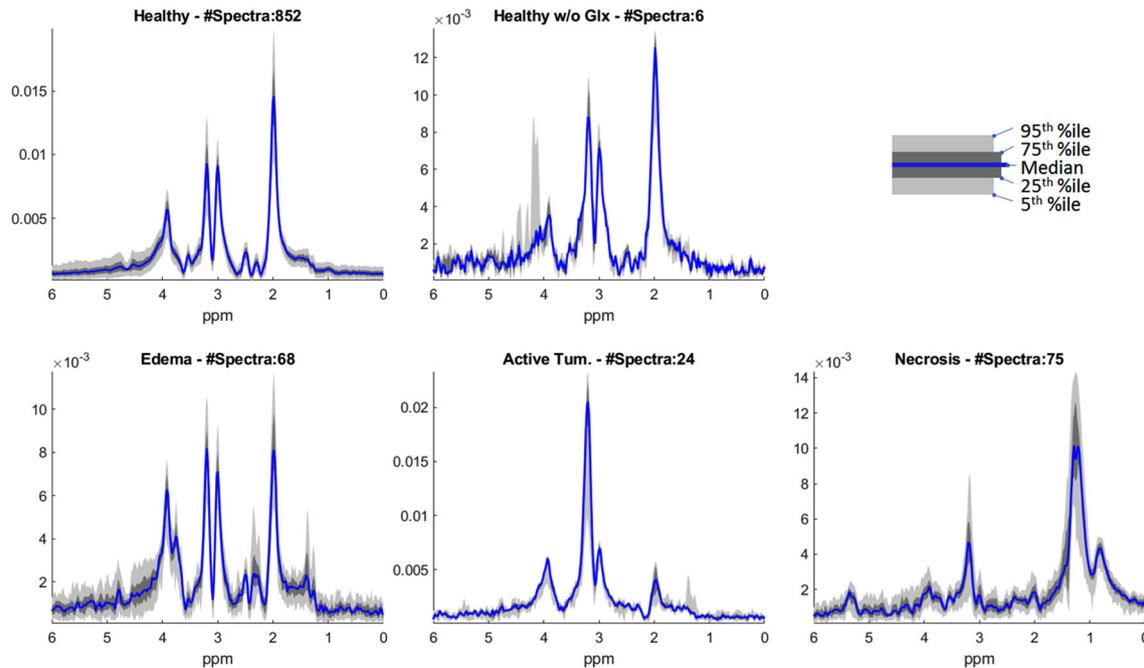
FIGURE 8 Magnitude spectra associated with each source. The spectra for each source corresponds to the selection of voxels that have a fraction of the corresponding source higher than 0.75. For the healthy group, the healthy fraction used to select the spectra corresponds to the sum of the fractions of all healthy sources. The title of each plot shows how many spectra were selected from the test set to represent each source

In the case of the source "healthy without Glx", the low number of spectra showing a high fraction of this source and the low SNR of the selected spectra may imply that this source originates from errors in fitting Glx when the SNR is low.

3.4 | Relationship between sources and MRI segmentation (BraTumIA)

Figure 9 shows the structural MRI (T_{1c} and FLAIR) and corresponding segmentation, the HGG source fraction maps, and the reconstruction error using only healthy sources and using both healthy and HGG sources. The maps of this figure correspond to the data from Patient 23 to Patient 25. The same maps for the remaining patients of the test set are included in the Supplementary Material (Figure S2). The maps of the reconstruction error using only healthy sources show a region of metabolic abnormalities containing the tumor and a margin around it. As expected, the reconstruction errors using all sources are considerably lower than when only healthy sources are used. Nevertheless, a few regions still show relatively large reconstruction errors indicating voxels that are not well represented by the basis set of patterns used.

In order to better understand the relation between the different HGG sources and the different MRI segmentation classes, Figure 10 shows the Pearson correlation coefficient between each source fraction and each segmentation class fraction for the non-interpolated voxels of the test set (statistically insignificant values marked with a cross). The fractions of each MRI class were determined based on the classes of the image voxels contained in each spectroscopic voxel. The results show that the healthy sources' fraction is positively correlated with the healthy MRI segmentation classes and negatively correlated with the tumor-related classes. Furthermore, the fraction of "healthy without Glx" is positively correlated with white matter. The two edema fractions, from MRI and MRSI, are positively correlated ($\rho = 0.44$), supporting the name chosen for the corresponding source. Active tumor is positively correlated with all tumor-related MRI classes, including edema. The correlation between active tumor and edema must arise as a consequence of the ability of MRSI to detect tumor cells beyond the tumor borders visible in MRI.[6,8] Finally, the fraction of the source necrosis correlates positively with all tumor compartment classes: non-enhancing tumor, enhancing tumor and necrosis. From all the pairwise correlations, the highest value measured was obtained for the correlation between the two necrosis fractions ($\rho = 0.58$). The fact that the necrosis source correlates not only with the necrosis in MRI but also with enhancing and non-enhancing tumor may be explained by the under-segmentation of necrosis performed by BraTumIA previously described by Rios Velasquez et al.[41]

3.5 | Validation with follow-up images

Figure 11 shows the data from two patients of the test set. The first row of each patient shows the MRSI-based maps and MRI images of a follow-up where MRSI indicates the presence of tumor in a region where there is no clear contrast enhancement. The remaining rows show the same
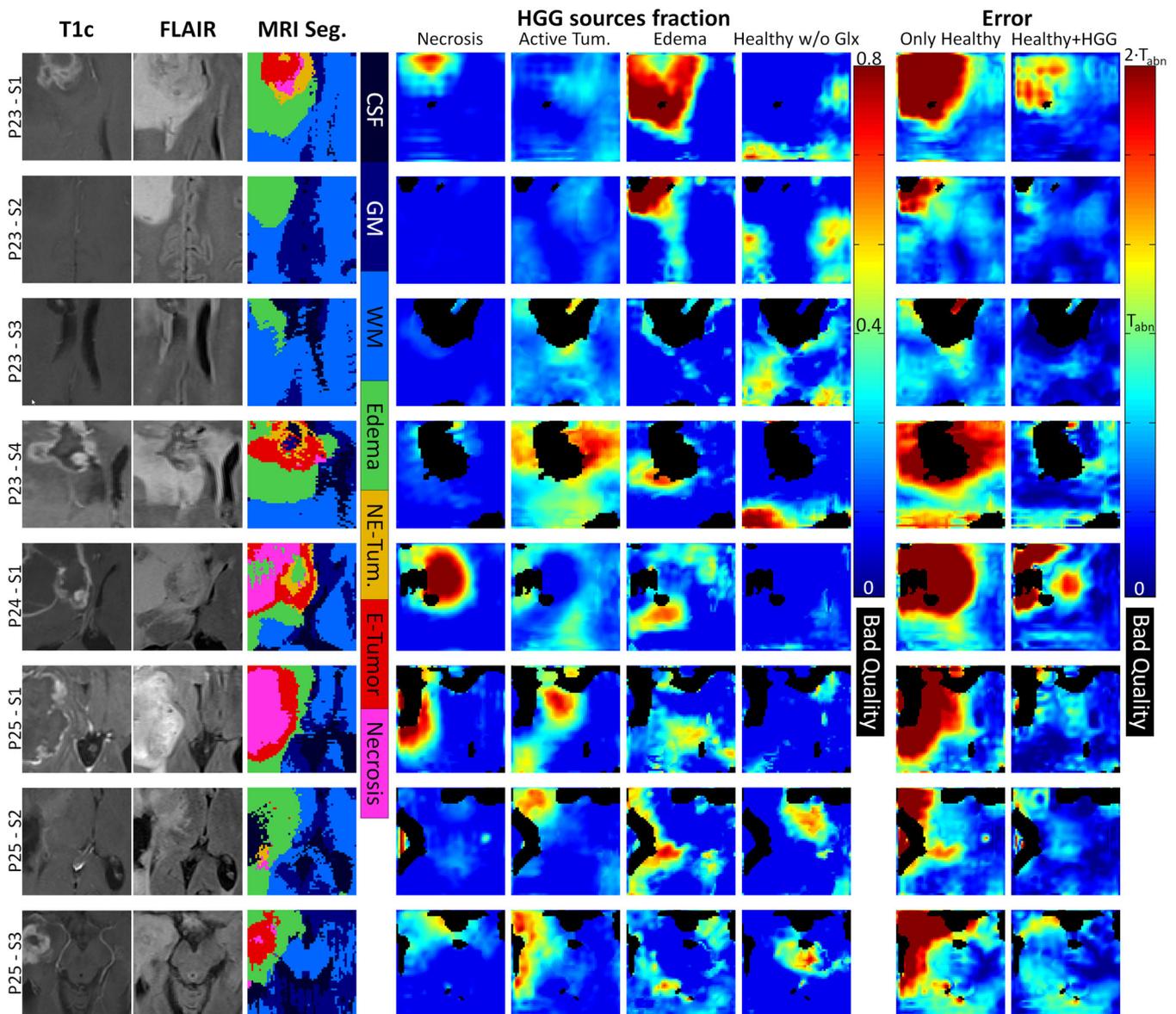
FIGURE 9 Comparison of structural MRI (T_{1c} and FLAIR), and corresponding segmentation, with the HGG source maps. Reconstruction error maps using (1) only healthy sources and (2) both healthy and HGG sources are shown on the right. The maps of this figure correspond to the data from patient 23 to patient 25. The same maps for the remaining patients of the test set are included in the supplementary material (Figure S2)

MRSI-based maps coregistered to T_{1c}-weighted follow-up images acquired several months later. This comparison was made to evaluate the ability of the introduced HGG sources' maps to predict future contrast enhancement. Besides the sources' maps, we also include Cho/NAA and expected distance to tumor (EDT)[42] maps for comparison. As in Reference 6, the Cho/NAA maps were divided by the value of Cho/NAA on the healthy contralateral side. The EDT maps show the predicted distance to the border of the solid tumor volume (necrosis, enhancing and non-enhancing tumor) based solely on the spectroscopic information of each voxel. As shown in the original paper where the method was presented,[42] the method allows recognition of spectroscopic patterns that are normally seen in the peritumoral region of GBM, where infiltration is most likely.

The first time point of Case 1 in Figure 11 corresponds to a post-operative scan acquired after tumor resection. After resection, this patient received concomitant radiotherapy and chemotherapy with TMZ, which could have caused pseudo-progression at Time-Point 2. However, the follow-up was classified as progressive disease following the RANO criteria,[43] and the tumor shows extremely elevated Cho/NAA and Cho/Cr at the second time-point, confirming tumor recurrence (see the last row of Figure 9). In the second case, the initial time point was acquired under chemotherapy with TMZ, which ended a few days before the second time point. During the third time point, the patient was under therapy with bevacizumab. For this second case, all follow-ups were classified as progressive disease and true tumor progression is confirmed by the growth of the contrast-enhanced tumor at subsequent time-points.

In both cases of Figure 11, despite no clear contrast enhancement being visible at $t = 0$, MRSI shows clear signs of the presence of tumor cells near the resection cavity. In the first case, new contrast enhancement is seen 4 months later in the region where (1) Cho/NAA was more than
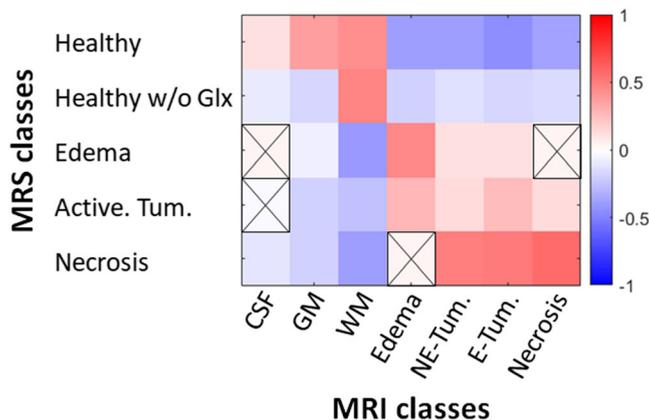
FIGURE 10 Pairwise Pearson correlations between the fraction of each source and the fraction of image voxels of each structural MRI segmentation class contained in the spectroscopic voxel. All the non-significant correlations, considering the Bonferroni correction and a significance level of 0.01, are marked with a cross. All values were calculated using the examinations of the test set

twice the value of the healthy contralateral side, (2) the active tumor fraction was higher than 0.5 and (3) the EDT was shorter than 5 mm. A similar relation is observed in the second case. However, in the frontal region of the grid, where each voxel should contain a considerable amount of CSF, the EDT map shows short predicted distances to the tumor. This false positive may originate from edema-like features, as suggested by the edema fraction map. In the same case, 11 months after the first follow-up, contrast enhancement appears in a region where, at the initial time point, the healthy without Glx fraction map shows a hotspot. Despite the striking relation seen in this particular case, if we consider the different maps of this source shown in Figure 9 it seems unlikely that this source is a good predictor for future contrast enhancement.

4 | DISCUSSION

4.1 | Automatic quality control and selection profile correction are essential pre-processing steps

To obtain the results presented in this paper, two processing steps were essential: automatic quality control and selection profile correction.

Bad quality data, if not removed, prevents meaningful data exploration. The quality of the signals was assessed using SpectrIm's quality control[44] algorithm, which can assess every grid almost instantaneously and with the same accuracy as an MRS expert. This not only saves a considerable amount of time compared with the manual assessment of the data, but also prevents subjective assessments of the quality of the signals and enables fully automatic MRS data processing.

With respect to the selection profile correction, Figure S1 shows the significant impact that this correction can have on one of the most commonly used metabolite ratios: Cho/NAA. MRSI pulse sequences suffering from these undesired effects are broadly used in clinical practice, emphasizing the need for such correction methods in the manufacturers' processing tools for the clinical routine use of MRS. The correction is particularly important when MRSI is used to identify tumor infiltration and for radiotherapy planning using Cho/NAA.[45]

4.2 | Using more than one source to describe healthy brain tissue

Several of the previous NMF-based methods used for describing brain tumor data, such as References [11, 12, 14, 16, 17, 24, and], have used a single source to represent healthy brain tissue. However, as demonstrated in this paper, healthy brain spectra are better approximated using multiple healthy sources. When only one source is used, any deviation from this reference pattern may be regarded as an abnormality. Therefore, it is expected that by considering a range of healthy patterns a more precise identification of metabolic abnormalities may be achieved. The impact of the number of healthy sources for the identification of metabolic abnormalities should be evaluated in further studies.

It should be noted that, in this paper, a very small number of healthy controls was used. To implement the presented method clinically, a larger dataset of healthy reference spectra would be recommended to guarantee that the model captures the normal range of healthy brain spectra seen in the normal population. Such a dataset should contain a wide range of spectra from different brain regions and be acquired from a large variety of healthy controls representative of the normal population. Ideally, a whole brain MRSI sequence should be used.
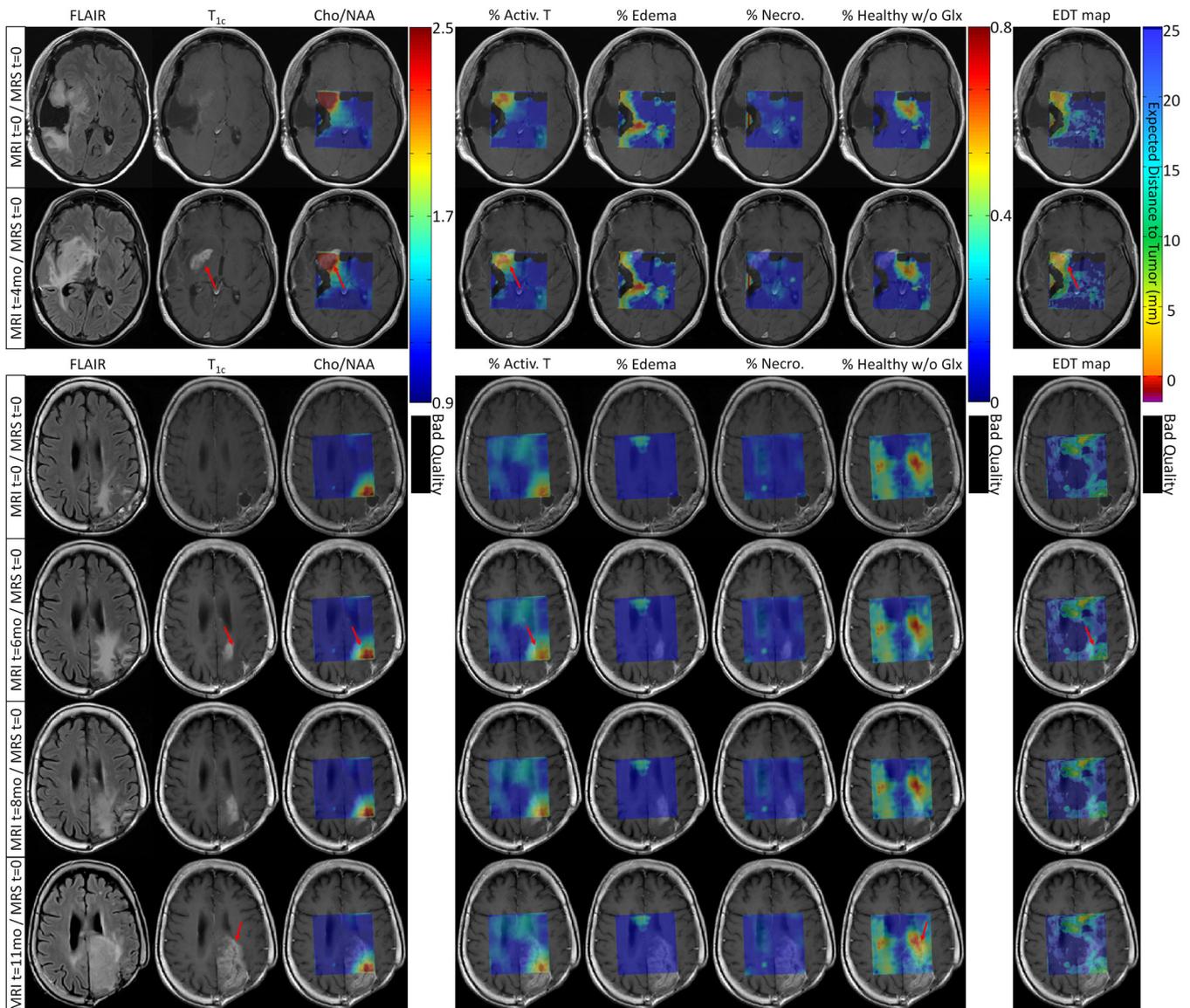
FIGURE 11 MRSI-based maps and MRI images of two patients, for which at a given time point (study 2 of patient 25 and study 3 of patient 26), here denoted $t = 0$, MRSI indicates the presence of tumor in a region where there is no clear contrast enhancement. In the subsequent time points, the rows show the follow-up structural images overlaid with the co-registered initial MRSI data. The image shows the ability of each map to predict future contrast enhancement in two different cases. Cho/NAA maps were divided by the value of Cho/NAA on the healthy contralateral side. EDT maps were determined as described in reference 42. The three follow-up images of the second case are not included in Table 1 since no spectroscopy was acquired in the corresponding studies

4.3 | Choosing the number of pathological patterns

One of the novel aspects of this paper is the approach used to decide the number of pathological sources required to describe the MRSI data of HGG patients. From our perspective, two main factors need to be considered in such a decision: (1) the reproducibility of the results and (2) how well the set of identified sources represents the data.

In order for the results to be reproducible, one needs to guarantee that, for two datasets acquired in the same fashion from two different populations with the same characteristics, cNMF converges to the same solution as the number of included patients increases. This is an important aspect for the clinical use of such methods, allowing results from independent centers to be compared. Therefore, the analysis of the stability of the sources for variations of the dataset is essential. To evaluate the stability of the results we suggest using the dissimilarity metric of Expression 2.

Besides the stability, the number of pathological sources considered dictates the types of metabolic alteration seen by the method. If not enough sources are used, important aspects of the data may be omitted from the set of source maps produced. Therefore, it is necessary to include all the sources that are relevant to the clinical interpretation of the data. Moreover, even if an extensive library of reference sources is

used, it is also important to analyze the reconstruction error maps when performing the clinical interpretation of the results. These may be regarded as "reliability maps" of the sources' maps.

4.4 | Edema source

One of the identified sources is characterized by elevated levels of Glx. This source was named edema given its high correlation with the MRI-segmentation class edema, as well as the previous literature[40] describing high Glx/Cr in brain tumor edema. This pattern was clearly identified in the data analyzed; however, to the best of our knowledge only another NMF-based method[18] considered a source for edema. Nevertheless, in that case, the pattern identified was different for each patient and was based not only on MRSI but also on other MR modalities (T_2, T_{1c}, FLAIR, perfusion and diffusion).

In the edema source, Gln and Glu are grouped as Glx; however, these metabolites have distinct roles in tumor development. Glu is cytotoxic,[46] and its secretion induces an inflammatory response in the surrounding tissue, possibly caused by the neuronal loss resulting from Glu. The inflammatory response is believed to contribute to tumor expansion.[46] Gln, on the other hand, is a tumor nutrient and contributes to several essential metabolic tasks in tumor proliferation: it is involved in bioenergetics, protects the cells against oxidative stress and complements glucose mechanism in the production of macromolecules.[47]

Despite the association between these two metabolites and tumor growth,[46] in the two selected cases of Figure 11 the fraction of the edema source does not seem to predict contrast enhancement.

4.5 | "Healthy without Glx" source

Another of the identified sources was the one named "healthy without Glx". The motivation for this name arises from the results of Figure 7. Despite its clear identification by the method proposed in this paper, it is unclear whether this class corresponds to a true metabolic pattern or simply originates as a consequence of a quantification error related to the difficulty in fitting Glx to low SNR spectra.

4.6 | Detection of tumor infiltration

One of the most interesting applications of MRSI is the detection of tumor infiltration not visible using structural imaging, allowing us to more accurately delineate infiltrative brain tumors such as GBM. The improvement in tumor delineation provided by MRSI has been shown to increase the effectiveness of radiotherapy in GBM patients, leading to increased patient survival.[45] Therefore, it is important to evaluate different approaches for translating the MRSI information into maps that allow for the accurate identification of tumor tissue.

Most *clinical* research on the use of MRSI to identify tumor infiltration[6-8,45,48] has focused on a single parameter: Cho/NAA. However, it would be interesting to further improve the detection of tumor infiltration by incorporating more of the available information. This was the motivation for a recently published method: EDT maps.[42] The method is based on the concept of learning the relationship between metabolic features and the distance to the tumor border visible in structural MRI. After training, the model can be used to identify metabolic patterns normally seen close to the tumor border, where tumor infiltration is most likely.

In Figure 11, we compared the maps of Cho/NAA, the HGG sources and EDT for prediction of future contrast enhancement. From the different maps presented, three maps seem to have predictive value for future contrast enhancement: Cho/NAA, active tumor fraction and EDT. In the cases presented, the fraction of active tumor and the EDT do not seem to provide an advantage over Cho/NAA. However, only two cases were included in this analysis. A more thorough comparison of these three approaches for the detection of tumor infiltration is needed.

5 | CONCLUSION

In this paper, we have presented a novel method for identifying the most relevant metabolic patterns present in MRSI data of HGG patients. The main contributions of this paper consist of (1) the approach introduced for the detection of metabolic abnormalities using reference data collected from healthy controls and (2) the strategy introduced to decide the optimal number of sources in cNMF. The proposed approach was successfully applied to MRSI data of HGG patients and may be applied to MRSI data of different diseases.

In this paper, special emphasis was given to the stability of the spectral sources, which is essential for different centers to be able to reproduce the results presented here. However, to guarantee consistent interpretation of MRSI data across different institutions, the reference libraries of metabolic patterns should be the same. In order to achieve this, it is required to create a framework that allows sharing libraries of reference spectra. The shared libraries should be derived, whenever possible, from large datasets collected from various institutions. For such a project to be implemented, efforts for standardization of MRS methodology such as those described in References [49] and [50] are essential.

By assisting the analysis and interpretation of MRSI data, the presented method may help to solve one of the main limitations for the clinical use of MRSI: the amount of expertise required for clinicians to translate MRS data into relevant clinical information.

ORCID

Nuno Pedrosa de Barros https://orcid.org/0000-0003-2403-6705

Johannes Slotboom https://orcid.org/0000-0001-5121-9852

REFERENCES

1. Zou QG, Xu HB, Liu F, Guo W, Kong XC, Wu Y. In the assessment of supratentorial glioma grade: the combined role of multivoxel proton MR spectroscopy and diffusion tensor imaging. *Clin Radiol.* 2011;66:953-960. https://doi.org/10.1016/j.crad.2011.05.001

2. Zeng Q, Liu H, Zhang K, Li C, Zhou G. Noninvasive evaluation of cerebral glioma grade by using multivoxel 3D proton MR spectroscopy. *Magn Reson Imaging.* 2011;29:25-31. https://doi.org/10.1016/j.mri.2010.07.017

3. Law M, Yang S, Wang H, et al. Glioma grading: sensitivity, specificity, and predictive values of perfusion MR imaging and proton MR spectroscopic imaging compared with conventional MR imaging. *Am J Neuroradiol.* 2003;24:1989-1998.

4. Matsusue E, Fink JR, Rockhill JK, Ogawa T, Maravilla KR. Distinction between glioma progression and post-radiation change by combined physiologic MR imaging. *Neuroradiology.* 2010;52:297-306. https://doi.org/10.1007/s00234-009-0613-9

5. Lichy MP, Henze M, Plathow C, Bachert P, Kauczor HU, Schlemmer HP. Metabolische Bildgebung zur Verlaufskontrolle stereotaktisch bestrahlter Gliome—Wertigkeit der [1]H-MR-Spektroskopie im Vergleich zur FDG-PET und IMT-SPECT. *RöFo.* 2004;176:1114-1121. https://doi.org/10.1055/s-2004-813194

6. Cordova JS, Shu HKG, Liang Z, et al. Whole-brain spectroscopic MRI biomarkers identify infiltrating margins in glioblastoma patients. *Neuro Oncol.* 2016;18:1180-1189. https://doi.org/10.1093/neuonc/now036

7. Guo J, Yao C, Chen H, et al. The relationship between Cho/NAA and glioma metabolism: implementation for margin delineation of cerebral gliomas. *Acta Neurochir.* 2012;154:1361-1370. https://doi.org/10.1007/s00701-012-1418-x

8. Stadlbauer A, Moser E, Gruber S, et al. Improved delineation of brain tumors: an automated method for segmentation based on pathologic changes of ^{1}H-MRSI metabolites in gliomas. *Neuroimage.* 2004;23:454-461. https://doi.org/10.1016/j.neuroimage.2004.06.022

9. Hollingworth W, Medina LS, Lenkinski RE, et al. A systematic literature review of magnetic resonance spectroscopy for the characterization of brain tumors. *Am J Neuroradiol.* 2006;27:1404-1411.

10. Preul MC, Caramanos Z, Leblanc R, Villemure JG, Arnold DL. Using pattern analysis of in vivo proton MRSI data to improve the diagnosis and surgical management of patients with brain tumors. *NMR Biomed.* 1998;11:192-200.

11. Su Y, Thakur SB, Sasan K, et al. Spectrum separation resolves partial-volume effect of MRSI as demonstrated on brain tumor scans. *NMR Biomed.* 2008;21:1030-1042. https://doi.org/10.1002/nbm.1271

12. Du S, Mao X, Sajda P, Shungu DC. Automated tissue segmentation and blind recovery of ^{1}H MRS imaging spectral patterns of normal and diseased human brain. *NMR Biomed.* 2008;21:33-41. https://doi.org/10.1002/nbm.1151

13. Luts J, Laudadio T, Idema AJ, et al. Nosologic imaging of the brain: segmentation and classification using MRI and MRSI. *NMR Biomed.* 2009;22:374-390. https://doi.org/10.1002/nbm.1347

14. Ortega-Martorell S, Lisboa PJG, Vellido A, et al. Convex non-negative matrix factorization for brain tumor delimitation from MRSI data. *PLoS ONE.* 2012;7:e47824. https://doi.org/10.1371/journal.pone.0047824

15. Ortega-Martorell S, Lisboa PJG, Vellido A, Julia-Sape M, Arus C. Non-negative matrix factorisation methods for the spectral decomposition of MRS data from human brain tumours. *BMC Bioinform.* 2012;13:38. https://doi.org/10.1186/1471-2105-13-38

16. Li Y, Sima DM, Van Cauter S, et al. Hierarchical non-negative matrix factorization (hNMF): a tissue pattern differentiation method for glioblastoma multiforme diagnosis using MRSI. *NMR Biomed.* 2013;26:307-319. https://doi.org/10.1002/nbm.2850

17. Raschke F, Fellows GA, Wright AJ, Howe FA. ^{1}H 2D MRSI tissue type analysis of gliomas. *Magn Reson Med.* 2015;73:1381-1389. https://doi.org/10.1002/mrm.25251

18. Sauwen N, Sima DM, Van Cauter S, et al. Hierarchical non-negative matrix factorization to characterize brain tumor heterogeneity using multi-parametric MRI. *NMR Biomed.* 2015;28:1599-1624. https://doi.org/10.1002/nbm.3413

19. Laudadio T, Croitor Sava AR, Sima DM, et al. Hierarchical non-negative matrix factorization applied to three-dimensional 3T MRSI data for automatic tissue characterization of the prostate. *NMR Biomed.* 2016;29:751-758. https://doi.org/10.1002/nbm.3527

20. Li Y, Liu X, Wei F, et al. An advanced MRI and MRSI data fusion scheme for enhancing unsupervised brain tumor differentiation. *Comput Biol Med.* 2017;81:121-129. https://doi.org/10.1016/j.compbiomed.2016.12.017

21. Sauwen N, Acou M, Sima DM, et al. Semi-automated brain tumor segmentation on multi-parametric MRI using regularized non-negative matrix factorization. *BMC Med Imaging.* 2017;17:1-14. https://doi.org/10.1186/s12880-017-0198-4

22. Ding C, Li T, Jordan MI. Convex and semi-nonnegative matrix factorizations. *IEEE Trans Pattern Anal Mach Intell*. 2010;32:45-55. https://doi.org/10.1109/TPAMI.2008.277

23. Mocioiu V, De Barros NMP, Ortega-Martorell S, et al. A machine learning pipeline for supporting differentiation of glioblastomas from single brain metastases. In: *ESANN 2016 Proceedings—European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*; Bruges, Belgium; 2016:247-252.

24. Pedrosa de Barros N, Mocioiu V, Ortega Martorell S, et al. Highlighting differences between GBM and brain metastasis using a blind source separation method applied to MRSI data. Paper presented at: ESMRMB 2015; October, 2015; Edinburgh, UK.

25. Mocioiu V, de Barros NM, Ortega Martorell S, et al. Differentiating glioblastomas from brain metastases by means of convex nonnegative matrix factorization. Paper presented at: ESMRMB; October, 2015; Edinburgh, UK.

26. Meier R, Knecht U, Loosli T, et al. Clinical evaluation of a fully-automatic segmentation method for longitudinal brain tumor volumetry. *Sci Rep*. 2016;6:23376. https://doi.org/10.1038/srep23376

27. Porz N, Bauer S, Pica A, et al. Multi-modal glioblastoma segmentation: man versus machine. *PLoS ONE*. 2014;9:e96873. https://doi.org/10.1371/journal.pone.0096873

28. jMRUI website. www.jmrui.eu

29. den Boogaart A, Van Ormondt D, Pijnappel WWF, De Beer R, Ala-Korpela M. Removal of the water resonance from 1H magnetic resonance spectra. *Math Signal Process*. 1994;3:175-195.

30. Pedrosa de Barros N, Mckinley R, Knecht U, Wiest R, Slotboom J. Automatic quality control in clinical ^{1}H MRSI of brain cancer. *NMR Biomed*. 2016. https://doi.org/10.1002/nbm.3470

31. Pedrosa de Barros N, Knecht U, McKinley R, Giezendanner J, Wiest R, Slotboom J. Automatic quality assessment of short and long-TE brain tumour MRSI data using novel spectral features. Paper presented at: ISMRM 24th Annual Meeting & Exhibition; May 9, 2016; Singapore.

32. Pedrosa de Barros N, Slotboom J. Quality management in *in vivo* proton MRS. *Anal Biochem*. 2017;529:98-116. https://doi.org/10.1016/j.ab.2017.01.017

33. Pedrosa de Barros N. *Improving the Clinical Use of Magnetic Resonance Spectroscopy for the Analysis of Brain Tumours using Machine Learning and Novel Post-Processing Methods*. Bern, Switzerland: University of Bern; 2018.

34. Ratiney H, Sdika M, Coenradie Y, Cavassila S, van Ormondt D, Graveron-Demilly D. Time-domain semi-parametric estimation based on a metabolite basis set. *NMR Biomed*. 2005;18:1-13. https://doi.org/10.1002/nbm.895

35. Lawson CL, Hanson RJ. *Solving Least-Squares Problems*. Upper Saddle River: NJ: Prentice Hall; 1974.

36. Li Y, Ngom A. The non-negative matrix factorization toolbox for biological data mining. *Source Code Biol Med*. 2013. https://doi.org/10.1186/1751-0473-8-10

37. Lange T, Roth V, Braun ML, Buhmann JM. Stability-based validation of clustering solutions. *Neural Comput*. 2004;16:1299-1323. https://doi.org/10.1162/089976604773717621

38. Vigneron D, Bollen A, McDermott M, et al. Three-dimensional magnetic resonance spectroscopic imaging of histologically confirmed brain tumors. *Magn Reson Imaging*. 2001;19:89-101.

39. Horska A, Barker PB. Imaging of brain tumors: MR spectroscopy and metabolic imaging. *Neuroimaging Clin N am*. 2010;20:293-310. https://doi.org/10.1016/j.nic.2010.04.003

40. Ricci R, Bacci A, Tugnoli V, et al. Metabolic findings on 3T ^{1}H-MR spectroscopy in peritumoral brain edema. *Am J Neuroradiol*. 2007;28:1287-1291. https://doi.org/10.3174/ajnr.A0564

41. Rios Velazquez E, Meier R, Dunn WD Jr, et al. Fully automatic GBM segmentation in the TCGA-GBM dataset: prognosis and correlation with VASARI features. *Sci Rep*. 2015;5:16822.

42. Pedrosa de Barros N, Meier R, Pletscher M, et al. On the relation between MR spectroscopy features and the distance to MRI-visible solid tumor in GBM patients. *Magn Reson Med*. 2018. https://doi.org/10.1002/mrm.27359

43. Wen PY, Macdonald DR, Reardon DA, et al. Updated response assessment criteria for high-grade gliomas: response assessment in neuro-oncology working group. *J Clin Oncol*. 2010;28:1963-1972. https://doi.org/10.1200/JCO.2009.26.3541

44. Pedrosa de Barros N, Mckinley R, Knecht U, Wiest R, Slotboom J. Automatic quality control in clinical ^{1}H MRSI of brain cancer. *NMR Biomed*. 2016;29:563-575. https://doi.org/10.1002/nbm.3470

45. Einstein DB, Wessels B, Bangert B, et al. Phase II trial of radiosurgery to magnetic resonance spectroscopy-defined high-risk tumor volumes in patients with glioblastoma multiforme. *Int J Radiat Oncol Biol Phys*. 2012;84:668-674. https://doi.org/10.1016/j.ijrobp.2012.01.020

46. Takano T, Lin JH, Arcuino G, Gao Q, Yang J, Nedergaard M. Glutamate release promotes growth of malignant gliomas. *Nat Med* 2001;7:1010–1015. https://doi.org/10.1038/nm0901-1010.

47. DeBerardinis RJ, Cheng T. Q's next: the diverse functions of glutamine in metabolism, cell biology and cancer. *Oncogene*. 2010;29:313-324. https://doi.org/10.1038/onc.2009.358

48. Ganslandt O, Stadlbauer A, Fahlbusch R, et al. Proton magnetic resonance spectroscopic imaging integrated into image-guided surgery: correlation to standard magnetic resonance imaging and tumor cell density. *Oper Neurosurg*. 2005;56. https://doi.org/10.1227/01.NEU.0000156782.14538.78

49. Oz G, Alger JR, Barker PB, et al. Clinical proton MR spectroscopy in central nervous system disorders. *Radiology*. 2014;270:658-679. https://doi.org/10.1148/radiol.13130531

50. Wilson M, Andronesi O, Barker PB, et al. A methodological consensus on clinical proton MR spectroscopy of the brain: review and recommendations. 2019. https://doi.org/10.1002/mrm.27742

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.