



LJMU Research Online

Ortega Martorell, S, Candiota, AP, Thomson, R, Riley, P, Julia-Sape, M and Olier-Caparroso, I

Embedding MRI information into MRSI data source extraction improves brain tumour delineation in animal models

<http://researchonline.ljmu.ac.uk/id/eprint/11142/>

Article

Citation (please note it is advisable to refer to the publisher's version if you intend to cite from this work)

Ortega Martorell, S, Candiota, AP, Thomson, R, Riley, P, Julia-Sape, M and Olier-Caparroso, I Embedding MRI information into MRSI data source extraction improves brain tumour delineation in animal models. PLoS One. ISSN 1932-6203 (Accepted)

LJMU has developed **LJMU Research Online** for users to access the research output of the University more effectively. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in LJMU Research Online to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain.

The version presented here may differ from the published version or from the version of the record. Please see the repository URL above for details on accessing the published version and note that access may require a subscription.

For more information please contact researchonline@ljmu.ac.uk

<http://researchonline.ljmu.ac.uk/>

1 Embedding MRI information into MRSI data source extraction
2 improves brain tumour delineation in animal models

3

4 **Authors and affiliations:**

5 Sandra Ortega-Martorell^{1,2,*}; Ana Paula Candiota^{2,3,4}; Ryan Thomson¹; Patrick Riley¹; Margarida Julia-
6 Sape^{2,3,4}; Ivan Olier¹.

7

8 ¹ Department of Applied Mathematics, Liverpool John Moores University, Liverpool, UK

9 ² Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina (CIBER-
10 BBN), Cerdanyola del Vallès, Spain

11 ³ Departament de Bioquímica i Biologia Molecular, Unitat de Biociències, Universitat Autònoma de
12 Barcelona, Cerdanyola del Vallès, Spain

13 ⁴ Institut de Biotecnologia i de Biomedicina, Universitat Autònoma de Barcelona, Cerdanyola del
14 Vallès, Spain

15

16 * Corresponding author

17 Email: s.ortegamartorell@ljmu.ac.uk

18

19 **Keywords:**

20 Magnetic Resonance (MR); MR spectroscopy (MRS); MR imaging (MRI); MR spectroscopic imaging
21 (MRSI); Glioblastoma; Semi-supervised analysis; Fisher Information metric; Multi-Layer Perceptron;
22 Non-negative Matrix factorisation (NMF); Convex-NMF; Semi-Supervised Source Extraction (SSSE)

23 **Abstract**

24 Glioblastoma is the most frequent malignant intra-cranial tumour. Magnetic resonance imaging is the
25 modality of choice in diagnosis, aggressiveness assessment, and follow-up. However, there are
26 examples where it lacks diagnostic accuracy. Magnetic resonance spectroscopy enables the
27 identification of molecules present in the tissue, providing a precise metabolomic signature. Previous
28 research shows that combining imaging and spectroscopy information results in more accurate
29 outcomes and superior diagnostic value. This study proposes a method to combine them, which builds
30 upon a previous methodology whose main objective is to guide the extraction of sources. To this aim,
31 prior knowledge about class-specific information is integrated into the methodology by setting the
32 metric of a latent variable space where Non-negative Matrix Factorisation is performed. The former
33 methodology, which only used spectroscopy and involved combining spectra from different subjects,
34 was adapted to use selected areas of interest that arise from segmenting the T2-weighted image. Results
35 showed that embedding imaging information into the source extraction (the proposed semi-supervised
36 analysis) improved the quality of the tumour delineation, as compared to those obtained without this
37 information (unsupervised analysis). Both approaches were applied to pre-clinical data, involving
38 thirteen brain tumour-bearing mice, and tested against histopathological data. On results of twenty-eight
39 images, the proposed Semi-Supervised Source Extraction (SSSE) method greatly outperformed the
40 unsupervised one, as well as an alternative semi-supervised approach from the literature, with
41 differences being statistically significant. SSSE has proven successful in the delineation of the tumour,
42 while bringing benefits such as 1) not constricting the metabolomic-based prediction to the image-
43 segmented area, 2) ability to deal with signal-to-noise issues, 3) opportunity to answer specific questions
44 by allowing researchers/radiologists define areas of interest that guide the source extraction, 4) creation
45 of an intra-subject model and avoiding contamination from inter-subject overlaps, and 5) extraction of
46 meaningful, good-quality sources that adds interpretability, conferring validation and better
47 understanding of each case.

48 **Introduction**

49 Magnetic Resonance (MR) is widely used for non-invasive investigations of brain tumours, in particular
50 *in vivo* diagnosis and grading, surgical planning and assessment of response to therapy. It is generally
51 applied as MR imaging (MRI), see Fig 1(a), which provides a morphologic characterisation of tissues,
52 and it is the modality of choice in diagnosis, aggressiveness assessment and follow-up. However, there
53 are examples (i.e. malignant gliomas) where current imaging techniques lack diagnostic accuracy [1].
54 In contrast, MR spectroscopy (MRS) provides biochemical information, resulting in a precise
55 metabolomic signature of the target tissue, which enables the identification of a wide array of molecules
56 present in tissues. MR spectroscopic imaging (MRSI), see Fig 1(b and c), produces a spatial distribution
57 of these metabolomic profiles, and thus delivers information about the spatial localisation of molecules
58 [2,3]. Typically, an MRSI acquisition of the brain consists of a spectral grid of varying dimensions (e.g.
59 10-by-10, 12-by-12) superimposed on the image, covering only a part of the full image.

60

61 **Fig 1. Two Magnetic Resonance approaches, MRI and MRSI, acquired in a murine glioblastoma model.** a)
62 Region of interest of the MRI. b) 10-by-10 MRSI grid of voxels, showing the metabolic composition of the tissue
63 at voxel level. c) MRSI grid superimposed on the MRI image. d) MRS example of the non-tumour area of this
64 mouse, indicating the position of a number of relevant metabolites: choline (Cho), creatine (Cr) and N-acetyl
65 aspartate (NAA). e) MRS example of the non-tumour area of this mouse, including lactate/mobile lipids
66 (Lac/ML), and highlighting the changes that occur to these relevant metabolites.

67 Glioblastoma (GB) is the most frequent malignant intra-cranial tumour, having a very poor prognosis.
68 Currently there is no cure for the disease. The standard treatment is surgery, with the aim of maximal
69 safe tumour removal, followed by radio- and chemotherapy [4]. Despite this, due to the infiltrative
70 nature of GB it is not always possible to assess the true degree of infiltration for a total tumour resection
71 using conventional MRI [5]. Additionally, chemo- and radiotherapy only provide a small increase in
72 the overall survival [6]. Eventually, most patients end up relapsing, and when this happens, they may
73 be offered the possibility of starting with either second-line or experimental treatment approaches. One

74 of the problems at this stage is the early assessment of progression using non-invasive methods such as
75 MRI, which may be inaccurate for some patients (e.g. distinguishing true progression from pseudo-
76 progression [7]). In this sense, non-invasive tools able to accurately distinguish the compromised area,
77 which in turn would allow early prediction of relapse, would be of maximal benefit.

78 Pattern recognition and machine learning methods have been extensively applied to MR data of brain
79 tumours, to assist with different clinical issues, from diagnosis and prognosis of several pathologies, to
80 delineation of tumour masses [8–11]. Most studies have been performed on single-voxel MRS [12–16],
81 where a single spectrum is acquired from the pathological area. Even though substantial advances have
82 been achieved, there are still challenges from the methodological point of view that need to be
83 addressed. One of them is the ability to combine, in a principled manner, information from two different
84 MR approaches such as images and spectra. As can be seen in Fig 1, the resolution of the MRSI (i.e.
85 number of voxels of each acquisition) is considerably lower than the resolution of the MRI, which is at
86 the pixel level. This combination of MR images and spectra poses an important challenge, as it is
87 necessary to deal with different types of signals that also have different resolutions [17,18].

88 In biomedical research, animal (preclinical) models are useful for testing new drugs or treatments. One
89 of the main advantages is the possibility of performing post-mortem studies – for example in the case
90 of brain studies, to excise the whole brain and perform a detailed histopathological analysis. This allows
91 the validation of imaging techniques and mathematical analyses or transformations of those, in contrast
92 to patient studies, in which such validation approaches are not feasible due to obvious ethical
93 restrictions.

94 The clinical value of a multivariate statistical analysis based on multiparametric MRI and MRSI for the
95 non-invasive analysis of brain tumours has been previously assessed in a number of publications, such
96 as [9,10,19]. They show that results combining MR approaches are more accurate and have superior
97 diagnostic value compared with single approach. Most of those works used supervised mathematical
98 models that explicitly require class information, i.e. tumour type and grade. This is often achieved in an
99 *ad-hoc* manner by concatenating or mixing selected characteristics from each approach (that can be MR
100 images and spectra), such as in [3,20,21]. A more recent study [22] makes use of Structured Data Fusion

101 (SDF) [23], which could be potentially used for a more appropriate coupling of the information by
102 joining the factors that are obtained in the factorisation process. However, authors in [22] do not seem
103 to have used SDF in that way, but instead chose to concatenate the information as in the aforementioned
104 studies. Providing multiparametric MRI and MRSI approaches in a principled manner may help to
105 overcome the instability in the tissue segmentation that may arise from intrinsic mixing in data space.
106 Therefore, the purpose of our study was to develop a new methodology for embedding morphological
107 information from MR images into the MRSI analysis of brain tumours in animal models, using a Semi-
108 Supervised approach to Source Extraction (SSSE). We applied this approach to retrospective MR data
109 from an orthotopic murine GB model (GL261 GB growing into C57/BL6 mice), widely used in
110 preclinical research, which mimics most of the human GB features [24]. In order to validate the
111 MRI+MRSI fusion, preclinical data allowed us to assess the quality of the tumour segmentation in
112 comparison to a third, independent technique (histopathology). We applied our methodology to both
113 control and treated tumour-bearing mice.

114 **Materials and methods**

115 **Ethics Statement**

116 No ethics approval was required for the current retrospective study. All studies with mice were approved
117 previously by the local ethics committee [*Comissió d'Ètica en l'Experimentació Animal i Humana*
118 (CEEAH). Available: <http://www.uab.cat/etica-recerca/>. Last accessed 29/06/2018], according to the
119 regional and state legislation (protocol DARP-3255/CEEAH-530). Mice were periodically subjected to
120 welfare inspections to check for any early symptoms of suffering and an objective scale for signs and
121 symptoms was established. Mice were obtained from Charles River Laboratories (France) and housed
122 at the animal facility of the *Universitat Autònoma de Barcelona (Servei d'Estabulari)*.

123 **MR studies**

124 MR studies were carried out at the joint nuclear MR facility of UAB and CIBER-BBN, Unit 25 of
125 NANBIOSIS (www.nanbiosis.es), with a 7 Tesla horizontal magnet (BioSpec 70/30, Bruker BioSpin,
126 Ettlingen, Germany). Details of the acquisition parameters can be found in the Supporting Information
127 file.

128 MRSI data were processed as described in [25,26]. The MRSI data grid was formed by an array of
129 10×10 voxels (MR spectrum from each voxel contained 692 data points), with an in-plane resolution of
130 0.55×0.55 mm and a 1 mm slice thickness in the 3rd dimension [25]. This volume of interest was
131 manually positioned approximately in the centre of the brain, based on the reference image, in a way
132 that it would include most of the tumour mass and also part of the normal/peritumoural brain
133 parenchyma.

134 The study was performed on retrospective data already acquired. Only the spectral information (MRSI)
135 from these data had been used in previous pattern recognition studies [25–28], performed for other
136 purposes. Reference T2w MRI for all mice had not been used previously in any of the analyses, except
137 for providing the anatomic overlay for the MRSI data. This is the first time that the images (MRI) were
138 also part of the pattern recognition analysis for these mice.

139 **Datasets**

140 **Pre-treatment data for analysis of tumour vs. non-tumour (groups A and B)**

141 In this section, we describe the control (untreated) mice data that was used in this study to evaluate the
142 ability of SSSE to delineate the tumour mass. This data is summarised in Table 1 and includes:

- 143 A. MRI and MRSI (short TE, 12ms) data of six mice from [25]. The MRSI data in the mentioned
144 study had been used for the purpose of discriminating the tumour from the healthy tissue, in a
145 fully unsupervised way.

146 B. MRI and MRSI (short TE, 14 ms) data of five mice from [26,28]. The MRSI data in the
 147 mentioned studies had been used as control (untreated) group to assess response to therapy, in
 148 a semi-supervised way.

149 **Follow-up, longitudinal data for analysis of tumour vs. non-tumour (group**
 150 **C)**

151 In this section, we describe the follow-up, longitudinal data that was used in this study. This is data
 152 obtained from mice under temozolomide (TMZ) treatment with the administration schedule described
 153 in [28] and in the Supporting Information file.

154 The aim was to assess SSSE’s ability to produce an accurate delineation of the tumour mass, as well as
 155 recognise the volume changes at different time points, and tumour response to therapy. This data is also
 156 summarised in Table 1 and includes:

157 C. MRI and MRSI (short TE, 14 ms) data of two treated mice from [26]. Again, only the MRSI
 158 data had been used previously. For this study, we selected the two mice receiving treatment for
 159 the longest period from [26] (survival time of 45 and 34 days, and with 9 and 7 MR
 160 explorations, respectively).

161

162 **Table 1. Data used for analysis in this study.**

Group	Analysis in previous studies	References	Unique ID of mice (D: day post-inoculation)	Number of mice
A	Performed on untreated cases	[25,27]	C69 (D: 15), C71 (D: 16), C32 (D: 16), C179 (D: 17), C233 (D: 17), C234 (D: 17)	6
B	Performed on untreated cases	[26,28,29]	C255 (D: 14), C288 (D: 18), C520 (D: 18), C529 (D: 18), C583 (D: 18)	5
C	Performed on treated cases, longitudinal study	[26]	C819 (D: 10, 15, 18, 21, 25, 30, 33, 41, 45), C821 (D: 10, 15, 18, 21, 25, 30, 33)	2

163 This table includes the references to original studies with this data, unique ID of animals, and their number of
 164 individuals.

165 **Histopathology**

166 After the MRSI study (or at after MRSI endpoint in the Group C cases), the animals were euthanised
167 according to the ethics protocol by experienced personnel and the brains were collected and analysed
168 by histopathology as described in [28]. Caspase 3 immunohistochemical staining was used for detecting
169 apoptosis. Ki67 immunohistochemical staining was used to determine the spatial proliferating
170 population of cells in each tumour mass [25,30], calculated as a proliferation index (PI). In this
171 particular murine model, and according to the veterinary pathologist, $PI > 30\%$ would correspond to a
172 safe threshold for identifying the solid tumour region, whereas a $PI \leq 5\%$ would correspond to definitely
173 non-tumour (excluding reactive gliosis and other phenomena). This is in agreement with other studies
174 with murine glioma, in which a PI of 23.9% was found in tumour core, and 9.6% in tumour periphery
175 [31].

176 The evaluation of necrosis was performed by the histopathologist on the haematoxylin and eosin stained
177 slides. Different features were considered: isolated necrotic cells, moderate amounts of eosinophilic
178 debris and large empty spaces. For each tumour, the percentage of the tissue section affected was semi-
179 quantified. When a tumour showed less than 20% of the mentioned necrotic features, a low grade of
180 necrosis was assigned, while high grade of necrosis was assigned to those having more than 20% of the
181 features. More details are included in [26].

182 Using all the available histopathological information such as Ki67 (which is information obtained *a*
183 *posteriori*, or *ex-vivo*), the preclinical bioimaging expert produced a set of images with the aim of using
184 them as the gold standard for this study. These images are compiled in Fig 2.

185

186 **Fig 2. Delineation of the tumour area used as gold standard in this study.** Images were produced *a posteriori*
187 by the preclinical bioimaging expert for Groups A (first row), B (second row), and C (third and fourth rows). Two
188 versions, (a) and (b), are available for C179, to study the main tumour mass first, and the two masses later,
189 considering they have different proliferation indices, i.e. the rostral mass (“secondary mass”) had a $5\% < PI \leq 30$

190 while the caudal mass (“core mass”) had a PI > 30%. Additionally, for C819 at day 41, a second (blue) area was
191 highlighted as abnormal.

192 **Non-negative Matrix Factorisation for source extraction**

193 In Non-negative Matrix Factorisation (NMF) methods [32,33] the non-negative data matrix X is
194 approximately factorised into two non-negative matrices: the matrix of sources or data basis S and the
195 mixing matrix H . The product of these two matrices provides an approximation to the original data
196 matrix in the form $X \approx SH$. There are different NMF variants, which mainly arise from using different
197 cost functions for computing the divergence between X and SH . While NMF describes the observed
198 data with positive-only mixtures of the latent variables or data sources, this does not apply to long echo
199 times where spectral phase-related signal modulation frequently results in negative values in the lactate
200 and alanine regions [12].

201 Convex Non-negative Matrix Factorisation (Convex-NMF) [34] is a variant of NMF that imposes a
202 restriction over the source matrix S to be a convex combination of the input data vectors. This restriction
203 significantly improves the quality of data representation of S . Unlike standard NMF, Convex-NMF
204 applies to both nonnegative and mixed-signed data matrices. What this means in practice is that: a) the
205 data are described by positive-only mixtures of the mixed-signed sources; and b) the sources, or latent
206 variables, are also positive-only mixtures of the data. This makes it easier to interpret both the mixing
207 and unmixing processes.

208 **Semi-supervised methodology for source extraction in MRSI data**

209 The semi-supervised methodology proposed in [13] involves three main stages and, in a nutshell, can
210 be described as follows:

- 211 a) Definition of a Fisher Information (FI) metric to model pairwise similarities and dissimilarities
212 between data points, using a Multi-Layer Perceptron (MLP) classifier to estimate the
213 conditional probabilities of class membership.

- 214 b) Approximation of the empirical data distribution in a Euclidean projective space in which we
215 can apply NMF-based techniques. Multi-dimensional Scaling (MDS) is one of the algorithms
216 proposed to do this mapping while retaining the distance structure generated by the FI matrix.
- 217 c) Application of Convex-NMF for the source decomposition of the data, which includes the
218 identification of the underlying sources and the calculation of the corresponding mixing matrix.

219 This semi-supervised methodology was previously applied to *single-voxel* data in [13], and *multi-voxel*
220 data in [26], although in a slightly different way. Both studies involved using the labels provided by
221 experts and combining spectra from different subjects to create a training dataset in which the sources
222 were extracted.

223 **Proposed methodology to embed MRI information into the source** 224 **extraction**

225 The proposed methodology, SSSE, builds upon the semi-supervised method proposed in [13] for the
226 extraction of relevant sources, which guides the source extraction in the direction of provided class
227 labels. In this new approach (see Fig 3), instead, we use the areas that arise from segmenting the MRI
228 (i.e. T2w images), hence using the normal parenchyma/peritumoral/ventricle/tumour structures
229 identified by MRI. We recommend this initial segmentation of the image to be performed manually by
230 a researcher (e.g. radiologist, clinician, data analyst, etc.), but other automatic approaches can also be
231 considered.

232

233 **Fig 3. Diagram of SSSE.** Details of the methodology proposed in this study for an extraction of sources in a semi-
234 supervised way, informed by knowledge hauled out from the T2w MRI.

235 Importantly, the semi-supervised nature of the proposed methodology has the benefit that it allows using
236 only partial regions, so that areas of uncertainty can be left outside the initial segmentation, while
237 allowing for concentration on the areas of maximum interest.

238 Following the initial selection of the areas of interest (or segmentation), we compute the posterior
239 probability of each class (or segmented region) for each pixel. As mentioned previously, our choice of
240 estimator of these probabilities, $p(c|x)$, is a Multi-Layer Perceptron (MLP), which is a feed-forward
241 artificial neural network. This is a semi-parametric non-linear probabilistic model of class membership,
242 for which a FI metric can be derived [35]. The parameters of the MLP were set as follows: We used
243 one hidden layer of 6 nodes plus the output layer. The initial weights and bias were generated randomly
244 in the interval $[-1, 1]$, the learning rate was set to 0.05, the momentum constant to 0.9, and weight decay
245 regularization was used to limit the size of the weights (lambda constant set to 0.01), which should
246 encourage the creation of simpler models with better generalisation capabilities. We sought model
247 convergence by allowing the MLP to iterate over a number of iterations (maximum epochs of 2000),
248 until the error was smaller than $5e-3$. The models were trained using 75% of the instances for training
249 and the rest for test.

250 After calculating the posterior probability of each class for each pixel, as the resolution of the T2w
251 image is higher than the one of the MRSI, we implement a voting system to bring the resolution of the
252 T2w image down to the spectroscopic imaging. Then, we train the neural network model using the
253 spectra from the MRSI and the labels provided by the segmentation, targeting the computed posterior
254 probability of each class. From this point onwards, the rest of SSSE is similar to the semi-supervised
255 methodology proposed in [13], but applied intra-subject (i.e. to an individual mouse) as opposed to a
256 number of them (as in our previous study [13]). Final resulting maps with the produced segmentations
257 were linearly interpolated to bring them to the resolution of the image, as in [25].

258 It is important to stress that SSSE was designed to create a model of an individual case, as the purpose
259 was to better understand and reflect the characteristics of each single patient. Hence, in order to avoid
260 contamination from inter-subject overlaps, SSSE does not involve combining spectra from different
261 subjects, thus focusing on intra-subject variation. What this means in practice is that every single case,
262 C_i , was studied independently from the others (even excluding information from the same case acquired
263 on a different day). From this data of case C_i , a proportion (detailed previously) was used for the initial

264 training step as required by the methodology, and from there the rest of the proposed pipeline is
265 unsupervised.

266 **Evaluation of the results**

267 To evaluate whether embedding information from the MRI into the data source extraction improves the
268 quality of the tumour delineation, we compare SSSE segmentations with the ones obtained without
269 prior knowledge (Convex-NMF). We also compared SSSE with the approach proposed by Sauwen et
270 al. in [22], as they both bear some similarities in their aim to propose a semi-supervised / semi-
271 automated method for the segmentation of brain tumours. For the latter comparison, some
272 considerations had to be made to allow for a fair comparison of both methods (please see details in the
273 Supporting Information file). The three approaches will be tested against the gold standard (see Fig 2,
274 and the Histopathology section for more details).

275 As measures used for comparison, we start by calculating the sensitivity and specificity of detecting the
276 tumour regions. The sensitivity, or true positive rate, is calculated as $TP/(TP+FN)$, where true positive
277 (TP) are tumour pixels correctly labelled as tumours; and false negative (FN) are the tumour pixels
278 labelled as non-tumours. The specificity, or true negative rate, is calculated as $TN/(TN+FP)$, where true
279 negative (TN) are the non-tumour pixels correctly identified as non-tumours, and false positive (FP) are
280 the non-tumour pixels labelled as tumours. A related measure frequently used in this area [36] for
281 comparing the similarity of two samples is the Dice score coefficient (DSC), also known as the
282 Sørensen–Dice coefficient [37], in order to show the effectiveness and robustness of proposed approach.
283 This is calculated from the values of TP, FP and FN, as follows:

$$DSC = \frac{2TP}{2TP + FP + FN} \quad (1)$$

284 As an overlap-based metric such as DSC can be dependent on the segmentation size, we also calculate
285 two distance-based metrics: the Euclidean distance, and the Hausdorff distance. These distances were
286 calculated between each of the resulting images and the corresponding one from the preclinical
287 bioimaging expert, considered the gold standard in this study. The purpose was to determine how far

288 (based on these distances) from the gold standard was the estimation of the tumour area when using
 289 each of the benchmarked approaches (i.e. Convex-NMF, Sauwen's and SSSE).

290 The Euclidean distance was calculated as in [38]. Hence, a pair of images E (estimated) and G (gold
 291 standard) having feature vectors f^E and f^G , respectively, have the following distance:

$$\text{Euc_dist}(E, G) = \sqrt{\sum_{i=1}^n (f_i^E - f_i^G)^2} \quad (2)$$

292 Where n is the number of voxels in the images. The distance between two identical images is zero, i.e.
 293 $\text{Euc_dist}(G, G) = 0$; and the larger the value (distance), the bigger the difference between them.

294 Turning the distance measure into a shape similarity score that is easier to interpret, we calculated then
 295 the number of pixels that match values in the two images with reference to the total number of pixels,
 296 giving us a percentage of success.

297 The Hausdorff distance, in turn, measures how far two sets of points are from each other and was used
 298 here to measure the most mismatched tumour delineation between the three segmentations (Convex-
 299 NMF, Sauwen's and SSSE) with respect to the gold standard used. Let us consider B_E and B_G the
 300 boundaries of the tumour areas of the estimated and the gold standard images, respectively, of which
 301 we want to calculate the Hausdorff distance. It was then calculated as follows:

$$\text{Haus_dist}(B_E, B_G) = \max \left\{ \sup_{e \in B_E} \inf_{g \in B_G} \text{dist}(e, g), \sup_{g \in B_G} \inf_{e \in B_E} \text{dist}(e, g) \right\} \quad (3)$$

302 where *sup* represents the supremum and *inf* the infimum. This distance will be zero if and only if B_E
 303 and B_G have the same closure, i.e. both tumour areas are exactly the same.

304 Finally, the Kruskal-Wallis test was used to determine whether the results between the three approaches,
 305 i.e. Convex-NMF, Sauwen's and SSSE, were statistically significant.

306 **Results**

307 The presentation of results is divided into two sections. Firstly, we show the results of applying SSSE
308 to the mice in Groups A and B, which include untreated and treated cases, respectively; and to the mice
309 in Group C, belonging to a longitudinal study with treated mice. (Please refer to Table 1 for more details
310 and references to these groups). This is followed by the evaluation of the presented results at the end of
311 this section.

312 **Brain tumour delineation**

313 Fig 4 shows that the sources obtained with the three approaches, for the different tissue types, are in
314 some cases very similar to the naked eye, which is backed by correlations above 97% in most cases
315 (e.g. comparing the red sources obtained with Convex-NMF, Sauwen's approach and SSSE). However,
316 even when subtle in some cases, these small differences between them are reflected in their
317 corresponding colour-coded maps, indicating the tumour delineation.

318

319 **Fig 4. Results for the cases in Group A, for two classes.**

320 For mouse C179, we also extracted three sources, in order to check whether a higher number of sources
321 could better represent its spatial heterogeneity – i.e. two tumour masses and three histologically
322 different regions with $PI \leq 5\%$ (normal brain), $5\% < PI \leq 30$ (secondary mass) and $PI > 30\%$ (core mass)
323 – see Fig 2. The results are presented in Fig 5. In this situation, both, the sources obtained and the
324 resulting maps, were more visually different. For example, when comparing the red source (representing
325 the main tumour mass) and the blue source (representing the non-tumour area) produced by SSSE with
326 the equivalent sources produced by Convex-NMF, they showed high similarity between themselves
327 (with a correlation of 97% for the main tumour mass and 95% for the normal tissue). However, the
328 yellow source obtained was less similar (with a correlation of 93%). More importantly, these two sets
329 of sources, i.e. red and yellow sources, were mainly representative of different areas of the brain, with
330 Sauwen's approach and SSSE providing a closer resemblance to the T2w image and the

331 histopathological PI values of each mass (see Fig 2) than Convex-NMF. The latter uses mainly the red
332 source to represent both tumour masses, according to the preclinical bioimaging expert, and the yellow
333 source is covering both the solid tumour area and the non-tumour area.

334

335 **Fig 5. Results for case C179 for three classes, from Group A.**

336 Fig 6 shows the results obtained for the cases in Group B. These results were in line to those from Group
337 A (excluding a few cases such as the three sources calculated for mouse C179), in the sense that the
338 sources were not strikingly visually different, except for case C583, in which both the unsupervised
339 tumour (red) and normal (blue) sources corresponded to patterns that matched with low signal-to-noise
340 spectra. This is a problem that has been encountered and characterised before [39]. In the case of the
341 non-tumour (blue) sources, they differed more than the tumour (red) source between the two approaches
342 in all five mice.

343

344 **Fig 6. Results for the cases in Group B, for two classes.**

345 Fig 7 compiles the resulting colour maps (tumour delineation) after applying Convex-NMF, Sauwen's
346 approach and SSSE to the two mice in Group C, at the different time points studied.

347

348 **Fig 7. Results for the cases in Group C, longitudinal study.** Under each case there is a colour-bar showing the
349 response stage determined by the RECIST criteria throughout the therapy protocol (orange means progressive
350 disease, yellow stable disease and green partial response). At the bottom it is indicated when the three TMZ cycles
351 were administered.

352 These colour-coded maps show how the volume of the tumour mass in each mouse changes as a result
353 of their response to the three cycles of therapy. The RECIST criteria [40], which provides an indication
354 of the response evaluation criteria in solid tumours, was determined for these two mice throughout the
355 course of the treatment [26].

356 The adapted RECIST criteria were applied for mouse C819 [26] and in results from day 10 to day 18,
357 the tumour was considered to be at the stage of progressive disease. At day 21, it entered the stable
358 disease stage, remaining in a transient response state until day 25, when tumour shrinkage indicated
359 partial response to the therapy. At days 30 and 33 the tumour was considered to be again in a stage of
360 stable disease, gradually halting its response to the therapy and starting to regrow/relapse. Finally, at
361 days 41 and 45 the tumour was considered again to be at the stage of progressive disease.

362 The adapted RECIST criteria for mouse C821 [26] indicated that the tumour was in the stage of
363 progressive disease from day 10 to 21, in which the tumour grew from occupying ca. 15% of the area
364 of the region of interest to nearly the 90% of it. This stage was followed by a short stable disease stage
365 during days 30-35, followed again by another stage of progressive disease at the last day. A summarized
366 explanation of the adapted RECIST criteria can be found in the Supporting Information file.

367 **Evaluation of the tumour delineation against the gold standard**

368 The results presented in the previous section for the three approaches were quantitatively evaluated
369 using two main criteria: i) their ability to delineate the tumour regions (therefore discriminating the
370 tumour from the non-tumour regions) with high sensitivity and specificity; and ii) their ability to
371 produce colour-coded maps that are as close as possible to the gold standard (see Fig 2), which in this
372 study, it is the set of images provided by the preclinical bioimaging expert (see more details in
373 Histopathology and Evaluation of Results sections from Methods).

374 Details of the sensitivity and specificity values corresponding to the ability of each approach to detect
375 the tumour masses can be found in Tables A in S1 File (for Groups A and B), B in S1 File (for more
376 details on mouse C179 from Group A), and C in S1 File (for Group C). These tables also include the
377 Dice score coefficient calculated for the three methods against the gold standard. A summary of these
378 results is presented in Table 2.

379 **Table 2. Overall sensitivity / specificity and Dice score of the correct delimitation of the tumour mass.**

Groups	Number of masses	Sensitivity / Specificity			Dice score		
		Convex-NMF	Sauwen et al.	SSSE	Convex-NMF	Sauwen et al.	SSSE
A and B	1	0.90 ± 0.08 / 0.68 ± 0.23	0.90 ± 0.08 / 0.76 ± 0.14	0.94 ± 0.05 / 0.82 ± 0.16	0.82 ± 0.11	0.84 ± 0.09	0.90 ± 0.05
A (C179)	2	0.70 ± 0.42 / 0.68 ± 0.04	0.87 ± 0.18 / 0.77 ± 0.11	1.00 ± 0.00 / 0.88 ± 0.05	0.45 ± 0.05	0.63 ± 0.13	0.82 ± 0.04
C	1	0.88 ± 0.17 / 0.60 ± 0.17	0.90 ± 0.08 / 0.72 ± 0.13	0.94 ± 0.10 / 0.81 ± 0.13	0.76 ± 0.13	0.82 ± 0.12	0.89 ± 0.10

380 Mean and standard deviation reported per group. First column indicates the group; second indicates the number
381 of tumour masses; third to fifth columns include the sensitivity / specificity results for the three approaches,
382 respectively; while sixth to eighth show their corresponding Dice scores. Shaded columns highlight the results
383 obtained with SSSE.

384

385 Next, we present the results of the Euclidean distance between each of the three methods to the gold
386 standard, followed by a shape similarity score (see Evaluation of the results in the Methods section),
387 and the Hausdorff distance between the same set of images. The results for the individual cases in
388 Groups A and B are shown in Table S4; while the results for Group C can be seen in Table S5. A
389 summary of them is presented in Table 3.

390 **Table 3. Overall Euclidean distance / shape similarity score (%) and Hausdorff distance.**

Groups	Number of masses	Euclidean distance / Shape similarity score (%)			Hausdorff distance		
		Convex-NMF	Sauwen et al.	SSSE	Convex-NMF	Sauwen et al.	SSSE
A and B	1	99.72 ± 24.99 / 79.07 ± 11.62	84.19 ± 14.46 / 85.72 ± 5.15	73.88 ± 17.67 / 88.75 ± 5.43	9.40 ± 1.35	8.22 ± 1.65	8.18 ± 1.49
C	1	122.23 ± 34.23 / 73.14 ± 14.52	102.41 ± 22.93 / 81.63 ± 7.57	76.76 ± 27.32 / 88.98 ± 8.18	10.22 ± 2.14	7.74 ± 1.80	6.61 ± 2.42

391 Distances between the produced colour-coded maps (Convex-NMF, Sauwen's and SSSE) and the expert's (mean
392 and standard deviation reported per group). Columns 1 and 2 as in Table 2. Third to fifth columns include the
393 Euclidean distance and the shape similarity score for the two approaches, respectively; while sixth to eighth show

394 their corresponding Hausdorff distance. Shaded columns highlight the results obtained with the proposed
395 methodology.

396 The results presented in Tables S4 and S5 were tested for significance (comparing the three approaches
397 against each other) using a Kruskal-Wallis test, to decide whether the population distributions are
398 identical (null hypothesis) without assuming them to follow the normal distribution. They resulted in a
399 p-value < 0.00001 (any p-value < 0.05 is deemed as significant).

400 **Discussion**

401 This section mirrors the structure of the Results section, to facilitate the discussion.

402 **Brain tumour delineation**

403 Starting with a discussion of the tumour delineation results for mice in Groups A and B, we can see that
404 the high resemblance between the sources representing the same tissue type for each approach (Figs 4
405 to 6), is generally because the areas that they are representing are largely the same. However, the
406 seemingly small differences displayed between them have an important impact in the resulting colour-
407 coded maps. This is not surprising as we have seen this effect in the past where very small differences
408 in the spectra led to the characterisation of the therapy response to temozolomide in preclinical
409 glioblastoma in [26]. In this study, SSSE achieved a sharper delimitation of the tumour masses when
410 comparing with Convex-NMF, probably due to the embedded information from the T2w image.
411 Additionally, it is worth mentioning that the inclusion of the MRI information did not affect the
412 biochemical interpretation of the sources, thanks to the way this information was embedded into the
413 model. Moreover, SSSE outperformed its counterpart approach proposed by Sauwen et al. when used
414 in equal conditions (i.e. same data), possibly because the initial information (labelling/segmentation)
415 fed into SSSE is preserving better the spatial configuration of the information.

416 Overall, for mice in Group A, SSSE and Sauwen's approach provided a more meaningful representation
417 than Convex-NMF. This is especially relevant for mouse C179. Metabolically, regarding the spectral

418 pattern represented, the yellow source obtained with Convex-NMF included metabolic features that
419 would be expected both in tumour and non-tumour areas, such as high mobile lipids and high NAA,
420 respectively. See relevant metabolites highlighted in Fig 1 d) and e).

421 In contrast, for mouse C179, SSSE yielded a source (coloured in yellow) that was very similar to the
422 red (tumour) source (Fig 5). The main difference between them was a higher peak signal from
423 lipids/lactate in the red source, which matched with the increased proliferation and necrosis, and the
424 higher choline to creatine ratio (3.21:3.03 ppm). The choline signal had the same intensity of the
425 lipid/lactate signal (1.28/1.33 ppm) in the yellow source, which is indicative of an active tumour, in
426 agreement with the intermediate proliferation indices that were recorded in the yellow region. This goes
427 in line to what we know about this case: this would be indeed expected, as the PIs calculated in zones
428 inside this secondary mass are indicative of tumour tissue, with PI values above 20%. This difference
429 in the spectral patterns are probably responsible for the production of the three coloured regions that
430 closely matched the three histologically different regions. In the case of Sauwen's approach, the
431 obtained results are not too far from those obtained by SSSE, which reinforces the value of using any
432 available knowledge (in this case MRI information) from the case/patient to guide the source extraction
433 in a semi-automated/semi-supervised way.

434 Most importantly, in the case of Convex-NMF, when looking at the area in the colour-coded map that
435 this yellow source is representing (Fig 5), it is considerably far from the secondary tumour mass and it
436 is actually infiltrating non-tumour area in the bottom of the region of interest, which explains why this
437 source exhibits some metabolic features of healthy tissue (see figure 11 available at [25]). The latter is
438 an important result as it shows that with both semi-supervised approaches, the system can be guided to
439 find a representation of the knowledge that the researchers would need modelling, giving them a tool
440 for testing different hypotheses – when left free in a completely unsupervised approach, the result will
441 not necessarily relate to anything of interest as they could be separating regions (e.g. artefacts,
442 ventricles, noise, etc.) that might not be relevant to the particular study.

443 The results for Group B show that, in general, the spectral pattern exhibited a great variation among
444 cases: for example, case C520, C529 and C583 show prominent peaks of lipids/lactate even in the non-

445 tumour region, whereas case C255 and C288 did not. This is not totally unexpected as the first three
446 cases had larger tumour volumes than the other ones, with most voxels within the MRSI grid being
447 represented either by the tumour itself or by the peritumoral infiltrating zone. Fig 8 shows a selection
448 of voxels from one of these mice, i.e. C520, in which a high peak of lipids/lactate at c.a. 1.3 ppm can
449 be seen, especially prominent in voxels a and b which could be due to tumour tissue infiltration.

450

451 **Fig 8. MR spectra of three selected voxels from the non-tumour area.**

452 However, even for the cases in which the similarities between the sources are higher, again the resulting
453 maps were quite different, as seen in the results reported for the Group A. Regardless of this apparent
454 high similarity, the colour maps show a different representation of the tumour areas for each approach,
455 which are visually more coincident with the initial manual segmentation, but with the benefit that they
456 also include the researcher-dubious (grey/uncertain) areas.

457 In addition to all the aspects mentioned before, both semi-supervised approaches have shown the ability
458 to learn more meaningful, better-quality sources, as a way to overcome the susceptibility to the presence
459 of artefacts and the lower signal-to-noise spectra issues reported in [39], with SSSE providing more
460 accurate results according to the gold standard.

461 The delineation of the tumour area in the maps for the Group C show that, as observed for Groups A
462 and B, SSSE results in more coincidental areas with the abnormal regions. Additionally, they are in
463 accordance with the RECIST criteria for each of these images. Indeed, for both mice, there was a much
464 higher correspondence between the RECIST criteria, and the maps obtained when using SSSE.

465 **Evaluation of the tumour delineation against the gold standard**

466 In most cases of Groups A and B, the sensitivity and specificity when using SSSE was better than the
467 other two approaches (Tables 2 and A in S1 File), and in some cases by a large difference. The
468 exceptions of a higher sensitivity with Convex-NMF and/or Sauwen's approach were seen in mice
469 C179, C520 and C583, but these are at the cost of a much reduced specificity. If we look closer at mouse

470 C520, for example, we can see that it showed some atypical features around the ventricle which the
471 imaging expert marked as abnormal, although this does not necessarily mean that it corresponds to core
472 tumour area, with $PI > 30\%$. SSSE then benefitted from representing the tumour as a larger volume, but
473 failed to do it accurately, as can be seen with a drop in specificity. In the case of specificity, SSSE also
474 exhibited a better performance overall, only running aground with case C529 in which it also failed to
475 identify the abnormal features around the ventricle (see Fig S4, supporting information), as described
476 by the preclinical bioimaging expert, possibly because they are not part of the core tumour area with PI
477 $> 30\%$.

478 When looking at the results of the analysis of the two abnormal masses from mouse C179 (see Fig 5
479 and Tables 2 and B in S1 File), we can see that in cases such as this one, semi-supervised approaches
480 can make a huge difference when creating a model that represents areas of interest (notice that the
481 resulting areas in the colour map, first row of Fig 5, do not represent such areas), with overall sensitivity
482 and specificity of SSSE (1.00 and 0.88, respectively) outperforming Sauwen's approach (0.87 and 0.77,
483 respectively) and Convex-NMF (0.70 and 0.68, respectively).

484 The sensitivity and specificity results for Group C are also consistent with the ones of Groups A and B.
485 A detailed discussion for this group can be found in the Supporting Information file (Discussion of the
486 tumour delineation sensitivity and specificity for Group C).

487 The Dice scores also show agreement with what was discussed previously. The semi-supervised
488 approaches greatly outperform the unsupervised one, with SSSE exhibiting better performance. As
489 recommended by Zijdenbos et al [41] in the literature of image validation, a good overlap occurs when
490 $DSC > 0.70$, which has been attained by the three approaches.

491 The Euclidean distance between the SSSE resulting maps and the gold standard were smaller in most
492 of the cases than the distances between the maps obtained with the other two approaches and the gold
493 standard (see Tables S4 and S5), which means that SSSE delineation of the tumour was more accurate
494 than the other methods. There were only four exceptions to this (from all the 28 delineation maps
495 produced in this study with the proposed methodology), two maps in group B, and another two in Group
496 C. The ones in Group B were for mice C288 and C529, for which the Euclidean distances were slightly

497 worse for SSSE, however the Hausdorff distance for both cases indicated otherwise. In addition, our
498 proposed method did spot the anomaly in the ventricles of C529, which was also identified by the expert
499 pathologist when analysing the samples *a posteriori*. The exceptions in Group C were case C819 at day
500 41 and case C821 at day 21, in which the differences were marginal, and again the Hausdorff distances
501 contrarily show a better performance by SSSE. It is worth mentioning that the distance between two
502 images will indicate how much they differ, meaning that the shorter the distance between them the
503 higher their similarity.

504 Therefore, considering that a) most of the delineation maps produced by SSSE showed a notorious
505 improvement over the unsupervised approach, and b) the maps produced by the counterpart, semi-
506 supervised approach proposed by Sauwen et al. did not generally outperform SSSE, we can confidently
507 say that overall, the distance of the maps produced by SSSE to the gold standard were considerably
508 shorter, meaning that SSSE was better suited for the problem and data at hand.

509 Previous results for the Euclidean distances between the three approaches to the gold standard
510 harmonise with those obtained for the shape similarity score, as these two evaluation methods are highly
511 related, except that the latter provides a better understanding and interpretability of what these distances
512 mean. Overall, for Groups A and B, the proposed methodology exhibited an improvement over the
513 unsupervised approach in the shape similarity score of nearly 10%; and Group C showed an
514 improvement of more than 15% (both with a much smaller standard deviation).

515 The results presented in Tables S4 and S5 (referring to the Euclidean distance, similarity score, and
516 Hausdorff distance between the three approaches to the gold standard) were tested for significance using
517 a Kruskal-Wallis test, obtaining a p-value < 0.00001 , which indicates that the difference between these
518 approaches are statistically significant (as $p < 0.05$). Furthermore, statistically significant differences
519 were found between the 28 delineation maps produced by the unsupervised approach and Sauwen's
520 approach in comparison to the 28 produced by SSSE.

521 One final note goes to the fact that the proposed methodology does not combine MRI and MRSI
522 information by concatenation, nor it constrains the MRSI to fit the MRI segmentation, as opposed to
523 several previous works, such as [3,20–22]. Instead, our proposed methodology embeds the information

524 coming from the MRI (e.g. the manually selected areas) into the analysis of the MRSI by guiding the
525 source extraction (which are MR spectra) in the direction of the areas of interest, according to the MR
526 image. Our reason to avoid direct concatenation of the information from these MR approaches was to
527 propose a model able to integrate them in a principled manner. Moreover, we use the full MR spectra
528 in the usual range of interest [4.5 – 0 ppm], as opposed to quantifying only a selection of metabolites
529 as in [3,20,21]. The added value of the latter is interpretability, providing not only the visualisation of
530 the resulting map with the delineation of the tumour areas and healthy parenchyma, but also the spectral
531 pattern associated to each of these segmented regions. This allows for extra validation and reassurance,
532 and more importantly, better understanding of the results.

533 **Conclusions**

534 Overall, the quality of the resulting colour-coded maps indicating the tumour delineation with the
535 proposed methodology was considerably better than when using only Convex-NMF in a completely
536 unsupervised approach, as in [25], or even an alternative semi-supervised approach proposed by
537 Sauwen et al. [22]. This was carefully assessed in a quantitative way, showing that, in most cases, the
538 sensitivity and specificity of delineating the tumour masses (which provide a measure of accuracy and
539 confidence in these delineations) was far superior when using the proposed methodology, SSSE. These
540 results were also consistent with the shape similarity scores and distances calculated for both sets of
541 images (in both approaches) to the gold standard (images from the imaging expert), which can be
542 considered statistically significant.

543 These results also come with additional advantages, namely: a) the proposed methodology is able to
544 effectively deal with signal-to-noise issues, which is not the case of the unsupervised approach in [25];
545 b) it allows radiologists/clinicians to define the area of interest to them and, with that, guide the process
546 of source extraction, which was not a possibility in [25]; c) while still creating an intra-subject model,
547 as opposed to [26] where the model was trained using a set of subjects, therefore needing to deal with
548 the tumour heterogeneity of GB, which is a well-recognised problem [42], and d) in addition to a
549 segmented images, SSSE also produce meaningful, good-quality sources that represent each of the

550 regions of interest, which adds an extra layer of interpretability to the results obtained, conferring not
551 only validation of the results, but also better understanding and analysis of each individual case.
552 Therefore, given the quality of the obtained results, and the advantages identified in the use of the
553 proposed methodology, we consider that the extra pre-processing steps related to the embedding of the
554 MRI information into the MRSI data source extraction are worthwhile, as they have shown to improve
555 the tumour delineation in the preclinical GB model. Once again, the multiparametric character of the
556 proposed approach (fusing MRI and MRSI) has shown to provide better results compared with using a
557 single approach (MRSI) [26].

558 **Abbreviations**

559 **FI**: Fisher Information. **GB**: Glioblastoma. **MDS**: Multi-dimensional Scaling. **MLP**: Multi-Layer
560 Perceptron. **MR**: Magnetic Resonance. **MRI**: Magnetic Resonance Imaging. **MRS**: MR spectroscopy.
561 **MRSI**: Magnetic Resonance Spectroscopic Imaging. **NMF**: Non-negative Matrix Factorisation. **PI**:
562 proliferation index. **RECIST**: response evaluation criteria in solid tumours. **SSSE**: semi-supervised
563 source extraction. **T2w**: T2-weighted.

564 **Authors' contributions**

565 SOM and IO conceived and designed the study. SOM implemented the methods and carried out the
566 experiments. PR assisted with the experiments carried out with Sauwen et al. approach. APC performed
567 the initial manual segmentations of mice from studies B and C. APC and MJS provided the data and
568 interpretations to the results. RT and PR assisted with the evaluation of the results. SOM, IO, PR, APC
569 and MSJ helped to draft the manuscript. All authors read and approved the final version of it.

570 **References**

- 571 1. Julià-Sapé M, Acosta D, Majós C, Moreno-Torres A, Wesseling P, Acebes JJ, et al. Comparison
572 between neuroimaging classifications and histopathological diagnoses using an international

- 573 multicenter brain tumor magnetic resonance imaging database. *J Neurosurg.* 2006;105: 6–14.
- 574 2. Segebarth CM, Balériaux DF, Luyten PR, Den Hollander JA. Detection of metabolic
575 heterogeneity of human intracranial tumors in vivo by ¹H NMR spectroscopic imaging. *Magn*
576 *Reson Med.* 1990;13: 62–76.
- 577 3. Luts J, Laudadio T, Idema AJ, Simonetti AW, Heerschap A, Vandermeulen D, et al. Nosologic
578 imaging of the brain: segmentation and classification using MRI and MRSI. *NMR Biomed.*
579 2009;22: 374–390. doi:10.1002/nbm.1347
- 580 4. Wen PY, Macdonald DR, Reardon DA, Cloughesy TF, Sorensen AG, Galanis E, et al. Updated
581 response assessment criteria for high-grade gliomas: response assessment in neuro-oncology
582 working group. *J Clin Oncol.* 2010;28: 1963–72. doi:10.1200/JCO.2009.26.3541
- 583 5. Widhalm G, Minchev G, Woehrer A, Preusser M, Kiesel B, Furtner J, et al. Strong 5-
584 aminolevulinic acid-induced fluorescence is a novel intraoperative marker for representative
585 tissue samples in stereotactic brain tumor biopsies. *Neurosurg Rev.* 2012;35: 381–391.
586 doi:10.1007/s10143-012-0374-5
- 587 6. Stupp R, Mason WP, van den Bent MJ, Weller M, Fisher B, Taphoorn MJB, et al. Radiotherapy
588 plus Concomitant and Adjuvant Temozolomide for Glioblastoma. *N Engl J Med.* 2005;352(10):
589 987–996.
- 590 7. Brandes AA, Tosoni A, Spagnolli F, Frezza G, Leonardi M, Calbucci F, et al. Disease
591 progression or pseudoprogression after concomitant radiochemotherapy treatment: pitfalls in
592 neurooncology. *Neuro Oncol.* 2008;10: 361–7. doi:10.1215/15228517-2008-008
- 593 8. Ortega-Martorell S, Julià-Sapé M, Lisboa P, Arús C. Pattern Recognition Analysis of MR
594 Spectra. *eMagRes.* Chichester, UK: John Wiley & Sons, Ltd; 2016. pp. 945–958.
595 doi:10.1002/9780470034590.emrstm1484
- 596 9. De Edelenyi FS, Rubin C, Estève F, Grand S, Décorps M, Lefournier V, et al. A new approach
597 for analyzing proton magnetic resonance spectroscopic images of brain tumors: nosologic
598 images. *Nat Med.* 2000;6: 1287–1289.

- 599 10. Simonetti AW, Melssen WJ, van der Graaf M, Postma GJ, Heerschap A, Buydens LMC. A
600 Chemometric Approach for Brain Tumor Classification Using Magnetic Resonance Imaging
601 and Spectroscopy. *Anal Chem.* 2003;75: 5352–5361.
- 602 11. Ortega-Martorell S, Vellido A, Lisboa PJG, Julià-Sapé M, Arús C. Spectral decomposition
603 methods for the analysis of MRS information from human brain tumors. *International Joint*
604 *Conference on Neural Networks (IJCNN 2011), Proceedings.* San José, California, USA; 2011.
605 pp. 3279–3284.
- 606 12. Ortega-Martorell S, Lisboa PJG, Vellido A, Julià-Sapé M, Arús C. Non-negative Matrix
607 Factorisation methods for the spectral decomposition of MRS data from human brain tumours.
608 *BMC Bioinformatics.* 2012;13.
- 609 13. Ortega-Martorell S, Ruiz H, Vellido A, Olier I, Romero E, Julià-Sapé M, et al. A Novel Semi-
610 Supervised Methodology for Extracting Tumor Type-Specific MRS Sources in Human Brain
611 Data. *PLoS One.* 2013;8: e83773. doi:10.1371/journal.pone.0083773
- 612 14. Ortega-Martorell S, Olier I, Julià-Sapé M, Arús C. SpectraClassifier 1.0: a user friendly,
613 automated MRS-based classifier-development system. *BMC Bioinformatics.* 2010;11: 106.
- 614 15. García-Gómez JM, Luts J, Julià-Sapé M, Krooshof P, Tortajada S, Robledo J, et al. Multiproject-
615 multicenter evaluation of automatic brain tumor classification by magnetic resonance
616 spectroscopy. *Magn Reson Mater Physics, Biol Med.* 2009;22: 5–18.
- 617 16. García-Gómez JM, Tortajada S, Vidal C, Julià-Sapé M, Luts J, Moreno-Torres A, et al. The
618 effect of combining two echo times in automatic brain tumor classification by MRS. *NMR*
619 *Biomed.* 2008;21: 1112–1125.
- 620 17. De Vos M, Laudadio T, Simonetti A, Heerschap A, Van Huffel S. Fast nosologic imaging of the
621 brain. *J Magn Reson.* 2007;184: 292–301.
- 622 18. Sauwen N, Acou M, Van Cauter S, Sima DM, Veraart J, Maes F, et al. Comparison of
623 unsupervised classification methods for brain tumor segmentation using multi-parametric MRI.
624 *NeuroImage Clin.* 2016;12: 753–764. doi:10.1016/j.nicl.2016.09.021

- 625 19. Postma GJ, Luts J, Idema AJ, Julià-Sapé M, Moreno-Torres Á, Gajewicz W, et al. On the
626 relevance of automatically selected single-voxel MRS and multimodal MRI and MRSI features
627 for brain tumour differentiation. *Comput Biol Med.* 2011;41: 87–97.
628 doi:10.1016/j.compbiomed.2010.12.003
- 629 20. Sauwen N, Sima DM, Van Cauter S, Veraart J, Leemans A, Maes F, et al. Hierarchical non-
630 negative matrix factorization to characterize brain tumor heterogeneity using multi-parametric
631 MRI. *NMR Biomed.* 2015;28: 1599–1624. doi:10.1002/nbm.3413
- 632 21. Li Y, Liu X, Wei F, Sima DM, Van Cauter S, Himmelreich U, et al. An advanced MRI and
633 MRSI data fusion scheme for enhancing unsupervised brain tumor differentiation. *Comput Biol*
634 *Med.* 2017;81: 121–129. doi:10.1016/J.COMPBIOMED.2016.12.017
- 635 22. Sauwen N, Acou M, Sima DM, Veraart J, Maes F, Himmelreich U, et al. Semi-automated brain
636 tumor segmentation on multi-parametric MRI using regularized non-negative matrix
637 factorization. *BMC Med Imaging.* 2017;17: 29. doi:10.1186/s12880-017-0198-4
- 638 23. Sorber L, Van Barel M, De Lathauwer L. Structured Data Fusion. *IEEE J Sel Top Signal*
639 *Process.* 2015;9: 586–600. doi:10.1109/JSTSP.2015.2400415
- 640 24. Cha S, Johnson G, Wadghiri YZ, Jin O, Babb J, Zagzag D, et al. Dynamic, contrast-enhanced
641 perfusion MRI in mouse gliomas: Correlation with histopathology. *Magn Reson Med.* 2003;
642 doi:10.1002/mrm.10446
- 643 25. Ortega-Martorell S, Lisboa PJG, Simões R V, Pumarola M, Julià-Sapé M, Arús C. Convex Non-
644 Negative Matrix Factorization for brain tumor delimitation from MRSI data. *PLoS One.* 2012;7:
645 e47824.
- 646 26. Delgado-Goñi T, Ortega-Martorell S, Ciezka M, Olier I, Candiota AP, Julià-Sapé M, et al.
647 MRSI-based molecular imaging of therapy response to temozolomide in preclinical
648 glioblastoma using source analysis. *NMR Biomed.* 2016;29: 732–743. doi:10.1002/nbm.3521
- 649 27. Simões R V., Delgado-Goñi T, Lope-Piedrafita S, Arús C. ¹H-MRSI pattern perturbation in a
650 mouse glioma: the effects of acute hyperglycemia and moderate hypothermia. *NMR Biomed.*

651 2010;23: 23–33. doi:10.1002/nbm.1421

652 28. Delgado-Goñi T, Julià-Sapé M, Candiota AP, Pumarola M, Arús C. Molecular imaging coupled
653 to pattern recognition distinguishes response to temozolomide in preclinical glioblastoma. *NMR*
654 *Biomed.* 2014;27: 1333–1345. doi:10.1002/nbm.3194

655 29. Ortega-Martorell S, Olier I, Delgado-Goñi T, Ciezka M, Julià-Sapé M, Lisboa P, et al. Semi-
656 supervised source extraction methodology for the nosological imaging of glioblastoma response
657 to therapy. 2014 IEEE Symposium on Computational Intelligence and Data Mining,
658 Proceedings. 2015. pp. 93–98. doi:10.1109/CIDM.2014.7008653

659 30. Ostertag CB, Volk B, Shibata T, Burger P, Kleihues P. The monoclonal antibody Ki-67 as a
660 marker for proliferating cells in stereotactic biopsies of brain tumours. *Acta Neurochir (Wien)*.
661 1987;89: 117–121.

662 31. Wang T-C, Cheng C-Y, Yang W-H, Chen W-C, Chang P-J. Characterization of highly
663 proliferative secondary tumor clusters along host blood vessels in malignant glioma. *Mol Med*
664 *Rep.* 2015;12: 6435–6444. doi:10.3892/mmr.2015.4228

665 32. Paatero P, Tapper U. Positive matrix factorization: A non-negative factor model with optimal
666 utilization of error estimates of data values. *Environmetrics.* 1994;5: 111–126.

667 33. Lee DD, Seung HS. Learning the parts of objects by non-negative matrix factorization. *Nature*.
668 1999;401: 788–791.

669 34. Ding C, Li T, Jordan MI. Convex and semi-nonnegative matrix factorizations. *IEEE Trans*
670 *Pattern Anal Mach Intell.* 2010;32: 45–55.

671 35. Ruiz H, Jarman IH, Martín JD, Lisboa PJG. The role of Fisher information in primary data space
672 for neighbourhood mapping. European Symposium on Artificial Neural Networks,
673 Computational Intelligence and Machine Learning (ESANN). Bruges, Belgium; 2011. pp. 381–
674 6.

675 36. Zou KH, Warfield SK, Bharatha A, Tempany CMC, Kaus MR, Haker SJ, et al. Statistical

- 676 Validation of Image Segmentation Quality Based on a Spatial Overlap Index. *Acad Radiol.*
677 2004;11: 178–189. doi:10.1016/S1076-6332(03)00671-8
- 678 37. Dice LR. Measures of the Amount of Ecologic Association Between Species. *Ecology.* 1945;
679 doi:10.2307/1932409
- 680 38. Safar MH, Shahabi C. Image Similarity Measures. In: *Shape Analysis and Retrieval of*
681 *Multimedia Objects. Multimed Syst Appl.* 2003;23: 9–11. doi:10.1007/978-1-4615-0349-1_2
- 682 39. Mocioiu V, Kyathanahally SP, Arús C, Vellido A, Julià-Sapé M. Automated Quality Control for
683 Proton Magnetic Resonance Spectroscopy Data Using Convex Non-negative Matrix
684 Factorization. *Bioinformatics and Biomedical Engineering Proceedings of IWBBIO 2016.*
685 Granada, Spain: Springer; 2016. pp. 719–727. doi:10.1007/978-3-319-31744-1_62
- 686 40. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response
687 evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). *Eur J Cancer.*
688 2009;45: 228–247. doi:10.1016/j.ejca.2008.10.026
- 689 41. Zijdenbos AP, Dawant BM, Margolin RA, Palmer AC. Morphometric Analysis of White Matter
690 Lesions in MR Images: Method and Validation. *IEEE Trans Med Imaging.* 1994;
691 doi:10.1109/42.363096
- 692 42. Sottoriva A, Spiteri I, Piccirillo SGM, Touloumis A, Collins VP, Marioni JC, et al. Intratumor
693 heterogeneity in human glioblastoma reflects cancer evolutionary dynamics. *Proc Natl Acad Sci*
694 *U S A.* 2013;110: 4009–4014. doi:10.1073/pnas.1219747110

695

696 **Supporting Information**

697 **S1 File. Supporting Information: Embedding MRI information into MRSI data source extraction**
698 **improves brain tumour delineation in animal models.** This document includes additional
699 information regarding the Magnetic Resonance studies, TMZ administration and preparation, how

700 Sauwen et al. method was used, how the initial segmentation was performed, and further details on the
701 evaluation and discussion of the results.