RESEARCH ARTICLE

Two-dimensional linear-combination model fitting of magnetic resonance spectra to define the macromolecule baseline using FiTAID, a Fitting Tool for Arrays of Interrelated Datasets

**Daniel G. Q. Chong · Roland Kreis ·
Christine S. Bolliger · Chris Boesch ·
Johannes Slotboom**

Abstract

Object To propose the determination of the macromolecular baseline (MMBL) in clinical 1H MR spectra based on T_1 and T_2 differentiation using 2D fitting in FiTAID, a general Fitting Tool for Arrays of Interrelated Datasets.

Materials and methods Series of localized inversion-recovery (IR) and 2DJ separation spectra of the brain were recorded at 3T. The MMBL was determined by three 2D evaluation methods based on (1) IR spectra only, (2) 2DJ spectra only, (3) both IR and 2DJ spectra (2DJ-IR). Their performance was compared using synthetic spectra and based on variability and reproducibility as obtained in vivo from 12 subjects in 20 examinations.

Results All methods performed well for synthetic data. In vivo, 2DJ-only yielded larger variations than the other methods. IR-only and 2DJ-IR yielded similar performance. FiTAID is illustrated with further applications where linear-combination model fitting of interrelated arrays of spectra is advantageous.

Conclusion 2D-Fitting offers the possibility to determine the MMBL based on a range of complementary experimental spectra not relying on smoothness criteria or global assumptions on T_1. Since 2DJ-IR includes information from spectra with different inversion and echo times, it is expected to be more robust in cases with more variable data quality and overlap with lipid resonances.

D. G. Q. Chong · R. Kreis (✉) · C. S. Bolliger · C. Boesch
Department of Clinical Research, Unit for MR-Spectroscopy and
Methodology, University Bern, Inselspital, P.O. Box 35,
3010 Bern, Switzerland
e-mail: roland.kreis@insel.ch

J. Slotboom
Department of Radiology, Neuroradiology, and Nuclear Medicine,
University Hospital Bern, 3010 Bern, Switzerland

Introduction

The large majority of in vivo magnetic resonance spectroscopy (MRS) examinations yield simple one-dimensional (1D) spectra (possibly from a multidimensional spatial array). Though there has been steady progress for localized two-dimensional (2D) spectroscopy, 2D and multidimensional methods have mainly remained in the realm of high-resolution spectroscopy. This is partly due to a lack in post-processing and fitting tools for quantitative evaluations of multidimensional MRS datasets. For clinical 1D MR spectra, modeling in the form of linear combinations of model spectra has become the gold standard and several software packages have been described and are available on a commercial or non-commercial basis [1–3] (recently reviewed in [4]). 2D modeling methods have only been used in isolated methodological studies [5–12] and the required tools are not broadly available. Furthermore, previous work was mostly focused on a single type of 2D experiment. However, simultaneous 2D fitting is not only beneficial in the obvious case of 2D MRS (e.g. separation or correlation spectroscopy) but also in any case of complementary interrelated MR spectra, where one gains in effective signal to noise (SNR) when combining all spectra in a single fit [8,13].

Here, the benefit of multidimensional fitting is demonstrated for the determination of the macromolecular baseline (MMBL) in ^{1}H MRS brain spectra. The broad MMBL underlying the metabolite spectra is probably the biggest confounding factor for accurate quantitation of short echo time spectra in vivo. Experimentally, it has been estimated with

the metabolite nulling technique [14] using single inversion-recovery spectra or using an extended set of saturation recovery spectra [15]. The former technique is not ideal because there is no single inversion time that can null all metabolite signals simultaneously, while the latter is time-consuming because it needs a multitude of spectra with very different TR. Alternatively, post-processing techniques [4,16–19] can be employed to define or eliminate this baseline, but all essentially rely on basic differences in signal shape between MMBL and narrow metabolite peaks, which are not guaranteed in all cases. Using a Fitting Tool for Interrelated Arrays of Datasets (FiTAID), the MMBL in localized brain spectra was determined and compared as obtained from different experimental datasets: (1) from a series of inversion-recovery spectra (i.e. using differences in T_1); (2) from a 2D-J separation spectrum [20–22] (i.e. based on differences in T_2); and 3) from a combination of these two datasets (using T_1 and T_2 differentiation).

In addition, FiTAID, a new flexible fitting tool, is being presented. It can equally well be used for traditional linear combination model fitting of one-dimensional spectra as well as for linear combination fitting of two-dimensional arrays of interrelated MRS datasets. Its major aim is to provide an extendable framework to fit general interrelated spectra by offering different model structures and prior knowledge connections in both dimensions. It offers a multitude of fitting options for the expert in order to optimize the fitting procedure for each special case. This includes the novel option of fitting either in time domain, or in arbitrary spectral windows in the frequency domain, or even subsequently in either domain according to arbitrary fit strategy steps. It also allows the user to fit limited parameter sets or limited parts of the 2D data matrix.

After presentation of the modeling background[1], some illustrations of the principles of the fitting tool are given along with some well-known examples of 1D and 2D fitting. The main emphasis of the presented work, however, lies on the evaluation of the MMBL as obtained from different experimental strategies and its variability and reproducibility in vivo and in synthetic data.

Materials and methods

Subjects

The local ethical committee approved the study protocol and written informed consent was obtained from all subjects. Data for the main application (definition of MMBL) were obtained from 12 healthy subjects (8 male, 4 female, 27 ± 8 years old), examinations were repeated 5 times in 2 of these

subjects. Further 5 subjects were examined to acquire data for the illustrations of FiTAID.

Data acquisition

All spectra were recorded on a Siemens 3 T Tim TRIO system (Siemens, Erlangen, Germany) using the manufacturers transmit/receive head coil and a locally developed localization sequence based on the product PRESS sequence. The local adaptations of the PRESS sequence include the features (1) to increment the echo time regularly by arbitrary delays to enable 2DJ spectroscopy or to allow to record a series of specifically preselected echo times, (2) to include a selective adiabatic inversion pulse [23] directly before the PRESS sequence to record non-water suppressed MR spectra according to the scheme of de Graaf [24], (3) to automatically record non-water suppressed reference scans at different echo times, (4) to include an adiabatic global pre-inversion pulse for IR spectroscopy with selectable delays, (5) to allow for extended phase cycling or phase rotation, (6) to store each acquisition before summation.

Experimental details are given in the text and figure captions for each of the illustrative examples later in the Methods section, whereas experimental parameters for the main application are listed in the following.

Axial fast spin echo and sagittal 3D FLASH MR images were acquired for the visualization of the placement of a $25 \times 25 \times 25\,\mathrm{mm}^3$ voxel in occipital cortex. The Siemens auto align feature was used to ensure consistent placement of the voxel in different subjects and reproducible placement for repeated scans (to within $\sim 1\,\mathrm{mm}$ in all directions for the particular ROI used in this application). To ensure exact 90-degree excitation pulses in PRESS, the transmitter voltage was corrected using the manufacturers B_1 map tool. Shimming was performed with the FASTESTMAP technique [25]. All spectra were recorded without water presaturation, but instead with the scheme by de Graaf [24] using a 22 ms long metabolite inversion pulse [23] directly before the PRESS sequence. With a bandwidth of 4,000 Hz 4,096 points were acquired. Specific details for 2DJ: 24 equally spaced echo times between 20 and 307.5 ms, 16 acquisitions per TE, TR 2,000 ms, total acquisition time 13 min. For IR: TE 20 ms, TR 2,000 ms, and inversion times TI of 30, 200, 450, 575, 700, 825, and 1,200 ms with 48 acquisitions each, total acquisition time 11 min. In addition, eight water spectra with varying TE between 20 and 1,500 ms were acquired to determine brain compartments for metabolite quantitation.

Data processing

Processing conditions are given in the text and figure captions for each of the illustrative examples in the Methods section,

---

[1] Implementation specifics are given in an Appendix.

whereas processing details for the main application are listed in the following.

The non-water suppressed spectra were summed with equal or alternating 180-degree phase to obtain water or metabolite spectra. The metabolite spectra were eddy current corrected [26] and zero-order phase corrected by subtraction of the instantaneous phase of the reference water signal between the start of the FID and a time where the phase of the water signal was ill-defined due to low SNR. To simplify the fitting model, the residual water peak was removed from the metabolite spectra with HLSVD filtering with up to 25 lines between 4.5 ppm and 10 ppm using jMRUI [3].

Three 2D methods of linear combination model fitting in FiTAID were employed, all using the same simulated metabolite basis set. IR-only using a 2D IR fit of the IR data only, 2DJ-only using a 2D fit of the multiple TE spectra in the 2DJ data series, and a combined method called 2DJ-IR where the parameter model is iteratively improved by going back and forth between the two datasets, one distinguishing metabolites from MMBL based on T_1 differences (IR) and the other based on T_2 effects (2DJ).

The details for the three fits in FiTAID are given in Appendix A. In order to be able to quantify the baseline, an analytical model for the MMBL was used in which it was described as a linear combination of equally spaced Voigt lines.

Synthetic spectra

Synthetic spectra were used to create a set of spectra with known true values where fit results could be tested for systematic deviations from true values. The spectra were set up as a linear combination of metabolites and a known macromolecular baseline. Metabolites included the following: glycero-phosphorylcholine (GPC), phosphorylcholine (PC), creatine (Cr), ethanolamine, GABA, glutamate, glutamine, glutathione, glycine, myo-inositol (mI), N-acetylaspartate (NAA), scyllo-inositol, and taurine. Relative concentrations of metabolites and macromolecules were used as typically obtained when fitting the in vivo spectra described in the main application. The following relaxation times, as typically obtained from fitting the experimental in vivo data sets, were used: MMBL T_2 22 ms, T_1 240 ms; NAA (CH$_3$) T_2 234 ms, T_1 1,540 ms; cholines (CH$_3$) T_2 209 ms, T_1 920 ms, Cr(CH$_3$) T_2 135 ms, T_1 1,040 ms; Cr(CH$_2$) T_2 117 ms, T_1 880 ms, all other metabolites T_2 108 ms, T_1 1,100 ms. The same parameter relations were applied as for the fitting defined in Appendix A. An SNR of 10 in time domain for the non-inverted TE 20 ms spectrum (i.e. first point of FID versus standard deviation of the noise) was used for the main comparisons since this approximates the SNR of the experimental data in the TE 20 ms spectrum of the 2DJ scan. Three sets of spectra with 10 random noise realizations each were simulated for the iterative 2DJ-IR, IR-only and 2DJ-only methods.

Unlike the in vivo data, the SNR was adjusted such that it corresponded to equal measurement time spent for each method.

Statistics

The performance of the three methods was judged based on the differences and variability of the resulting frequency domain spectrum at TE 20 ms for the MMBL, visually using plots of mean ± 1SD and numerically by calculating average standard deviations over the whole spectrum. Univariate two-way ANOVA (with the main effects methods and subjects or repetitions, but no cross-term) was used spectral point-by-point (from −0.1 to 4.8 ppm; 0.016 ppm/point) to test for significant differences between the resulting MMBLs from the three methods. A post hoc analysis was used to identify which method was responsible for the significant differences. $P < 0.05$ was considered significant, but—given the multitude of point-by-point tests—results were only considered truly significant if more than an arbitrary threshold of 5% of the point-by-point tests yielded $P < 0.05$. Levene's test for homogeneity of variance [27] was applied to determine whether the methods differ with respect to variance.

For the simulated spectra, the absolute value of the difference between the true and the fitted MMBL was calculated and the mean result compared between the methods, while a Wilcoxon sign test was used to determine whether the median resulting value at any of the spectral points for any of the methods differed significantly from the true values.

Theory of signal models, prior knowledge constraints, and fit strategies

The basic experimental data structure as obtained in the measurement domain is treated in the following as a two-dimensional array of complex numbers, recorded as a series of 1-dimensional data trains.

1D model

A Lorentzian line describes the basic free induction decay (FID), however, without accounting for magnetic field inhomogeneity effects. A Voigt line shape [Refs. in [2]] is a more general model with an additional Gaussian width parameter that can accommodate an imperfect lineshape in a first approximation. Hence, it was chosen as the basic model, given in time domain (TD) as:

$$\xi^V(t_n) = A \cdot \exp(-i\omega t_n) \cdot \exp(-\alpha t_n - \beta^2 t_n^2) \cdot \exp(i\phi) \quad (1)$$

where in the following $\xi(t_n)$ always stands for a digitized model signal, the (first) superscript (here V) for the type of signal entity. (As explained below, signal entities covered in this context, include simple Voigt lines, but also collections of

Voigt lines referred to as parametric patterns, numeric patterns or collections of these entities to describe complete metabolite spectra)

$t_n = n \cdot \Delta t$, $n = 0,1,\dots n$ reflects the sampling times, Δt the sampling interval, A the area, ω the offset frequency, α the exponential or Lorentzian damping, β^2 the Gaussian damping, and ϕ the signal phase.

A standard linear combination model in 1D will then be

$$\xi(t_n) = \sum_{m=1}^{M} \xi_m(t_n) \tag{2}$$

where M is the number of entities in the model. In TD, the cost function is

$$\chi^2 = \sum_{n=ns}^{ne} (x(t_n) - \xi(t_n))^2 \tag{3}$$

where $x(t_n)$ is the experimental data in TD, n runs over the fitting range, which excludes truncation ranges at the start (ns) and end (ne) of the TD signal, where the experimental data may deviate from the theoretical model (e.g. phase distortions from filtering) or not add any information (noisy signal tail).

For frequency domain (FD) fitting, the points "truncated" in TD are replaced by zeroes for both the TD model and the experimental TD data before Fourier transformation.

$$x(t_n) = \begin{cases} 0 \\ x(t_n) \end{cases}; \ \xi(t_n) = \begin{cases} 0 & n < ns, n > ne \\ \xi(t_n) & ns \le n \le ne \end{cases} \tag{4}$$

and

$$\chi^2 = \sum^{ranges} (FFT(x(t_n)) - FFT(\xi(t_n)))^2 \tag{5}$$

where *ranges* are the FD ranges (i.e. multiple non-overlapping regions of the spectrum) selected to fit and FFT stands for fast Fourier transformation.

Prior knowledge in 1D

Prior knowledge is a term for presumably known relationships between model parameters that can be enforced during fitting and result in a reduced number of fitting parameters, increased model robustness, and reduced error bounds. Prior knowledge makes a set of model parameters depend on the free parameters. The following are the prior knowledge relations implemented for each parameter type:

$$A_d = R_d \cdot A_i \tag{6}$$
$$\omega_d = \Delta\omega_d + \omega_i \tag{7}$$
$$\alpha_d = \Delta\alpha_d + \alpha_i \tag{8}$$
$$\beta_d^2 = \Delta\beta_d^2 + \beta_i^2 \tag{9}$$
$$\phi_d = \Delta\phi_d + \phi_i \tag{10}$$

where R_d is the ratio between area parameters and $\Delta\omega_d, \Delta\alpha_d$, $\Delta\beta_d^2$, and $\Delta\phi_d$ are shifts between frequencies, Lorentzian dampings, Gaussian dampings, and phases. Subscript d refers to the dependent variables and subscript i to the independent variables.

Patterns and hierarchical models

Patterns are defined as a collection of model signals combined to describe a subset of the overall spectrum, e.g. a metabolite. They can originate either from a measured spectrum, cast into the form of a numerical vector ($Num(t_n)$), or from a simulated spectrum, filtered into the form of a small set of Lorentzian or Voigt lines. FiTAID allows for different pattern types, some of which are of a hierarchical nature [28], i.e. composed of parent and children, where the parent parameters describe overall properties (e.g. concentration of a metabolite) and the children's parameters relate to differences between them (e.g. allowing for different relaxation properties for different protons within a metabolite). For a parent p and C children, the hierarchical model $\xi^{hier}(t_n)$ has the form.

$$\xi^{hier}(t_n) = \xi_p(t_n) \sum_{c=1}^{C} \xi_c(t_n) \tag{11}$$

Expanding the expression results in

$$\xi^{hier} = \sum_{c=1}^{C} A_p \cdot A_c \cdot \exp(-i(\omega_p+\omega_c)t) \cdot \exp\left(-(\alpha_p+\alpha_c)t \right.$$
$$\left. -(\beta_p^2 + \beta_c^2)t^2\right) \cdot \exp(i(\phi_p + \phi_c)) \tag{12}$$

which is equivalent to a model with C independent Voigt lines, but allows for accommodating differences between overall parameters of the children, while nicely keeping an intuitive model structure.

The different patterns are:
Simple Line:

$$\xi^{Line}(t_n) = \xi^V(t_n) \tag{13}$$

Parametric pattern:

$$\xi^{Para}(t_n) = \xi^V(t_n) \sum_{d=1}^{D} \xi_d^{Line}(t_n) \tag{14}$$

d is the line number and D is the number of lines in this parametric pattern.
Numeric pattern:

$$\xi^{Num}(t_n) = \xi^V(t_n) \cdot Num(t_n) \tag{15}$$

Metabolite pattern:

$$\xi^{Met}(t_n) = \xi^V(t_n) \cdot \left(\sum_{e=1}^{E} \xi_e^{Line}(t_n) + \sum_{f=1}^{F} \xi_f^{Para}(t_n) \right.$$
$$\left. + \sum_{g=1}^{G} \xi_g^{Num}(t_n) \right) \tag{16}$$

A structure composed of arbitrary numbers of lines, parametric, and numeric patterns with e the running line number and E the total number of simple lines in this metabolite pattern. f and F, as well as g and G, are the respective labels for the parametric and numeric patterns. An example for using a metabolite pattern with mixed pattern types is NAA where the CH_3 signal can be modeled with a simple line, the amide signals as a parametric pattern with 2 lines and—because of its complexity—the aspartate signals as numerical pattern.

Note that the indices should fulfill the conditions $D \geq 2$ and $E + F + G \geq 2^2$ in order to justify using these patterns, otherwise, they can be replaced by simpler forms. Furthermore, care has to be taken not to create an overdetermined system with interdependent fit parameters, if for example a doublet of lines is described by a parent with two children, only two of the three possible parameters of each type should be used as free fitting parameter, the remaining ones being fixed or locked by prior knowledge restraints. The concept of these patterns and hierarchical models is illustrated for lactate in the following illustrations section.

With the use of patterns, equation [2] can be re-expressed as a sum of H lines, I parametric patterns, J numeric patterns, and K metabolite patterns:

$$\xi(t_n) = \sum_{h=1}^{H} \xi_h^{Line}(t_n) + \sum_{i=1}^{I} \xi_i^{Para}(t_n)$$
$$+ \sum_{j=1}^{J} \xi_j^{Num}(t_n) + \sum_{k=1}^{K} \xi_k^{Met}(t_n) \tag{17}$$

Note that H, I, J, K can have zero value[2], but their sum must be greater than 1.

2D model

In the two-dimensional case, the model is an array of 1D models, the data an array of 1D datasets:

$$\xi = \begin{bmatrix} \xi^1 \\ \vdots \\ \xi^L \end{bmatrix}; \quad \mathbf{x} = \begin{bmatrix} x^1 \\ \vdots \\ x^L \end{bmatrix} \tag{18}$$

---

[2] Not all entity types have to be present. For missing entities the corresponding maximum index is symbolically set to 0 and the term eliminated from the summation.

and χ^2 becomes a sum over the χ^2s of the rows, as defined in Eqs. (3) and (5):

$$\chi^2 = \sum_{l=1}^{L} \chi^{l^2} \tag{19}$$

where l is the spectrum number and L is the number of spectra in the 2D array.

Prior knowledge and relations in 2D

2D modeling is not very useful if the model parameters of all rows are independent of each other. On the contrary, the 2D model normally only introduces a few new parameters compared to the 1D model, with most parameters between rows being linked by a defined relationship. The differences between rows are mostly due to changed experimental conditions, with e.g. differing individual echo time (TE), repetition time (TR), saturation time (TS), inversion time (TI), or mixing time (TM). Each row will therefore have its own base models for linear combination modeling, e.g. in a 2DJ data array each row will access base spectra simulated with the appropriate TE. In addition, the 2D relationship can include an additional amplitude factor f, dependent on the type of 2D experiment. For row l in the 2D model, the mth 1D model entity $\xi_m(t_n)$ from Eq. (2) transforms into $\xi_m^l(t_n)$, and spelled out for simple Voigt lines:

$$\xi^{V^l}_m(t_n) = f_m^l \cdot A_m^l \cdot \exp(-i\omega_m^l t_n)$$
$$\cdot \exp(-\alpha_m^l t_n - \beta_m^{l^2} t_n^2) \cdot \exp(i\phi_m^l) \tag{20}$$

where the factor f_m^l describes the second dimensional amplitude relationship.

The following idealized 2D experiments have been implemented:

Saturation Recovery (SR):

$$f_m^l = (1 - \exp(-TS^l \cdot r_{1m})) \tag{21}$$

Inversion Recovery (IR):

$$f_m^l = 1 - \left[2 - \exp(-Tr^l \cdot r_{1m}) \right] \cdot \exp\left(-TI^l \cdot r_{1m} \right) \tag{22}$$

2DJ:

$$f_m^l = \exp(-TE^l \cdot \alpha_m) \tag{23}$$

Constant relationship:

$$f_m^l = 1 \tag{24}$$

where Tr is the signal recovery time (before the inversion pulse), and longitudinal and transverse relaxivities are defined as

$$r_1 = 1/T_1 \tag{25}$$

Fig. 1 Illustration of signal patterns and hierarchical models to demonstrate possible pattern combinations for modeling the lactate spectrum. The lactate signal has two groups of peaks, one for the CH$_3$ protons (doublet) and one for the CH proton (quartet); due to strong coupling all peaks with their own phase and intensity. The signals can be represented as a combination of basic (Voigt) lines (*black square*), numerical patterns (*asterisk*), parametric patterns (*lozenge*), and metabolite patterns (*dagger*). (i) consists of one numerical pattern, described by the basic numerical vector and 5 parameters (i.e. area, frequency, phase, Gaussian width, Lorentzian width). (ii) consists of two numerical patterns (potentially 10 parameters). (iii) introduces a metabolite pattern as parent level and case (ii) as child level (15 parameters overall). (iv) includes 2 separate parametric patterns (40 parameters), while (vii) makes use of 3 hierarchical levels with a metabolite pattern as parent and case (iv) as descendants (total of 45 parameters). (viii) represents the plain description as 6 simple lines (30 parameters). Many further possibilities exist (e.g. (v), (vi)). Models (iii–vii) are of a hierarchical structure while (i–ii, viii–ix) are flat. For meaningful fitting, the maximal number of parameters is reduced to the minimum needed for each case by using prior knowledge relations within each hierarchical level

and

$$\alpha = r_2 = 1/T_2 \tag{26}$$

In addition to the prior knowledge included in the amplitude factor f_m^l, all prior knowledge relations P_m for any of the standard 1D parameters describing an entity and valid for entity m in the first trace are valid for all traces, independent of the row number l.

$$P_m^l = P_m^1, \, P \in \{A, \omega, \alpha, \beta, \phi, r_1\} \tag{27}$$

Fit strategy steps

In most fitting programs, the fit is done in a single automatic fit with selectable model, data, algorithm, and potentially pre-incorporated preprocessing or subsequent optimization of parametric and non-parametric parts of the fitting problem—all steps without options for a user strategy. FiTAID offers fit strategy steps that allow to consecutively improve the fit following a user-defined path in which at each step the result of the previous fitting step is used as initial parameter state for the current step while different pre-defined fitting conditions can be applied (see last part of following illustra-

tions for examples). The aim of using fit strategy steps is to avoid local χ^2 minima and to speed up convergence. (For details on this option and other implementation issues, see Appendix B).

Illustrations of signal models, prior knowledge constraints, and fit strategies

To demonstrate the described signal models, prior knowledge implementations and fit strategies, as well as general features of 2D modeling of interrelated MRS datasets, selected examples are presented in the following. Modeling of 1D spectra—a trivial case of 2D fitting—can be accomplished just as with TDFDFIT [2] or LCModel [1] but is not illustrated for brevity.

Hierarchical model

Figure 1 shows the different model definitions possible for a single metabolite. Lactate was chosen for this example because of its clear division into 2 signal groups, one of them potentially affected by water suppression and pH
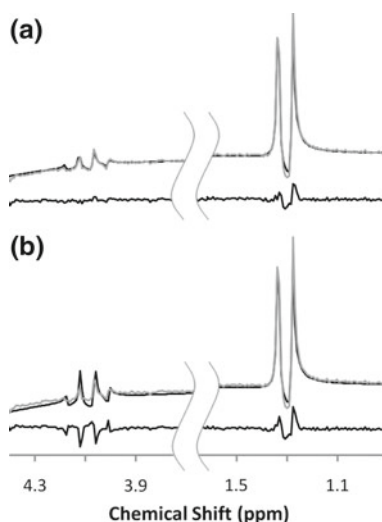
Fig. 2 Effect of split basis spectra in a hierarchical model allowing for deviations between basis and in vivo spectra. The measured signal from a lactate phantom (TE 20 ms) was fitted with a metabolite pattern with a two-part model for the quartet and doublet parts of lactate (**a**) and with a single numerical pattern for lactate (**b**). In case (**b**), the residues show a large difference for the quartet signal because of a difference between simulated and measured spectra due to the effect of the water suppression pulses. The residuals for the doublet are probably mainly due to chemical shift offset effects, not considered in the spectral simulation

changes. Using the patterns described above, nine different combinations are shown. Depending on the situation, different combinations may be useful. For example, (i) a single numerical pattern is the simplest model with fixed prior knowledge between the doublet and quartet (least free parameters). (ii) allows for variation between them by splitting them into two numerical patterns. (iii) is the same as (ii) but with a metabolite pattern as parent and the main fitting parameters would be on this level, while small parameter variations would occur on child level (e.g. a water suppression effect on the quartet amplitude, see Fig. 2). However, the total number of free parameters must be the same as in (ii) to prevent an under-determined condition with no true solution. (v–vii) show the deepest hierarchical structure as currently implemented.

To demonstrate the benefit of having hierarchical base patterns available, a lactate phantom signal was measured with PRESS at TE 20 ms and fitted using the model versions (i) and (iii) from Fig. 1. Figure 2b shows that there are differences between the idealized simulation and the actually measured spectrum. Model (iii) allows for these differences with an extra fitting parameter (area ratio) for one of the child patterns, hence improving the fit considerably, as shown in Fig. 2a. The prime cause for these differences is the fact that the quartet signal near water is affected by the water suppression pulses in the measurement. Since this effect will be identical in vivo, the optimized model (iii) can be used as basis spectrum for brain measurements. There could also be pH or temperature differences between the simulation and measurement conditions that might affect the chemical shift differences [29], which could be taken into account in a similar

manner. The remaining residues for the doublet with model (iii) are caused by imperfect simulation of the lactate spectrum that did not take chemical shift offset effects of the RF pulses into account [30]. Further relaxation of the originally simulated lactate model—as feasible with e.g. pattern (v)—would be needed to compensate for this.

2D linear combination model fitting: saturation recovery

The advantage of fitting interrelated spectra simultaneously instead of sequentially can be demonstrated with a set of saturation recovery spectra, reflecting longitudinal relaxation. Simulated data (Fig. 3a) is used for illustrative purposes and to verify that FiTAID converges to the true values. The overlapping spectra of glutamate, glutamine, and GABA were assumed to feature clearly distinct T_1s. Three different methods to fit this data set are investigated: (1) Fitting each spectrum separately. (2) Modeling all the spectra together as an array with prior knowledge between the spectra to ensure they have identical chemical shifts and line widths. (3) Introducing in addition the known relationship for the amplitudes along the delay time dimension. Figure 3c demonstrates the deviations of the fitted from the true peak areas in the dataset. Deviations mainly occur at short recovery times, i.e. low SNR. It is obvious that the fit works best for the 3rd method when all prior knowledge is enforced (and is correct). The main reason for the differences in this example is the restriction to a common chemical shift, which is well defined in the spectra with large SNR. Of course, if the amplitude relationship is introduced, no outliers are possible for any single amplitude; hence the resulting fit is close to ideal. An
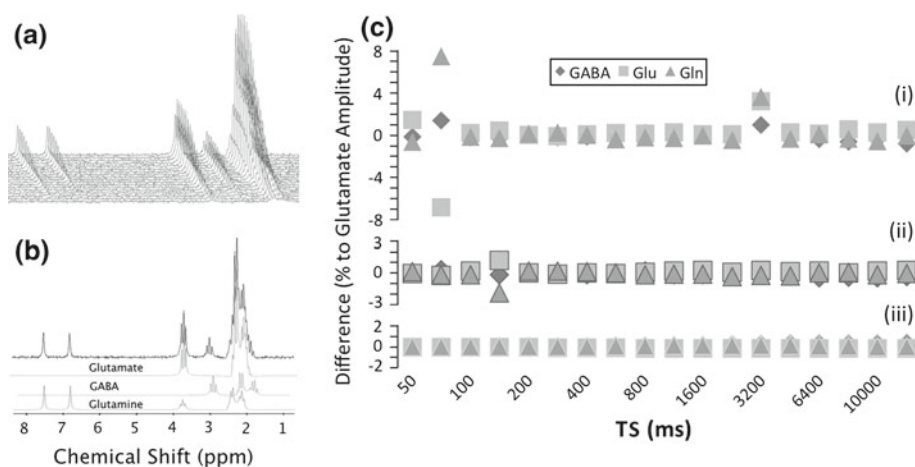
Fig. 3 Example of 2D fitting for simulated saturation recovery data. Simulated saturation recovery spectra (**a**) with distinctly different T_1s for glutamine, glutamate, and GABA and the respective basis spectra (**b**) are shown. In (**c**), the difference between expected and fitted values as ratio to the glutamate time domain amplitude are displayed for three different fit models: (i) independent fit for each spectrum, (ii) 2D fit of the spectral array with prior knowledge (see Text) but no amplitude relation along TS; (iii) 2D fit of the spectral array with prior knowledge and known relationship. (Model with relative amplitudes of 1, 3.5, and 1.2 for GABA, glutamate, and glutamine, respectively, and with T_1s of 1,200, 1,000, 2,500, and 2,000 ms for GABA, glutamate, glutamine (upfield), and glutamine(downfield), respectively. The noise was set with a standard deviation in frequency domain of 1/10 of the downfield glutamine singlet intensities

advantage of method 3 versus method 2 would have to be evaluated based on fitting errors for the T_1s. A certain benefit for maximum prior knowledge along the second dimension is the reduced number of fitting parameters, which decreases the computation time significantly.

2D linear combination model fitting: In vivo time series

Time series of spectra are necessary to study kinetic processes. The recorded data are related and hence can be combined into an array for fitting [8]. Figure 4 shows a series of 24 downfield ^{1}H brain spectra obtained over 12 hours after an oral histidine load [31]. Histidine features 2 singlet peaks in the downfield region [32] with less overlap with background signals than for the upfield portion. It can be assumed that only the histidine contribution to the spectrum is varying, while the baseline spectrum remains unchanged. The fitting model was set up to include two Voigt lines for the two histidine peaks, a doublet to represent NAA at 7.8 ppm, and eight broad Voigt lines for the rest of the downfield baseline evident in the pre-load spectrum. Two lines were used to model residual water and three broad peaks and a sharp NAA (2.0 ppm) peak to coarsely represent the main features of the upfield spectrum. These lines are clearly insufficient to model the upfield spectrum as such, but were judged sufficient to represent the potential baseline contributions to the downfield spectrum fitted in a frequency-selective FD fit. Using an approximate model outside the range of χ^2 minimization reduces the model complexity considerably and was used here in combination with frequency-selective fitting and

appropriate parameter restrictions. The model enforced common frequency and Lorentzian line width as prior knowledge because T_2s did not change, and frequency offsets were corrected using eddy current correction with the reference water spectrum for each spectrum. A variable shared Gaussian width per spectrum allowed for small differences in shim efficiency in different measurement sessions. Simultaneous fitting of all spectra allowed for concurrent determination of the baseline signals and the amplitude dynamics of the histidine peaks.

Fitting was done in three fit strategy steps: (1) in TD for the large signals of water only (sequential 1D fits). (2) 2D fit in TD, restricting the upfield part in terms of common area, width, and offset parameters for all spectra (except for individual lipid and water contributions). (3) Finally in FD, restricted to the range of the downfield metabolites (5.8 to 9.2 ppm). This allowed for a high precision fit for the small variations of the histidine contributions in spite of fairly large signals of water and the upfield region. Evidence for the success of this fitting strategy comes from the fact that the residues (shown for one subject in Fig. 4)—also if summed over all spectra (not shown)—appear mostly random, from the smooth increase of the histidine signals over time, and from the variance measures presented next.

Time and frequency domain fit

The relative performance of the TD and FD fit in the previous example of histidine kinetics is evaluated in two ways. First, the reproducibility of the determined histidine concen-
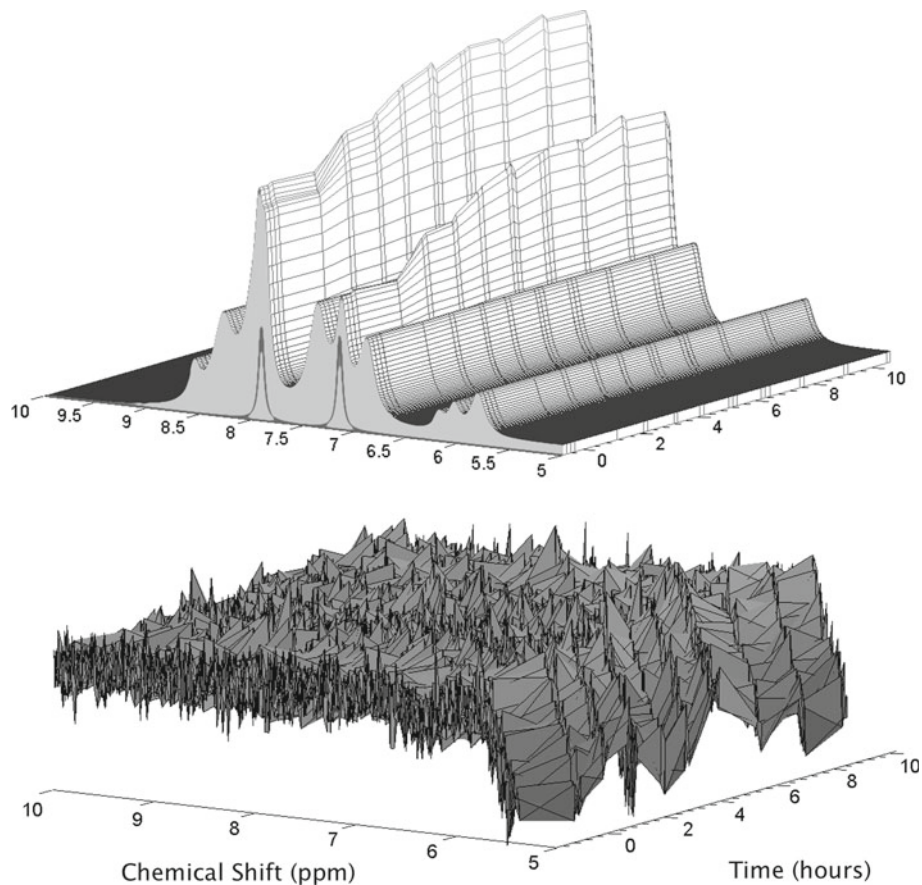
Fig. 4 Illustration for 2D fitting of a time series of spectra. The plot contains fitted spectra and fit residues obtained in a single healthy subject for a series of spectra recorded after an oral load with histidine at time 0. The model consists of two resonances with varying amplitude from histidine (shown at 7.06 and 7.79 ppm), the 7.82 ppm doublet from the amide proton of NAA, and heuristically placed lines to describe the rest of the downfield spectrum. There are a total of 24 spectra recorded at intervals indicated by the *vertical lines* and fitted after scaling with unsuppressed water. The spectra, fitted with individually varying Gaussian widths, were plotted with identical Gaussian width for visual clarity. Fit and residues were drawn on the same scale reminding that downfield spectra have limited signal to noise, a typical case where there is intrinsically a large potential benefit for combined 2D fitting

trations was expressed as the coefficient of variance (CV) for the values found in single measurement sessions (i.e. values obtained without repositioning the subject and over a time interval without large true concentration variations [i.e. ~15 min]). On average over all 10 sessions and all 8 volunteers, the TD fit yielded a CV of 32%, while the FD fit was much more stable with a CV of 17%.

Second, the obtained values and their variations are displayed as a correlation plot in Fig. 5. Two features are striking. (A) The TD values show an offset, i.e. a large histidine content before the load, where we know [His] to be exceedingly small. (B) The standard deviation (inter-subject variation) for each time point is consistently larger for TD fitting. Both effects are probably due to a baseline contribution from residual water where the incomplete model for water is used to fit the main features of water at 4.7 ppm, not the baseline for the minor downfield signals. However, the trend line for TD versus FD values is linear and close to parallel to
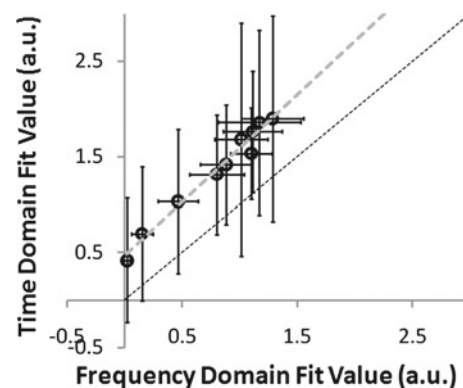
Fig. 5 Correlation between fitted histidine levels (arbitrary unit, a.u., obtained by scaling with the unsuppressed water signal) as found in TD and FD fitting. Mean ± one standard deviation *error bars* for a cohort of 8 volunteers' data recorded at 10 different time points after histidine load. The TD fit yielded a consistent offset compared to the FD fit, suggesting that the histidine level was overestimated. Also, the inter-subject variation at each time point is larger for the TD fit

the diagonal, suggesting that both methods yielded similar concentration dynamics.

In the present case, a simplistic model was used for most of the spectrum instead of a complete model; hence, the distinct performance differences. Of course there would be other solutions for a more robust TD fit, in particular pre-filtering (Refs in [4]). Completing the model would probably not work in this case, because it would be very complex for the upfield part and no true model exists for residual water.

Fit-strategies

It might be argued that fit strategy steps are superfluous and that a fitting algorithm should be able to converge to the global χ^2 minimum in a single step allowing for adjustment of all free variables simultaneously. As it turns out, localized MR spectra, and in particular 2D arrays of spectra, involve a multitude of mutually dependent fitting variables, such that finding the overall χ^2 minimum is a daunting task, where human strategy implemented into the strategy steps in FiT-AID may be superior to brute force optimization (particularly if using a least squares technique that necessitates good starting values).

Sequential fit: in 2D arrays of spectra, the residual water can be the dominating signal in many traces. It can be of a complex and ill-defined form, and without any prior knowledge relationship between the spectra. Hence, it adds a huge number of free fit parameters to the problem. In most of the presented 2D cases, it turned out to be beneficial to include a first fitting step in which to adjust the model for the residual water only in a sequential spectrum-by-spectrum fashion, followed by actual model optimization with and without further fitting of the water lines.

Partial data selection: in some 2D cases it may be beneficial to start fitting on selected traces of the full matrix, which may reduce computational burden tremendously. For IR spectra, one can optimize the metabolite signals (widths, shifts, relative amplitudes) primarily using the non-inverted spectrum and have this information cascade through the model to all other traces due to the second dimensional prior knowledge. Subsequently, their T_1 can be roughly estimated when only using the non-inverted plus one metabolite-nulled signal. A final fitting step including all data and all free parameters will then start with near optimal starting values leading to faster convergence.

Fit strategy steps: Besides switching between χ^2 calculation in TD and FD as discussed before, fit strategy steps also allow to optimize some types of parameters only, while keeping the others fixed for the moment. For example, it has been found useful to keep relaxation parameters at the starting values (found in earlier similar cases) initially while optimizing the other parameters (amplitudes, frequencies, widths) and only later include all free parameters into the fit. Further-
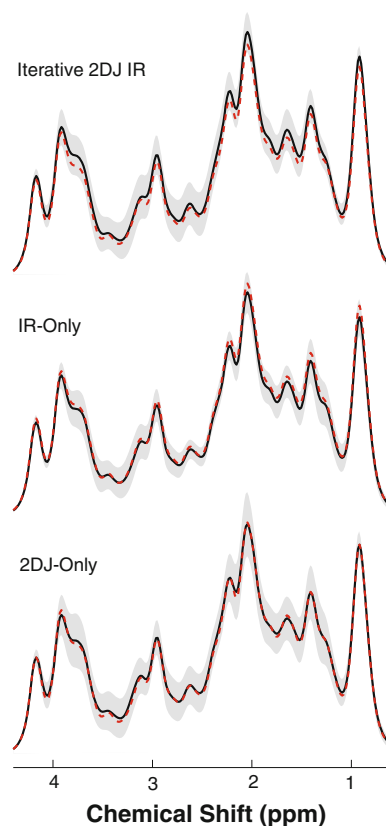

Fig. 6 Fit results for the synthetic data as specified in Methods and with similar spectral appearance as the in vivo case in Fig. 7. Macro-molecule baseline fits for the three methods, 2DJ-only, IR-only, and iterative 2DJ-IR. The *dashed line* corresponds to the noise-less true spectrum, while the full *black line* represents the mean fitting result enveloped by a *light gray* band of ±1 standard deviation over different noise realizations

more, fit strategy steps allow fitting of different data channels. It is reasonable to optimize frequency shifts initially using the magnitude data channel (and keep phase variables untouched) and let the signal phases and widths only be fitted subsequently using the complex data.

Results

Since there is no gold standard to obtain the true MMBL in vivo and hence to judge the accuracy of the fits, the performance of the three methods was tested on simulated data with different noise levels and noise realizations. Figure 6 visualizes for one noise level (similar to the SNR in Fig. 7) and a set of different noise realizations that all three methods return results close to the true values and with similar SD. Quantitatively, the 2DJ-only method shows higher variation than the other two methods (mean variation of 0.6% (2DJ-only) versus 0.1% (IR-only) and 0.2% (iterative), respectively. The Levene test, indeed, confirms a difference of variance between methods.

Fig. 7 Representative experimental 2DJ and IR in vivo spectra and corresponding fit and residues. For clarity, the data are only shown in part. For 2DJ, only TE = 20, 32.5, 45, 57.5, and 307.5 ms spectra are shown, while for IR spectra with TI = 30, 450, 575, 700, 825, and 1,200 ms are presented
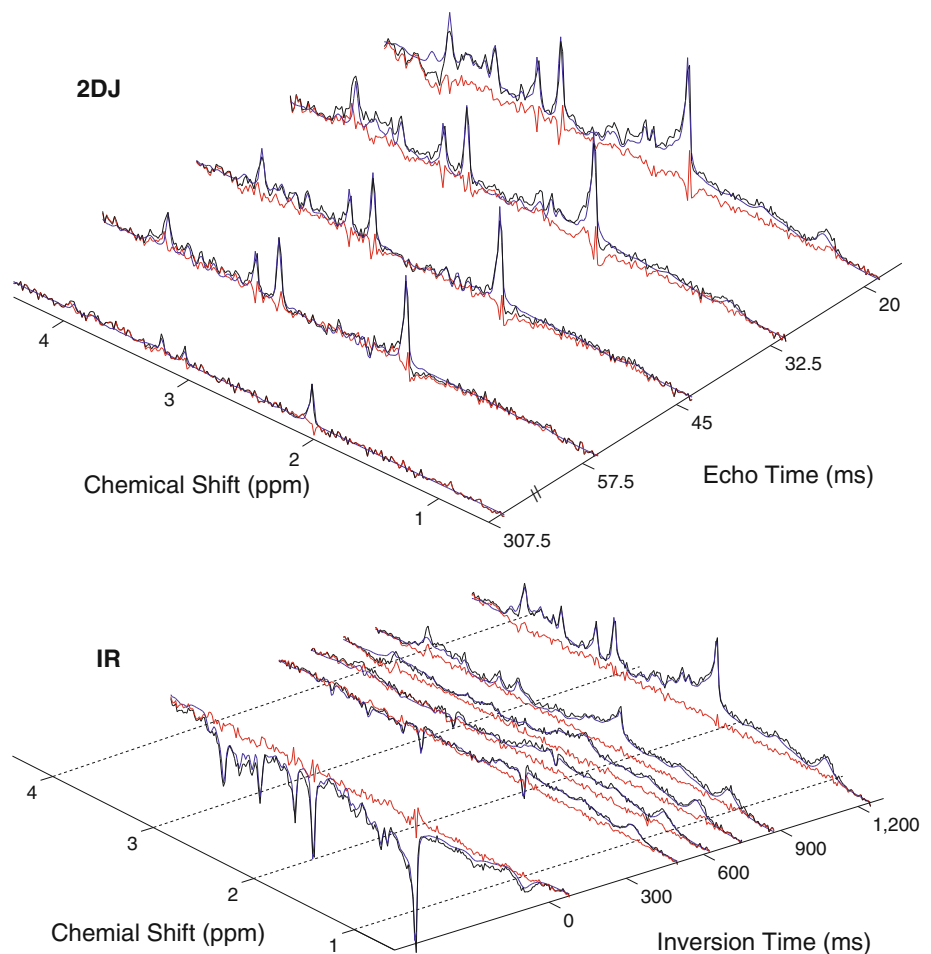
The mean absolute deviation from the true values over the whole spectrum was 2.6, 1.3, and 1.6% of the 0.9 ppm peak amplitude for the 2DJ-only, the IR-only, and the iterative method, respectively. The Wilcoxon test indicates that some of the small deviations from the true means are indeed systematic—worst for the iterative method, where 35% of all data points are systematically deviating from true values— though most of them outside the spectral range of interest (i.e. < 0.7 ppm and > 4.1 ppm where there is intrinsically very little signal) and only 10% within that range. For IR-only 14% and for 2DJ-only 4% of the full spectral range show a systematic deviation from the true value. The differences between the methods are reproduced also at higher and lower noise levels.

Connected to the variations in determined MMBL are variations in the obtained metabolite contents. It was found that for this particular setup the variations in the metabolite contents are comparable for all three methods of MMBL determination. For all three methods, the metabolite contents vary between <1% and ∼ 4% of the creatine concentration.

The quality of the experimental data used to determine the MMBL is illustrated in Fig. 7 where the experimental IR and 2DJ spectra for a single subject are portrayed together with respective fits and residues.

The average resulting MMBLs and their inter- and intra-individual variations as obtained from the three techniques are shown in Fig. 8. The top row shows the average over 12 subjects and the gray area represents the range of the mean ±1 SD of inter-individual variation, while the lower two rows include results for intra-individual repetitions in two subjects, where the gray area refers to intra-individual variation. As evident from the figure, the results differ strongly between the methods; in particular, the mean baseline obtained from 2DJ-only differs over most part of the spectrum from what is obtained from the other two methods. This is confirmed statistically with ANOVA yielding significant differences between methods for 88% of the spectral range and the post hoc analysis identifying the 2DJ-only method as the reason for the significant difference for the average of all subjects and similar results for the intra-subject repetitions. The graphs also indicate that the variance between subjects and for the repetitions is larger for the 2D-J only method. Quantitatively, the mean variance over the whole spectrum as percentage of the amplitude of the 0.9 ppm baseline

Fig. 8 Fit results for the human in vivo data. Macromolecule baseline fits for the three methods, 2DJ-only, IR-only, and iterative 2DJ-IR. Three sets of data are plotted representing the cohort of 12 healthy subjects (*top row*) and 5 repeated examinations in two subjects. The *black line* represents the mean result enveloped by a *light gray* band of ±1 standard deviation over subjects (*top*) or repetitions
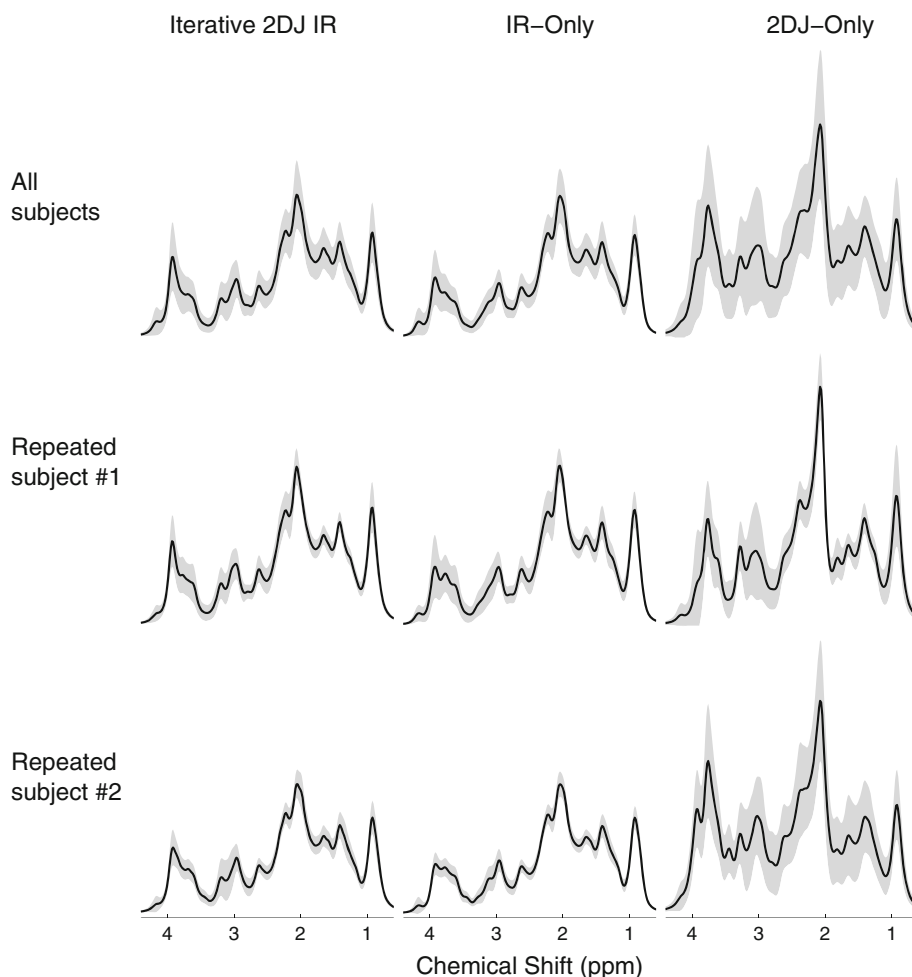
peak is indeed about equal for both methods that include T_1 information (0.7% for the inter-subject comparison; 0.1 and 0.4% for the intra-subject comparisons) and much larger for 2DJ-only (1.6% inter-subject, 1.5 and 1.0% intra-subject, respectively). The Levene test confirms statistical differences between methods for the standard deviations for most of the spectral range in the inter-subject data, and in one of the two intra-subject repetitions. The difference in variance for inter- and intra-subject data suggests that there are differences in the MMBL for different subjects. This was tested by ANOVA. It showed that for 29% of the spectral points there were significant differences between subjects, mostly in the spectral range between 1.0 and 2.5 ppm.

Discussion

Novel schemes for the determination of the macromolecular baseline in localized brain spectra are presented, which rely on simultaneous fitting of series of related datasets. Using FiTAID, a general Fitting Tool for Arrays of Interrelated

Datasets, it is possible to define the MMBL either based on differences in T_1 or T_2, or both characteristics. We have presented both a description of the general fitting tool and the results of applying this tool for the particular aim of baseline characterization in localized brain spectra obtained from healthy human subjects at 3 T.

Macromolecular baseline estimation methods

The three methods of determining the MMBL were compared for both in vivo and synthetic data. In vivo, the methods were judged based on reproducibility measures for intra- and inter-subject variance in the resulting baseline. Since there is no gold standard for the baseline, accuracy in terms of deviation of fit from true values was investigated for synthetic data closely derived from the in vivo spectra. Even though it was found that there are clear differences in performance of the three methods, the prime target of this contribution was not to arrive at an optimal method for baseline determination because much more work can and will be put into optimizing all three methods. Rather, this report aimed at

introducing the concept and describing the tools for combined multi-parametric determination of the baseline.

The most used experimental technique for MMBL definition is metabolite nulling based on zeroing metabolite signals by inversion recovery. This is problematic, given that T_1's are different for different metabolites or proton groups. Hence, there will always be residual contributions from metabolites in a thus determined MMBL. Methods to approximately correct for this in post-processing have been presented [33]. In addition, diffusion weighting has been proposed as an experimental alternative for baseline determination [34]. This is currently only useful for animal scanners with very high gradient power. Pure post-processing methods for baseline elimination rely on differences in spectral shape and/or fast T_2^* decay and are prone to systematic effects [33,35] on metabolite estimation and/or lead to a loss in SNR.

Similar to a method based on simultaneous evaluation of saturation recovery data [9], all currently presented methods rely on simultaneous evaluation of a series of spectra. (1) IR-only uses an extended set of short-TE IR and non-inverted spectra, thus relies on differences in T_1 between metabolites and baseline, but does not depend on metabolite T_1s to be equal. (2) 2DJ-only is based on the fast TE decay of the MMBL to distinguish it from the metabolites, which are described as a linear combination of basis 2DJ spectra obtained from simulation that decay with co-determined metabolite T_2s. (3) 2DJ-IR, finally, is designed to make use of both, the distinction based on T_1 and T_2, in an iterative manner.

Thus, the third method should in principle provide the best means to determine the true MMBL. The actual results did not confirm this for the currently tested version of this method. It rather seemed that using IR-only gave slightly more stable and trustworthy results, while using the TE series only to determine the MMBL yielded the largest variance in the results for both the intra-subject repetitions and the synthetic data. In addition, it yielded systematically larger baselines in the in vivo case than the IR-based methods. There are several reasons why the 2DJ-only method performed least favorably. The main cause is probably the fact that the TE increment was directed more at defining metabolite signals and their T_2 than at an efficient determination of the baseline. Out of the 24 TE traces, only those three with shortest TE had substantial amplitudes for the MMBL, and the large TE increments did not provide a good basis to extrapolate back to TE 0 for the MMBL. The baseline would probably be much better determined with less, but more densely sampled TE increments. Another reason could be that some of the contributing MM resonances are subject to J-couplings and hence do not adhere to the simple mono-exponential decay described by the model. A more sophisticated model of the true baseline than a set of overlapping and exponentially decaying Voigt lines might be called for. But also the

IR-only method can be improved. Instead of using a full IR picture with the lengthy procedure of recording seven IR times, it may be more efficient to define the MMBL based on just one or two IR times around the metabolite nulling time plus a non-inverted spectrum. Enforcing the large difference in T_1 between metabolites and MMBL in the fit, small residual positive or negative metabolite peaks in the quasi-nulled spectra would be unproblematic, since they would be incorporated in the model. The information from the standard metabolite spectrum, recorded with better SNR than in the current approach, would enforce proper chemical shift and width information for those residual peaks in the IR spectra. Clearly, the combined 2DJ-IR method can therefore also be improved by using optimized 2DJ and IR elements, but also by concurrent instead of iterative modeling. This would also be less labor-intensive since currently the switching between models is done manually. For a concurrent approach, it will also be much easier to pre-define the best combination of acquisition parameters in the IR and 2DJ series to achieve the smallest errors for the area parameters of the MMBL.

The accuracy of the fitting methods has been evaluated with synthetic data. However, the presently used synthetic data has the disadvantage that it conforms exactly with the models used for the fit, hence there exists a true solution for all methods, whereas in reality the baseline is nonparametric and highly complex. The results from the fit of the synthetic data should therefore not be over-interpreted. At first, it seems surprising that the 2DJ technique almost works as well as the iterative and IR-only techniques, although most traces do not show much baseline intensity. However, this is not unreasonable since the model exactly represents the reality of the simulated spectra. Stability in estimating the in vivo data may be a better criterion—though with the caveat that the results for baseline and metabolites might be systematically wrong.

Individual macromolecular baselines

An important finding, independent of the methodology, is that the variance for the MMBL is clearly much larger in the inter-subject comparison than for the within subject data. This indicates that the MMBL cannot be assumed to be constant between subjects, even in health. The increased variance may just reflect small differences in voxel composition or true differences in background metabolite content. Inclusion of a single MMBL spectrum into standard 1D linear combination model fitting may therefore introduce some bias.

Variation in co-determined metabolite content

The main emphasis has presently been laid on the determination of the MMBL for use in subsequent separate anal-

ysis of metabolite contents, though of course the presented methodology can be used for simultaneous evaluations. A fair comparison of the proposed methods in terms of achievable accuracy and variability of the determined metabolite content is therefore beyond the scope of this work. For the simulated data, where the methods can at least be compared in terms of measurement time, it was found that there was no consistent overall trend for one method to be better in terms of associated metabolite reproducibility. This is promising for the combined method, given that it yields additional information on both longitudinal and transverse relaxivities.

FiTAID

The fitting tool FiTAID, which has been used in this evaluation and which is also presented in this report, is one of the first fitting tools for evaluation of multidimensional in vivo MRS datasets. It has a very general scope and provides multiple features that are not available in this manner in other evaluation packages. This includes the possibilities to fit multiple one- and two-dimensional experiments, as presented above, the use of parameter restrictions, the implementation of prior knowledge in both dimensions, the option of fit strategies, the elective use of fitting in time or frequency domain, and hierarchical models.

In contrast to traditional (often non-quantitative) evaluations of 2D NMR spectra or the Matlab package by Schulte et al. [10], FiTAID is based on a 2D signal model in the measurement domain, i.e. without Fourier transformation, and on χ^2 evaluation in either time/time or frequency/time domain. χ^2 calculation after 2D Fourier transformation would allow for selective FD fitting in both dimensions but has the disadvantage of either being slow or of only approximate modeling. Whether or not this is an advantage remains open at this time, probably the same arguments hold as for FD versus TD in 1D fitting. If the model is correct and complete the two approaches are equivalent. The modeling is certainly less complex and more direct in terms of T_2 determination without the second Fourier transformation. An additional benefit of FiTAID is the possibility to combine irregularly spaced TE data because no FT is required along TE. This is useful to optimize TE combinations for largest metabolite to background contrast.

The time required for the fit of a spectral array depends on many factors, including the minimization algorithm used, the number of model spectra, the number of data points, as well as the number of free fitting parameters. In our experience, calculation times in FiTAID on standard PC's may vary from seconds to many hours depending on the application. Vanhamme et al. [8] have nicely shown that simultaneous fitting can be very costly in terms of CPU time.

In general, concurrent fitting of arrays of spectra will improve overall fit performance (i.e. improved accuracy of fit variables, i.e. better use of scan time) only if the spectra are interrelated in a way that can be included as prior knowledge relations between the traces, otherwise separate modeling yields equivalent results and is much faster. Seen in this light, 2D fitting is no other trick than using another form of prior knowledge.

One might argue that 2D fitting is most often equivalent to fitting in a two-step procedure, where all available spectra would be summed and fitted to determine chemical shifts and widths in a first round, followed by sequential fitting of individual traces using the prior knowledge from the summed fit. This is true for specific cases, but not in general. For example, a TE-averaged spectrum [36] allows for easy glutamate determination but other metabolites may be averaged away and hardly contribute to the sum. Gonenc et al. [12] have shown that fitting 2D spectra in an array yield lower standard deviations for a 2DJ experiment compared to the TE-averaging technique. Similarly, the summed signal will often be broadened because of motion or other instabilities—effects that can be eliminated in 2D fitting.

Conclusion

We have described three methods for the determination the MMBL of short-TE brain spectra that depend on multidimensional fitting of interrelated datasets. In the current implementation, it appeared that fitting of IR data is just as reliable as a more complex approach integrating T_2 and T_1 differentiation for MMBL determination. However, all three techniques can be further optimized such that indeed the combined method is expected to be more useful eventually. We also presented a versatile and extendable tool for the simultaneous evaluation of arrays of interrelated datasets. Because sensitivity restrictions almost always force clinical MRS examinations to consist of multiple repetitive acquisitions, it appears that it will often be beneficial to use the multiple scans for a more advanced use than for a mere average. Artifact detection [37] or error determination by bootstrapping procedures [38] may be two direction, the combination of multiply encoded information in simultaneous evaluations is another that may either give additional information (e.g. relaxation or diffusion) or may allow to lower intrinsic error bounds by combining data providing less correlation between the metabolite spectra of interest [39]. The main example emphasized the possibility for better distinction between MMBL and metabolites based on relaxation properties, but collections of multiple TE or TM spectra will certainly provide the potential for targeted metabolites to be better distinguishable from background and other metabolites—similar

to editing scans, but without sacrificing the overall information.

Appendix A: Fitting conditions for macromolecule estimation

Model

2DJ: In FiTAID, the experimental multi-TE array of spectra is modeled in a way as to optimize linear combinations of 2DJ basis spectra of known metabolites (see Eqs. 19, 20 and 23). The basis spectra were simulated for ideal hard RF pulses with fixed TE1 of 4 ms and varying TE2 from 16 ms to 303.5 ms in steps of 12.5 ms to imitate the measurement condition.

IR: The IR data were modeled according to Eqs. (19), (20) and (22) using identical base spectra simulated for TE 20 ms.

Metabolites included in the models were GPC, phosphorylcholine, aspartate, creatine, ethanolamine, GABA, glucose, glutamate, glutamine, glutathione, glycine, myoinositol, NAA, NAAG, scyllo-inositol, serine, and taurine. Metabolite basis sets were calculated using GAVA [39]. Because of eddy current correction they were assumed to have correct phase for all TEs and the signal phase was not fitted. All metabolite base spectra were fit with a common frequency shift and a common Gaussian width. Metabolite T_1's and T_2's: peaks were divided up into 5 groups with separate relaxation times. (1) acetyl groups of NAA and NAAG, (2) trimethyl groups of GPC and phosphorylcholine, (3) methyl group of creatine, (4) methylene group of creatine, (5) all other protons. T_1's were restricted to values >0.1 s, T_2's >16 ms. Macromolecules were modeled as a set of Voigt lines with freely adjustable individual amplitudes equal adjustable Gaussian width at fixed 10 Hz separation placed between 0.6 and 4.2 ppm. The Gaussian width was allowed to vary from 5 to 30 Hz in order to accommodate smoothing of the MMBL spectrum and to compensate the specific shim setting. To allow for inherent differences between terminal methyl groups and other protons in the MMBL, the baseline was split into two parts. One part below 1 ppm and another above 1 ppm, both with their own T_1 and T_2 parameters, T_1's restricted to >0.1 s and T_2's to >10 ms.

Starting values

For the synthetic data, the fitting started with values $\pm 5\%$ off the true values except for the area parameter for the MMBL; which had equal area for all Voigt lines in the pattern (mean

true value). In vivo, the starting values were chosen based on previous experience to be close to the expected mean values for metabolites, and equal area for the MMBL components.

Fitting method

The iterative 2DJ-IR method consisted of 3 IR fitting iterations interleaved with two 2DJ fitting iterations. Fitted parameters after each iteration were copied as starting values for the next fitting iteration. The 2DJ-only and IR-only methods consisted of only 1 fitting iteration.

Fit strategy steps within each iteration

Fitting was always performed using the complex channel of the frequency domain data using the Levenberg Marquardt algorithm with boundary conditions. The first point of the FIDs, which is often distorted by digital or analog filtering in the receive chain, was truncated and the fit ranges restricted in FD to 0.5 ppm to 4.3 ppm.

IR

1. Fit only the amplitudes and r_1 parameter of all macromolecule models. (47 free variables to fit)
2. Fit frequency offset for all metabolites. (1 free variable)
3. Fit amplitude and r_1 for all metabolites and macromolecules. (70 free variables)
4. Fit amplitude, frequency offset, Gaussian width, and r_1 for all metabolites and macromolecules; Lorentzian widths only in case of IR-only method. (80 free variables)

2DJ

1. Fit amplitude and Lorentzian width for MMBL only. (47 free variables)
2. Fit frequency offset for all metabolites. (1 free variable)
3. Fit amplitude and Lorentzian width for all metabolites. (22 free variables)
4. Fit all parameters except r_1 for all metabolites and macromolecules. (73 free variables)

Calculation times

CPU times for modeling do not vary primarily with different 2D methods, but mainly depend on the number of free fitting variables and the number of spectral points and traces in a 2D spectrum. Furthermore, overall calculation times depend on the starting conditions. For the particular modeling cases described above, the CPU times per iteration on a iMac (2009 model, 2.93 GHz Intel Core i7 CPU) computer were around

33 sec for the 2DJ-only (24 spectra, 73 free parameters) and 5 sec for the IR-only (7 spectra, 80 free parameters) cases in vivo.

Appendix B: Software implementation

Project definition

Experiment type, experimental parameters, the specific hierarchical model (incl. prior knowledge relations in both dimensions, fit strategy, and numerical patterns), and optionally accompanying data are defined and stored as projects. Hierarchical structures can be collapsed into single numeric patterns to reduce model complexity. Prior knowledge can be applied on the same hierarchical levels only.

Fit strategies and minimization algorithms

For each fit strategy step, the following fitting specifics can be defined: fit domain (χ^2-summation in TD or FD), data channel (in FD domain: absorption, dispersion, complex, magnitude, or power channel), data truncation range (leave out points at start and/or end of TD signal, Eqs. (3, 4), data range for χ^2-summation in frequency-selective fitting in FD (Eq. 5), fitting algorithm, parameter sets, and spectra. Model parameters of the previous step are used as initial values for a subsequent step.

An option for optimization of starting values for 2D fitting is a sequential fit, where each spectrum is fitted separately in a time-saving [8] manner. However, this option may lead to violated prior knowledge along the second dimension.

The model is calculated in TD. Several minimization algorithms are available: Levenberg-Marquardt (LM) and LM with simple parameter boundary conditions (Bounded LM) [40]. L-BFGS [41], random search [42], conjugate direction, and conjugate gradient algorithms. The bounded LM in TD and FD was predominantly used; it enforces simple boundary constraints for all free parameters, while the other implemented algorithms either do not include parameter bounds or are slow.

Basis spectra and data processing

Import of simulated or measured model spectra is implemented. GAVA [43] is used for simulations based on published parameters [44]. During import, metabolite patterns can be split at predefined locations into multiple numeric subpatterns to allow for hierarchical model spectra.

FiTAID contains standard spectral processing tools, including zero filling, phasing, apodization, frequency shifting, or eddy current correction with a reference spectrum.

File formats supported

Siemens spectroscopy DICOM format, jMRUI ASCII Text Format [3], GAVA FITT format [43], Siemens RAW floating point format.

Other aspects

FiTAID was implemented in java. The Graphical User Interface [28] allows the user to quickly import, model, view, and fit a spectrum. During modeling, parameters can be rapidly changed with immediate visual feedback on the model.

It is planned to make FiTAID available as executable as well as open source software as soon as a stable performance on multiple platforms can be assured, some of the code has been refactored for clearer structure and as soon as a manual or tutorial is available. FiTAID allows for extension by the user to define new 2D experiments.

System requirements

Windows, Mac OSX, or Linux with JVM 1.5 or later. Minimum 1 GB memory available to FiTAID.

Source code used in FiTAID

dcm4che2[3]: Open Source Clinical Image and Object Management. v2.0.19
The RISO Project[4]: Limited-memory Broyden-Fletcher-Goldfarb-Shanno (LBFGS) algorithm.
BEAST[5]: Conjugate Gradient and Conjugate Direction Algorithm.
Minpack_f77[6]: LM Algorithm.
SQLiteJDBC[7]: Java JDBC driver for SQLite. v0.54
MySQL Connect/J[8]: Official JDBC driver for MySQL. v5.1.6
Disclosure Triangle[9] GUI component.
Colt[10]: High Performance Scientific and Technical Computing in Java. v2.0

---

[3] http://www.dcm4che.org/.

[4] http://riso.sourceforge.net/.

[5] http://code.google.com/p/beast-mcmc/.

[6] http://www1.fpl.fs.fed.us/Minpack_f77.java.

[7] http://www.zentus.com/sqlitejdbc/.

[8] http://dev.mysql.com/downloads/connector/j/5.1.html.

[9] http://www.dishevelled.org/disclosure-triangle/source-repository.html.

[10] http://acs.lbl.gov/~hoschek/colt/.

Current limitations

Among the current limitations one should mention: a) There is presently no batch-processing mode, which is a highly desirable feature in a clinical setting. b) Only the area parameter is implemented as a factor that is accessible for manipulation in the model calculation along the second dimension, changes in other parameters are currently only possible as a change in the base patterns (e.g. TE variation that affects phase and overall pattern). c) FFT, and hence frequency range selection is not implemented for the second dimension. d) No baseline elimination method other than description as a number of Voigt lines is available.

References

1. Provencher SW (1993) Estimation of metabolite concentration from localized in vivo proton NMR spectra. Magn Reson Med 30:672–679
2. Slotboom J, Boesch C, Kreis R (1998) Versatile frequency domain fitting using time domain models and prior knowledge. Magn Reson Med 39:899–911
3. Stefan D, Di Cesare F, Andrasescu A, Popa E, Lazariev A, Vescovo E, Strbak O, Williams S, Starcuk Z, Cabanas M, van Ormondt D, Graveron-Demilly D (2009) Quantitation of magnetic resonance spectroscopy signals: the jMRUI software package. Measure Sci Technol 20:104035
4. Poullet JB, Sima DM, van Huffel S (2008) MRS signal quantitation: A review of time- and frequency-domain methods. J Magn Reson 195:134–144
5. van Ormondt D, de Beer R, Marien AJH, den Hollander JA, Luyten PR, Vermeulen JWAH (1990) 2D approach to quantitation of inversion-recovery data. J Magn Reson 88:652–659
6. de Beer R, van Ormondt D, Pijnappel WWF (1992) Quantification of 1-D and 2-D magnetic-resonance time domain signals. Pure Appl Chem 64:815–823
7. Kiefer AP, Govindaraju V, Matson GB, Weiner MW, Maudsley AA (1998) Multiple-echo proton spectroscopic imaging using time domain parametric spectral analysis. Magn Reson Med 39:528–538
8. Vanhamme L, van Huffel S, van Hecke P, van Ormondt D (1999) Time-domain quantification of series of biomedical magnetic resonance spectroscopy signals. J Magn Reson 140:120–130
9. Kreis R, Slotboom J, Hofmann L, Boesch C (2005) Integrated data acquisition and processing to determine metabolite contents, relaxation times, and macromolecule baseline in single examinations of individual subjects. Magn Reson Med 54:761–768
10. Schulte RF, Boesiger P (2006) ProFit: two-dimensional prior-knowledge fitting of J-resolved spectra. NMR Biomed 19:255–263
11. Thomas MA, Lange T, Velan SS, Nagarajan R, Raman S, Gomez A, Margolis D, Swart S, Raylman RR, Schulte RF, Boesiger P (2008) Two-dimensional MR spectroscopy of healthy and cancerous prostates in vivo. Magn Reson Mater Phy 21:443–458
12. Gonenc A, Govind V, Sheriff S, Maudsley AA (2010) Comparison of spectral fitting methods for overlapping J-coupled metabolite resonances. Magn Reson Med 64:623–628
13. Garcia-Gomez JM, Tortajada S, Vidal C, Julia-Sape M, Luts J, Moreno-Torres A, van Huffel S, Arus C, Robles M (2008) The effect of combining two echo times in automatic brain tumor classification by MRS. NMR Biomed 21:1112–1125
14. Behar KL, Rothman DL, Spencer DD, Petroff OAC (1994) Analysis of macromolecule resonances in 1H NMR spectra of the brain. Magn Reson Med 32:294–302
15. Hofmann L, Slotboom J, Boesch C, Kreis R (2001) Characterization of the macromolecule baseline in localized ^{1}H-MR spectra of human brain. Magn Reson Med 46:855–863
16. Ratiney H, Coenradie Y, Cavassila S, van Ormondt D, Graveron-Demilly D (2004) Time-domain quantitation of (1)H short echo-time signals: background accommodation. Magn Reson Mater Phy 16:284–296
17. Ratiney H, Sdika M, Coenradie Y, Cavassila S, van Ormondt D, Graveron-Demilly D (2005) Time-domain semi-parametric estimation based on a metabolite basis set. NMR Biomed 18:1–13
18. Elster C, Schubert F, Link A, Walzel M, Seifert F, Rinneberg H (2005) Quantitative magnetic resonance spectroscopy: semi-parametric modeling and determination of uncertainties. Magn Reson Med 53:1288–1296
19. Rabeson H, Ratiney H, van Ormondt D, Graveron-Demilly D (2007) Semi-parametric estimation in magnetic resonance spectroscopy: automation of the disentanglement procedure. Conf Proc IEEE Eng Med Biol Soc 2007:662–665
20. Kreis R, Boesch C (1994) Liquid-crystal-like structures of human muscle demonstrated by in vivo observation of direct dipolar coupling in localized proton magnetic resonance spectroscopy. J Magn Reson Series B 104:189–192
21. Thomas MA, Ryner LN, Mehta MP, Turski PA, Sorenson JA (1996) Localized 2D J-resolved 1H MR spectroscopy of human brain tumors in vivo. J Magn Reson Imag 6:453–459
22. Hurd RE, Gurr D, Sailasuta N (1998) Proton spectroscopy without water suppression: the oversampled J- resolved experiment. Magn Reson Med 40:343–347
23. Hwang TL, van Zijl PC, Garwood M (1999) Asymmetric adiabatic pulses for NH selection. J Magn Reson 138:173–177
24. De Graaf RA, Sacolick LI, Rothman DL (2006) Water and metabolite-modulated MR spectroscopy and spectroscopic imaging. In: Proceedigs of 14th meeting of the ISMRM, Seattle, p 3063
25. Gruetter R, Tkac I (2000) Field mapping without reference scan using asymmetric echo-planar techniques. Magn Reson Med 43:319–323
26. Klose U (1990) In vivo proton spectroscopy in presence of eddy currents. Magn Reson Med 14:26–30
27. National Institute of Standards and Technology (NIST)/SEMA-TECH e-Handbook of Statistical Methods (2010) 1.3.5.10. Levene Test for Equality of Variances. http://www.itl.nist.gov/div898/handbook/
28. Slotboom J, Brekenfeld C, Kiefer C, Kreis R, Remonda L, Schroth G (2004) A graphic tool for fast interactive spectroscopic prior knowledge modeling and prior knowledge database management. In: Proceedings of 12th Meeting of the ISMRM, Kyoto, p 2263
29. Reynolds G, Wilson M, Peet A, Arvanitis TN (2006) An algorithm for the automated quantitation of metabolites in in vitro NMR signals. Magn Reson Med 56:1211–1219
30. Yablonskiy DA, Neil JJ, Raichle ME, Ackerman JJ (1998) Homonuclear J coupling effects in volume localized NMR spectroscopy: pitfalls and solutions. Magn Reson Med 39:169–178
31. Chong DGQ, Nuoffer J, Vermathen P, Slotboom J, Boesch C, Kreis R (2009) Histidine transport dynamics across the healthy human blood-brain barrier investigated by 1H MRS. In: Proceedings of 17th meeting of the ISMRM, p 2346
32. Vermathen P, Capizzano AA, Maudsley AA (2000) Administration and (1)H MRS detection of histidine in human brain: application to in vivo pH measurement. Magn Reson Med 43:665–675
33. Cudalbu C, Bucur A, Graveron-Demilly D, Beuf O, Cavassila S (2007) Comparison of two strategies of background-accommodation: influence on the metabolite concentration estimation from

in vivo Magnetic Resonance Spectroscopy data. Conf Proc IEEE Eng Med Biol Soc 2007:2077–2080

34. Kunz N, Cudalbu C, Mlynarik V, Huppi PS, Sizonenko SV, Gruetter R (2010) Diffusion-weighted spectroscopy: a novel approach to determine macromolecule resonances in short-echo time (1)H-MRS. Magn Reson Med 64:939–946

35. Cudalbu C, Mlynarik V, Xin L, Gruetter R (2009) Quantification of in vivo short echo-time proton magnetic resonance spectra at 14.1 T using two different approaches of modelling the macromolecule spectrum. Measure Sci Technol 20:104034

36. Hurd R, Sailasuta N, Srinivasan R, Vigneron DB, Pelletier D, Nelson SJ (2004) Measurement of brain glutamate using TE-averaged PRESS at 3T. Magn Reson Med 51:435–440

37. Slotboom J, Nirkko A, Brekenfeld C, van Ormondt D (2009) Reliability testing of in vivo magnetic resonance spectroscopy (MRS) signals and signal artifact reduction by order statistic filtering. Measure Sci Technol 20:104030

38. van Ormondt D (2006) Error estimation. Role of CRB's in a semi-parametric context. In: Proceedings of ISMRM workshop on data processing for MR spectroscopy and imaging, Warrenton, VA, USA

39. Chong DGQ, Slotboom J, Kreis R (2007) Feasibility of using predicted Cramer Rao lower bounds for the design of optimized in vivo MR spectroscopy sequences targeting multiple metabolites. In: Proceedings of 15th meeting of the ISMRM, Berlin, p. 1406

40. Lourakis MIA (2004) Levenberg-Marquardt nonlinear least squares algorithms in {C}/{C}++. http://www.ics.forth.gr/~lourakis/levmar

41. Zhu C, Byrd R, Lu P, Nocedal J (1997) Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization. ACM Trans Math Software 23:550–560

42. Hofmann L, Slotboom J, Jung B, Maloca P, Boesch C, Kreis R (2002) Quantitative 1H-magnetic resonance spectroscopy of human brain: Influence of composition and parameterization of the basis set in linear combination model fitting. Magn Reson Med 48:440–453

43. Soher BJ, Young K, Bernstein A, Aygula Z, Maudsley AA (2007) GAVA: spectral simulation for in vivo MRS applications. J Magn Reson 185:291–299

44. Govindaraju V, Young K, Maudsley AA (2000) Proton NMR chemical shifts and coupling constants for brain metabolites. NMR Biomed 13:129–153