Comparison of linear combination modeling strategies for GABA-edited MRS at 3T

Helge J. Zöllner[1,2], Sofie Tapper[1,2], Steve C. N. Hui[1,2], Peter B. Barker[1,2], Richard A. E. Edden[1,2], Georg Oeltzschner[1,2,*]

[1] *Russell H. Morgan Department of Radiology and Radiological Science, The Johns Hopkins University School of Medicine, Baltimore, MD, United States*

[2] *F. M. Kirby Research Center for Functional Brain Imaging, Kennedy Krieger Institute, Baltimore, MD, United States*

***Corresponding author:**

Georg Oeltzschner, Ph.D.

Division of Neuroradiology, Park 367G

The Johns Hopkins University School of Medicine

600 N Wolfe St

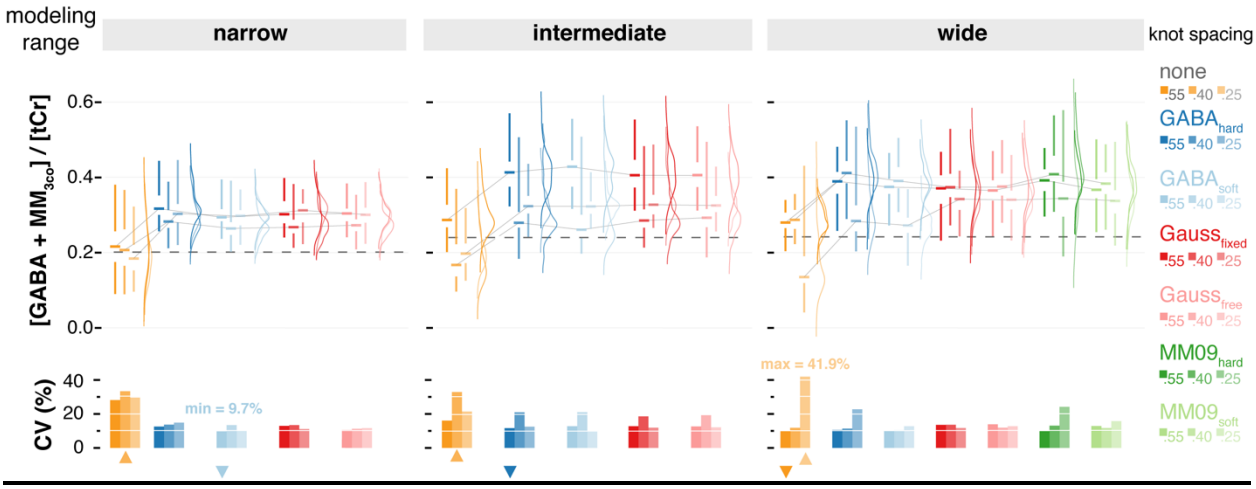
Baltimore, MD 21287

goeltzs1@jhmi.edu

Running title: Linear combination modeling of GABA-edited MRS

Word count: 7239

Graphical Abstract



102 strategies to model GABA-edited MRS with linear combination modeling were evaluated to quantify GABA and GABA+ in Osprey. Significantly different GABA and GABA+ estimates were found when a well-parameterized macro-molecule at 3 ppm was included. The findings suggest that linear combination modeling needs to be adapted for quantification of GABA-edited MRS.

Abbreviations

γ-aminobutyric-acid – GABA; Linear combination modeling – LCM; macromolecule – MM; GABA + MM - GABA+; homocarnosine – HCar; glutamate – Glu; glutamine – Gln; glutathione – GSH; N-acetylaspartylglutamate – NAAG; N-acetylaspartate – NAA; Hankel singular value decomposition – HSVD; full-width at half-maximum – FWHM; creatine – Cr; negative creatine methylene – -$CrCH_2$; phosphocreatine – PCr; SD – standard deviation; Akaike Information Criterion – AIC; coefficients of variation – CVs;

38

39 **Keywords**

40 • Magnetic resonance spectroscopy

41 • Linear combination modeling

42 • GABA-edited MEGA-PRESS

Abstract

Purpose

J-difference-edited spectroscopy is a valuable approach for the in vivo detection of γ-aminobutyric-acid (GABA) with MRS. A recent expert consensus article recommends linear combination modeling (LCM) of edited MRS, but does not give specific details of implementation. This study explores different modeling strategies to adapt LCM for GABA-edited MRS.

Methods

61 medial parietal lobe GABA-edited MEGA-PRESS spectra from a recent 3T multi-site study were modeled using 102 different strategies combining six different approaches to account for co-edited macromolecules, three modeling ranges, three baseline knot spacings, and the use of basis sets with or without homocarnosine. The resulting GABA and GABA+ estimates (quantified relative to total creatine), the residuals at different ranges, SDs and CVs, and Akaike information criteria, were used to evaluate the models' performance.

Results

Significantly different GABA+ and GABA estimates were found when a well-parameterized MM_{3co} basis function was included in the model. The mean GABA estimates were significantly lower when modeling MM, while the CVs were similar. A sparser spline knot spacing led to lower variation in the GABA and GABA+ estimates, and a narrower modeling range – only including the signals of interest – did not substantially improve or degrade modeling performance. Additionally, results suggest that LCM can separate GABA and the underlying co-edited MM_{3co}. Incorporating homocarnosine into the modeling did not significantly improve variance in GABA+ estimates.

Conclusion

GABA-edited MRS is most appropriately quantified by LCM with a well-parameterized co-edited MM_{3co} basis function with a constraint to the non-overlapped $MM_{0.93}$, in combination with a sparse spline knot spacing and a modeling range between 0.5 and 4 ppm.

Introduction

A recent expert consensus paper recommended that linear combination modeling (LCM) should be used for the quantification of edited MRS data[1] , stating that standard fitting approaches originally optimized for short-TE MRS should be adapted for edited MRS. Further, it was recommended that quantum-mechanical simulations should be used to confirm the co-edited profile of all metabolites in the edited spectrum, and contributions from macromolecule (MM) signals should be specified. Despite these recommendations, little detail was given regarding several unique features of edited spectra, and how they should be appropriately modeled. These features include:

1) The MEGA-PRESS experiment is well-known to co-edit MM signals with coupled spins at 1.7 and 3 ppm, causing substantial contamination of the edited GABA signal, and forcing researchers to report the composite measure GABA+MM (GABA+)[1]. Because the co-edited MM signal is poorly characterized, there is currently no consensus or recommendation on how to appropriately account for it during spectral modeling. Instead, the most widely used analysis algorithms implement entirely different strategies to fit the composite 3-ppm signal. For example, the Gannet software uses a single Gaussian model[2], while a double-Gaussian is used in Tarquin[3], and LCModel[4] defaults to a basis set that only includes the GABA basis function.
2) Another co-edited compound contributing to the 3 ppm signal is homocarnosine (HCar), a dipeptide of GABA and histidine. While the 3 ppm multiplets of GABA and homocarnosine are separated by just 0.05 ppm (which are therefore unlikely to be successfully separated), inclusion of a homocarnosine basis function may be warranted based on its reported concentration in vivo (~0.5 mmol/kg [5], compared to ~1-2 mmol/kg for GABA), but it has not been investigated whether doing so has a stabilizing or destabilizing effect on the modeling[6].
3) Unedited spectra are typically modeled over a restricted frequency-domain range covering the visible upfield peaks, including macromolecular and lipid resonances between 0 and 1 ppm, but usually avoiding the water suppression window above ~4 ppm. The choice of frequency-domain modeling range for edited spectra is less obvious. Since

the main advantage of spectral editing is the isolation of a single target resonance, modeling signals outside the immediate surrounding of the target may dilute the resolving power of editing. On the other hand, increasing the modeling range may offer useful constraints to stabilize the solution of the modeling problem. The difference is highlighted by the different strategies encountered in common software tools – while the Gannet software fits the GABA-edited difference spectrum over a narrow range (only including the 3-ppm GABA+ and 3.75 ppm glutamate and glutamine peaks), the LCModel recommendation is to include the strong co-edited signals from glutamate (Glu), glutamine (Gln), glutathione (GSH), N-acetylaspartylglutamate (NAAG), and N-acetylaspartate (NAA), which heavily overlap with GABA around 2.25 ppm. The effects of limiting the modeling range have not been assessed systematically to date.

4) Linear combination modeling methods commonly include terms to account for smooth baseline curvature, usually parametrized from cubic B-spline or polynomial functions, or by smoothing residuals. The flexibility of the baseline model substantially affects metabolite estimates from unedited spectra[7]; while baseline terms are necessary to account for e.g. lipid contamination, poor water suppression etc., they are potential sources of overfitting if awarded too many degrees of freedom. Baseline modeling may have an even greater influence when modeling difference spectra, since only *co-edited* lipid and MM signals contribute to the smooth background variation. Importantly, the co-edited MM background of the GABA-edited difference spectrum has not been appropriately characterized (e.g., through metabolite-nulled acquisition), suggesting that the choice of baseline flexibility can drastically influence modeling results through two highly susceptible regions of the spectrum. First, in the absence of an appropriate model for the co-edited broad MM signal at 3 ppm, this signal may be absorbed into the baseline depending on its flexibility. Second, strong MM and lipid signals in the region between 0.5 and 2.5 ppm may be affected by the 1.9 ppm editing pulse (either directly through saturation or indirectly through coupling), likely leading to an unknown, but substantial, MM contribution in this spectral region[8,9]. This is especially important considering that the co-edited signals from NAA, NAAG, Glu, Gln, and GSH overlap with GABA in this region. Overly rigid baselines may provide insufficient flexibility to capture these signals, in turn compromising the accuracy of the estimation of the co-edited metabolites.

The aim of this study was to evaluate different strategies for linear combination modeling of GABA-edited MEGA-PRESS difference spectra, and to establish initial 'best practices' substantiating the recommendations of the expert consensus on spectral editing. To this end, different approaches to account for co-edited MM signals, various modeling ranges and baseline knot spacings, as well as the inclusion of homocarnosine were compared. In the absence of a 'gold standard', the performance of each modeling strategy was assessed by comparing descriptive statistics of the metabolite estimates, calculating the Akaike information criteria, and assessing the fit residuals.

Methods

Study participants & data acquisition

In this study, 61 publicly available GABA-edited MEGA-PRESS datasets originating from 7 sites from a recent 3T multi-center study[10] were analyzed (see **Supplementary Material 1** for subject list). All datasets were acquired on Philips 3T scanners with the following acquisition parameters: TR/TE = 2000/68 ms; 320 excitations (10m 40s scan time); 16-step phase-cycle; 2 kHz spectral width; 2000 samples; 27-ml cubic voxel volume in the medial parietal lobe. For this heuristic approach of exploring the GABA modeling the data homogeneity (SNR, FWHM, tissue composition, and absence of fat contamination) was increased while reducing the overall number of subjects by including only 61/298 subjects of the original dataset[10]. All sites except for P8 used a similar sequence implementation with interleaved water referencing for prospective frequency correction[11]. For the edit-ON transients, the editing pulses with 15 ms pulse duration and 82.5 Hz inversion bandwidth (FHWM) were applied at a frequency of 1.9 ppm to refocus the coupling evolution of the GABA spin system. For the edit-OFF transients, the editing pulses were applied at a frequency of 7.5 ppm. Edit-ON and edit-OFF transients were acquired in alternating order. An additional water reference scan was acquired for each dataset using interleaved water referencing [11], i.e. one excitation with water suppression and editing pulses deactivated every 40 water-suppressed excitations (total of 8 averages).

Data pre-processing

Data were analyzed in MATLAB using Osprey[12,13] (v.1.0.1.1), a recently published open-source MRS analysis toolbox. Raw data were eddy-current-corrected [14] based on the water reference, and individual transients were aligned separately within the edit-ON and edit-OFF conditions using the robust spectral registration algorithm[15]. Averaged edit-ON and edit-OFF spectra were aligned by optimizing relative frequency and phase such that the water signal in the difference spectrum was minimized. The final difference spectra for quantification were generated by subtracting the edit-OFF from the edit-ON spectra. Finally, any residual water signal was removed with a Hankel singular value decomposition (HSVD) filter[16] to improve data quality in the edit-OFF spectra and to reduce residual baseline roll in the difference spectra.

Basis set

The basis set used for modeling was generated from a fully localized 2D density-matrix simulation of a 101 x 101 spatial grid (voxel size: 30 mm x 30 mm x 30 mm; field of view: 45 mm x 45 mm x 45 mm) implemented in a MATLAB based simulation toolbox FID-A [17], using vendor-specific refocusing pulse shape and duration, sequence timings, and phase cycling. It contains 17 metabolite basis functions (ascorbate, aspartate, creatine (Cr), negative creatine methylene (-$CrCH_2$), GABA, glycerophosphocholine, GSH, Gln, Glu, water, myo-inositol, lactate, NAA, NAAG, phosphocholine, phosphocreatine (PCr), phosphoethanolamine, scyllo-inositol, and taurine) and 8 Gaussian MM and lipid resonances ($MM_{0.94}$, $MM_{1.22}$, $MM_{1.43}$, $MM_{1.70}$, $MM_{2.05}$, Lip09, Lip13, Lip20, details in **Supplementary Material 2** with similarly defined parametrization as described in the LCModel software manual[18]) for the edit-OFF spectrum.

For the difference spectrum, $MM_{0.94}$ and the co-edited macromolecular signal at 3 ppm (MM_{3co}) were parametrized as Gaussian basis functions ($MM_{0.94}$: 3-proton signal; chemical shift 0.915 ppm, full-width at half-maximum (FWHM) 11 Hz; MM_{3co}: 2-proton signal; chemical shift 3 ppm; FWHM 14 Hz). The MM_{3co} amplitude was defined under the assumptions of a pseudo-doublet GABA signal at 3 ppm and the MM_{3co} contribution to the 3-ppm GABA peak to be around 50%[1,6,8,19]. The optimum FWHM used to parametrize the MM_{3co} basis function was determined to be 14 Hz by fitting the mean difference spectrum of all datasets with a composite GABA+ basis function (GABA + MM_{3co}) with varying FHWM (between 1 and 20 Hz). The parameterized Gaussian MM_{3co} basis function was integrated into the modeling process using different assumptions and constraints described in the following paragraphs.

Linear combination modeling of GABA-edited difference spectra

Osprey's frequency-domain linear combination model was used to determine the metabolite estimates. Model parameters include metabolite basis function amplitudes, frequency shifts, zero/first order phase correction, Gaussian and Lorentzian linebroadening, and cubic spline baseline coefficients. All parameters are determined by Levenberg-Marquardt[20,21] non-linear least-squares optimization, using a non-negative least-squares (NNLS) fit [22–24] to determine the metabolite amplitudes and baseline coefficients at each iteration of the non-linear optimization.

200 Amplitude ratio soft constraints are imposed on MM and lipid amplitudes, as well as selected
201 pairs of metabolite amplitudes, as defined in the LCModel manual[4,18]. The strength factor of the
202 amplitude ratio soft constraint λ is set to 0.05 by default.

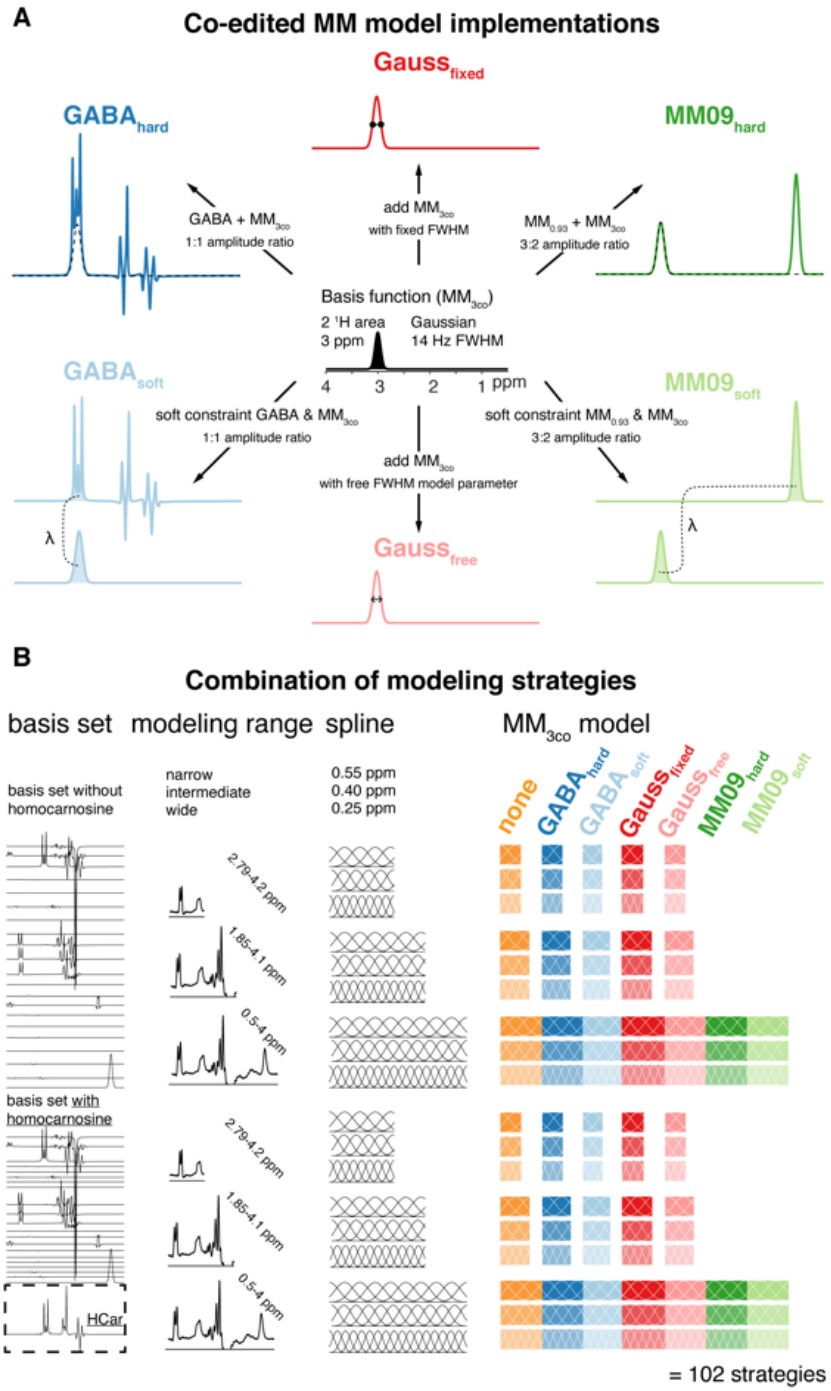


Figure 1 – Different linear combination modeling strategies for GABA-edited spectra. (A) Different co-edited MM_{3co} modeling approaches derived from a Gaussian function at 3.0 ppm (B)

All combinations of basis set composition, modeling range, spline knot spacing, and MM_{3co} modeling leading to 102 different modeling strategies.

A range of modeling strategies for the GABA-edited difference spectrum was included in this study, covering various aspects of the modeling process (**Figure 1**). The different parametrizations and soft constraints to account for the co-edited MM_{3co} signal are shown in **Figure 1A**. All possible combinations for the modeling strategies: i) inclusion of homocarnosine in the basis set; ii) different parametrizations and soft constraints to account for the co-edited MM_{3co} signal; iii) different modeling ranges and iv) different baseline spline knot spacings are tabulated graphically in **Figure 1 B**. Each modeling aspect is described in detail below:

Including homocarnosine in the basis set

To assess the effects of including homocarnosine in the linear combination model, we repeated all analysis steps with two different basis sets: the default Osprey basis set *with* and *without* an additional HCar basis function. Chemical shift and scalar coupling parameters describing the HCar spin system were taken from literature[6].

Varying the modeling range and baseline knot spacing

Two aspects of linear combination modeling are suggested to have a considerable influence on metabolite estimates[7,25]. First, the choice of the modeling range, i.e., the frequency interval that defines the part of the frequency-domain spectrum that is considered to calculate the least-squares difference between model and data. Second, the baseline knot spacing, i.e., the frequency difference between two adjacent knots of the cubic spline basis that is used to approximate the smooth baseline.

Three different modeling range scenarios were considered, reflecting common choices in the literature and widely used software tools: a) a wide modeling range typically used to analyze unedited spectra, including all signals in the GABA-edited difference spectrum (0.5 to 4 ppm – *"wide"*); b) an intermediate modeling range excluding signals below 1.9 ppm (e.g. co-edited lipids and macromolecules), but including strong co-edited signals from NAA, NAAG, Glu, Gln, and GSH (1.85 and 4.1 ppm, *"intermediate"*), similar to the range recommended in LCModel's dedicated 'mega-press-3' option; and c) a narrow modeling range only including the co-edited

signals from GABA+ and Glx (2.79 – 4.2 ppm, *"narrow"*), the default modeling range in Gannet[2].

Three spline knot spacings were included in the analysis, with 0.4 ppm being the default Osprey option, shown to create reproducible and comparable metabolite estimates for conventional MRS [26], as well as sparser (0.55 ppm) and denser (0.25 ppm) spline knot spacings.

Co-edited macromolecule models

Seven different strategies to model the GABA-edited difference spectrum were implemented (**Figure 1 A**). The trivial approach – not accounting for the co-edited signal MM_{3co} at all – is labeled **none**. The other six modeling strategies all include a dedicated parametrized Gaussian MM_{3co} basis function. This basis function is given different degrees of freedom in the different strategies, e.g. hard- or soft-constrained relative to the amplitude of the GABA or the $MM_{0.94}$ basis functions, and with a fixed or free width. Here, strategies with fewer degrees of freedom reflect the frequently made assumption that the GABA-to-MM ratio (and the MM background itself) is relatively stable across subjects and anatomical region, and assumed to be known, while strategies with more degrees of freedom or soft constraints relax these assumptions:

- The **$GABA_{hard}$** model uses a single composite GABA+MM basis function by adding the GABA and MM_{3co} (initial FWHM of the basis function = 14 Hz) basis functions with a fixed 1:1 amplitude ratio. The 1:1 ratio reflects the widely used empirical assumption that 50% of the 3-ppm signal in a conventional GABA-edited difference spectrum can be attributed to co-edited macromolecules[6,19].
- The **$GABA_{soft}$** model uses separate GABA and MM_{3co} (initial FWHM of the basis function = 14 Hz) basis functions and imposes a soft constraint on the ration of the amplitudes of both basis functions during the optimization (1:1 ratio).
- The **$Gauss_{fixed}$** model uses separate GABA and MM_{3co} (initial FWHM of the basis function = 14 Hz) basis functions. No further constraints are imposed. This means possible changes in the contributions to the 3-ppm GABA peak are modeled.
- The **$Gauss_{free}$** model uses separate GABA and MM_{3co} basis functions. In contrast to the $Gauss_{fixed}$ model, the FWHM of the Gaussian MM_{3co} signal is represented by an additional

model parameter. This means that the MM_{3co} basis function itself is not static, but dynamically modified during optimization.

- The **$MM09_{hard}$** model uses separate GABA and MM basis functions. The MM_{3co} basis function is replaced by a composite $MM_{0.94}$ + MM_{3co} basis function (i.e., the $MM_{0.94}$ (initial FWHM of the basis function = 11 Hz) and MM_{3co} (initial FWHM of the basis function = 14 Hz) basis functions are added in a 3:2 ratio). Resulting in a single composite basis function for $MM_{0.94}$ and MM_{3co}, adapted from the soft constraint model described in literature [9].
- The **$MM09_{soft}$** model uses separate GABA, $MM_{0.94}$ and MM_{3co} basis functions. In contrast to the $MM09_{hard}$ model, soft constraints enforce a ~3:2 amplitude ratio for the $MM_{0.94}$ and MM_{3co} amplitudes during optimization. Resulting in two separate basis functions for $MM_{0.94}$ and MM_{3co}, which is similar to previously described implementations [9].

The models $MM09_{hard}$ and $MM09_{soft}$ [27] as well as $Gauss_{fixed}$ [28] correspond to models previously investigated using the LCModel software and the amplitude assumptions were derived empirically. It is worth noting that each basis function receives a separate Lorentzian linebroadening, frequency shift, and amplitude parameter during the optimization. For the $Gauss_{free}$ model, the MM_{3co} basis function is dynamically updated as an explicit modeling parameter during the optimization, therefore the MM_{3co} basis function has effectively two separate parameters to account for its linewidth (the Lorentzian linebroadening term and the FWHM of the MM_{3co} basis function). Finally, the composite models ($GABA_{hard}$), which do not have separate GABA and MM functions, only have one linebroadening, one frequency, and one amplitude parameter compared to twice the parameters for the soft constraint counterparts.

Combining the various MM_{3co} models (5 + 2 that were used for the wide modeling range only), modeling ranges (3), baseline spline knot spacings (3), and basis sets (2), a total of 102 different modeling strategies were investigated in this study. All models were implemented in Osprey[12] and are available on GitHub[13].

Quantification, visualization, and statistics

Quantification

For the basis set without homocarnosine, GABA refers to the model amplitude estimate for the GABA basis function, which is of course only available for the modeling strategies with separate basis functions for GABA and MM_{3co} (none, $GABA_{soft}$, $Gauss_{fixed}$, $Gauss_{free}$, $MM09_{soft}$). GABA+ refers to the sum of the amplitude estimates for GABA and MM_{3co} ($GABA_{soft}$, $Gauss_{fixed}$, $Gauss_{free}$, $MM09_{hard}$, $MM09_{soft}$) or the amplitude estimate for the composite basis function including both MM and GABA ($GABA_{hard}$) and is therefore calculated for all strategies with an explicit MM_{3co} model. For comparison, the GABA amplitude for the `none` strategy is included in the figures reporting GABA + MM_{3co}. However, it still refers to a GABA-only estimate.

For the basis set that included homocarnosine (HCar), the difference in GABA and MM_{3co} estimates between the modeling strategies with and without HCar (ΔGABA and ΔMM_{3co}, respectively) were investigated to evaluate whether the inclusion of HCar has a systematic effect on the estimation of those signals with which it overlaps. All estimates were quantified relative to the total creatine (Cr + PCr) amplitude from the edit-OFF spectrum with the wide modeling range and a spline knot spacing of 0.4 ppm. Differences in GABA(+)/tCr between modeling strategies are therefore only related to the modeling of the difference spectra, but not to the reference compound modeling. No further tissue or relaxation corrections were applied.

Further, the relative contributions of MM_{3co} to the GABA+ estimate and the relative contributions of HCar to the sum of GABA+ and HCar estimate were calculated.

Visualization

The modeling performance and systematic characteristics of each modeling strategy were visually assessed through the mean spectra, mean fit, mean residual, and mean models of GABA+, GABA, $MM_{3co,}$ HCar (if included) and the baseline, i.e., averaged across all datasets.

The metabolite estimate distributions were visualized as violin plots including boxplots with median, 25^{th}/75^{th} quartile ranges, and smoothed distributions to identify systematic differences between modeling strategies. In addition, the mean value of the ‘none’ model across the three spline knot spacings was added for each modeling range as a dashed horizontal line. Bar plots

were created to visualize quality metrics, including the standard deviation if appropriate. All plots were generated with R[29] (Version 3.6.1) in RStudio (Version 1.2.5019, RStudio Inc.) using SpecVis[26,30], an open-source package to visualize linear combination modeling results with the ggplot2 package[31]. All scripts and results are publicly available[32].

Statistics

Significant differences in the mean and the variance of the GABA, GABA+, and MM_{3co} estimates were assessed between all modeling strategies. The statistical tests were set up as paired without any further inference. Differences of variances were tested with Fligner-Killeen's test, with a post-hoc pair-wise Bonferroni-corrected Fligner-Killeen's test. The means were compared with an ANOVA or a Welch's ANOVA, depending on whether variances were different or not. Post-hoc analysis was performed with a paired t-test with equal or non-equal variances, respectively.

Additionally, Pearson's correlation was used to investigate the impact of including HCar in the basis set. The strength of the correlation was considered substantial for $R > 0.25$.

Model evaluation criteria

The performance of each modeling strategy was evaluated in different ways, including the impact of the different modeling strategies on the GABA, GABA+ , and MM_{3co} estimates, as well as several quality measures:

1) Visual inspection: Mean model, residual, and baseline were assessed for characteristic features.
2) SD fit quality: The SD of the residual was determined, and then normalized by the noise level (calculated as the SD of the noise between -2 and 0 ppm). This is done over the entire modeling range of the difference spectrum and termed $\mathbf{residual_{SD\ range}}$.
3) Amplitude fit quality: the difference between the maximum and minimum of the residual was determined, and then normalized by the noise level [25] (similarly calculated as in the second criterion).This was done over the entire modeling range of the difference spectrum and termed $\mathbf{residual_{ampl\ range}}$.

4) Amplitude 3-ppm peak fit quality: Similar to the third criterion, the residual was calculated over the range of 3.027 ± 0.15 ppm to assess the fit quality of the 3-ppm GABA peak and termed **residual$_{\text{ampl 3ppm}}$**.
5) Consistency of metabolite estimates: The across-subject coefficients of variation (CV = SD/mean) for all metabolite estimates (GABA/tCr, GABA+/tCr) were calculated for each modeling strategy.
6) Akaike Information Criterion (AIC): The Akaike information criterion [33], which takes the number of model parameters into account, is defined as follows:

$$logLikelihood_i = -0.5 * N_i * (\log(2 * \pi) + 1 - \log(N_i) + \log(SSE_i))$$

$$AIC_i = -2 * \frac{1}{N_i} * logLikelihood_i + 2K_i$$

Here, N_i is the number of points in the modeling strategy i, SSE_i is the sum of squared error (i.e., squared residual) of that strategy, and K_i is the number of free model parameters for that strategy. The $logLikelidhood_i$ was divided by the number of points N_i to reduce the strong weighting of the datapoints and to make the AIC_i values comparable for different modeling ranges. Soft constraint model parameters were included with a value of 0.5. Lower AIC_i values indicate a more appropriate model. Subsequently, $\boldsymbol{\Delta AIC_i}$ scores were calculated as the difference of AIC_i of modeling strategy i and the model with the lowest AIC_{min}:

$$\Delta AIC_i = AIC_i - AIC_{min}$$

Results

All 61 datasets were successfully processed and modeled with all 102 modeling strategies. No data were excluded from further analysis. The data quality assessment indicated consistently high spectral quality for all spectra (NAA_{SNR} = 272 ± 70; NAA_{FWHM} = 5.29 ± 1.09 Hz) without lipid contamination (data overview in **Supplementary Material 4**). Other data quality measures were extracted from the Gannet[2] analysis performed in recent multi-site studies[10,34]. For the Philips-only subset of datasets in the present study, the tissue composition (fGM = 0.60 ± 0.04; fWM = 0.27 ± 0.03; fCSF = 0.13 ± 0.04) and across-subject CV (GABA+/Cr = 9.99%) indicate consistency in the dataset and the modeling. across-subject CV was interpreted as a measure of modeling performance, assuming that increased CVs are mainly introduced by variability in the modeling and do not reflect biologically meaningful variance of GABA+ estimates.

Summary and visual inspection of the modeling results

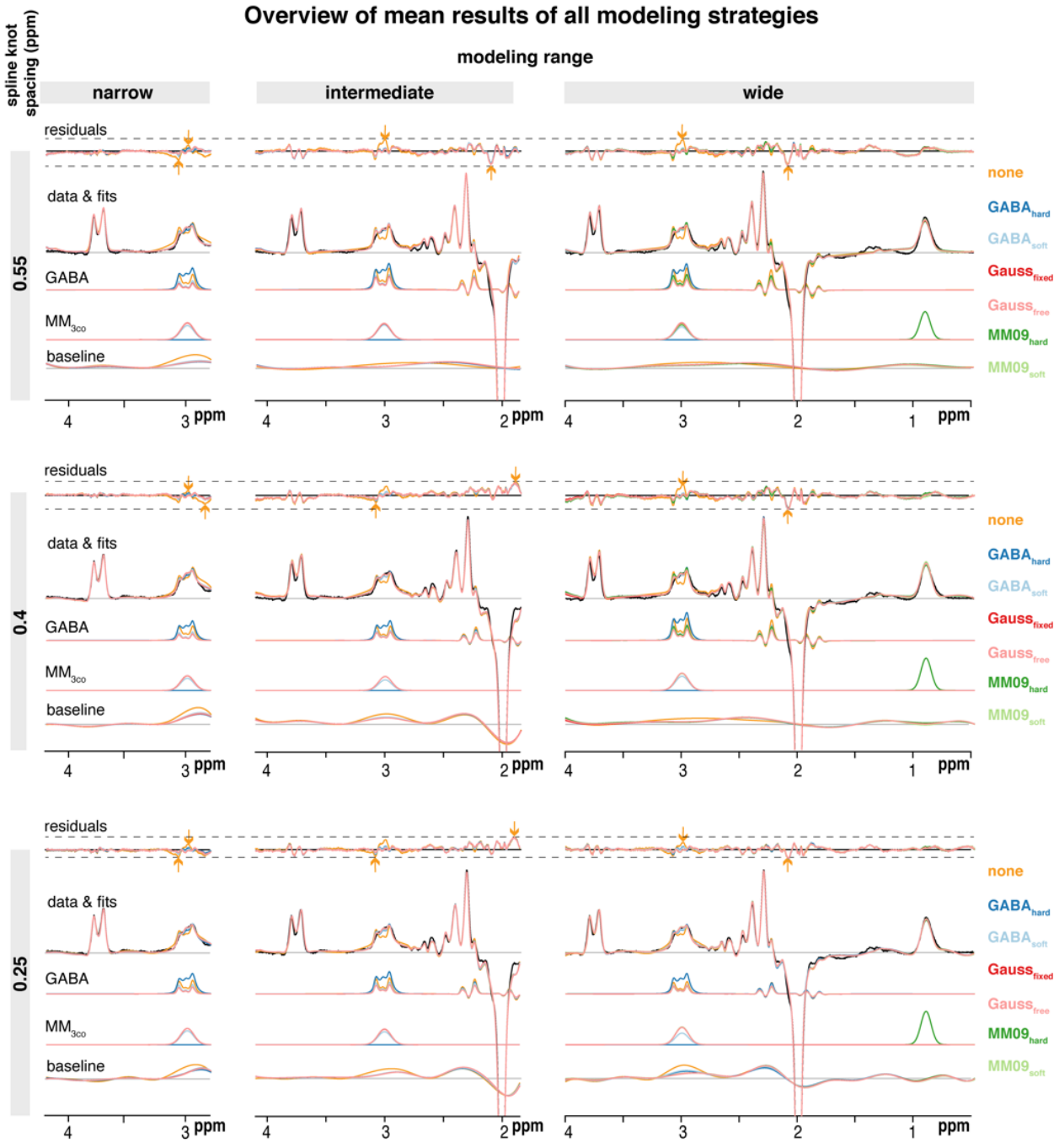


Figure 2 – Mean modeling results for all modeling strategies without homocarnosine. A substantial structured residual is apparent at 3 ppm if no MM modeling strategy is included. All three modeling ranges (columns), three spline knot spacings (rows), and MM_{3co} model (color-coded) are presented with mean residuals and fits, as well as the GABA+, GABA, MM_{3co}, and spline baseline models. The mean data is included in black. The dashed lines indicate the range of the residual across one row. The arrows indicate the range of values for a specific modeling range and spline knot spacing with the color corresponding to the MM_{3co} model with minimum/maximum value.

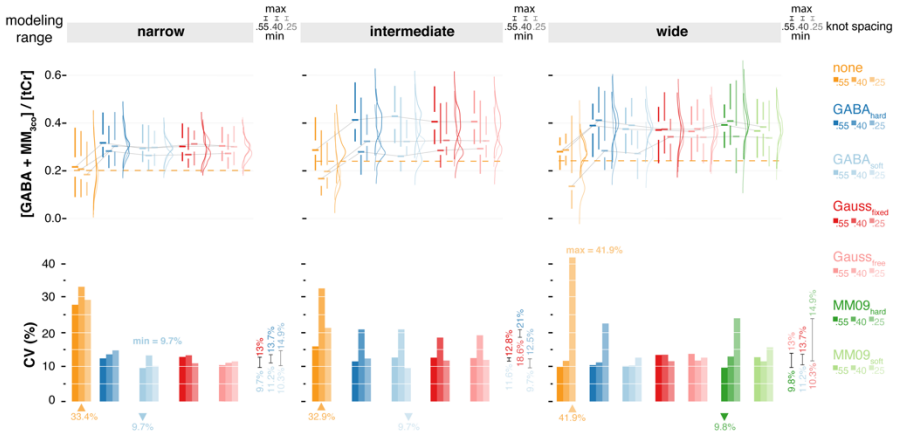
Figure 2 shows the mean modeling results for all modeling strategies without homocarnosine. Not including MM_{3co} leads to a substantial structured residual around 3 ppm for all knot spacings and modeling ranges. In contrast, all modeling strategies with MM_{3co} appear to reflect the lineshape of the 3-ppm signal more accurately, with very similar results for the complete fit (metabolites, MMs, and baseline) and the individual components. Modeling strategies with the intermediate and wide modeling range further show strong residuals around 2 ppm, suggesting slightly inaccurate lineshape modeling of the methyl singlets from NAA and NAAG, or inaccurate modeling of co-edited MM signals in this region. Structured residuals appear also in the region of the 3.75 ppm Glx signals, although they are much less pronounced in strategies with the narrow modeling range, suggesting that including the 2.25 ppm multiplets (and underlying baseline fluctuation) has a considerable impact on phase estimation.

In general, the residuals are consistent between different MM_{3co} models for any given knot spacing and modeling range. Notably, residuals tend to be smaller on an absolute scale for denser knot spacing and narrower modeling range.

Mean GABA models agree well between all strategies with a separate MM_{3co} model. The $GABA_{hard}$ strategy appears to produce a larger signal as its GABA basis function includes the MM_{3co} signal, but does not model it separately, while the strategies that do so produce comparable mean MM_{3co} models.

The mean baseline is consistently flatter around 3 ppm for modeling strategies with an explicit MM_{3co} model, while absorbing substantially more signal for the 'none' approach without an MM model. This behavior is particularly obvious for the dense knot spacing (0.25 ppm) over the wide modeling range. Baseline curvature generally increases for denser knot spacings around 2.2 ppm for the intermediate and wide range.

409 Metabolite level distribution



410

Figure 3 – Distribution and across subject coefficients of variation (CVs) of GABA+ estimates for all modeling strategies. Including a MM_{3co} model significantly increases the mean estimates for all modeling strategies, while giving similar or reduced CVs. The mean estimates across the three spline knot spacings of the 'none' approach are indicated as a dashed line for each modeling range. All three modeling ranges (column) and three spline knot spacings (within each column), and co-edited MM models (color-coded) are presented. Distributions are shown as half-violins (smoothed distribution), box plots with median, interquartile range, and 25th/75th quartile. The median lines of the box plots are connected to visualize trends within a specific baseline knot spacing. CVs are summarized as bar plots. Minimum/maximum CVs for each modeling range are indicated as downwards/upwards triangle in the color corresponding to the MM_{3co} model. Minimum/maximum CVs for each baseline knot spacing within a specific modeling range are reported on the right side of each column. Global minimum and maximum CVs across all models are added as text.

411

412 **Figure 3** shows distributions and coefficients of variation (CVs) of the GABA+ estimates for all
413 modeling strategies. Table 1 summarizes the mean and SD GABA/GABA+ estimates as well as
414 the statistics. GABA+ estimates are significantly higher than GABA-only estimates of the 'none'
415 modeling strategy for all modeling ranges and knot spacings, supporting the notion from **Figure**
416 **2** that not including an MM model leaves a considerable fraction of the edited 3-ppm signal un-
417 modeled, resulting in substantial residuals or increased baseline amplitude flexion.
418 Notably, all modeling strategies with MM_{3co} return comparable mean estimates and CVs within
419 the same knot spacing (see Minimum/Maximum column of **Figure 3**). In addition, sparser knot
420 spacing leads to lower CVs. The intermediate modeling range does not appear to perform more
421 consistently than both other modeling ranges.

422 *Table 1 – GABA+ mean and SDs for all modeling strategies (ratios to tCr). Significant differences ($p <$*
423 *.05) between the corresponding model and the 'none' are shaded in gray.*
424

modeling range		**narrow**			**intermediate**			**wide**		
knot spacing (ppm)		0.55	0.4	0.25	0.55	0.4	0.25	0.55	0.4	0.25
none	[GABA]	0.22 *.062*	0.213 *.071*	0.193 *.057*	0.297 *.048*	0.186 *.061*	0.204 *.044*	0.284 *.029*	0.288 *.034*	0.158 *.066*
GABA$_{hard}$	[GABA+]	0.328 *.041*	0.291 *.04*	0.312 *.046*	0.414 *.048*	0.296 *.062*	0.329 *.041*	0.389 *.041*	0.414 *.047*	0.3 *.068*
GABA$_{soft}$	[GABA+]	0.298 *.029*	0.267 *.036*	0.295 *.03*	0.417 *.053*	0.279 *.058*	0.323 *.031*	0.375 *.039*	0.397 *.042*	0.271 *.035*
Gauss$_{fixed}$	[GABA+]	0.306 *.04*	0.272 *.037*	0.304 *.034*	0.412 *.052*	0.298 *.055*	0.332 *.04*	0.369 *.05*	0.382 *.052*	0.343 *.04*
Gauss$_{free}$	[GABA+]	0.305 *.032*	0.277 *.031*	0.3 *.035*	0.409 *.052*	0.3 *.058*	0.331 *.04*	0.364 *.05*	0.382 *.045*	0.341 *.043*
MM09$_{hard}$	[GABA+]	-	-	-	-	-	-	0.396 *.039*	0.422 *.055*	0.345 *.084*
MM09$_{soft}$	[GABA+]	-	-	-	-	-	-	0.367 *.047*	0.386 *.045*	0.337 *.053*

Model evaluation

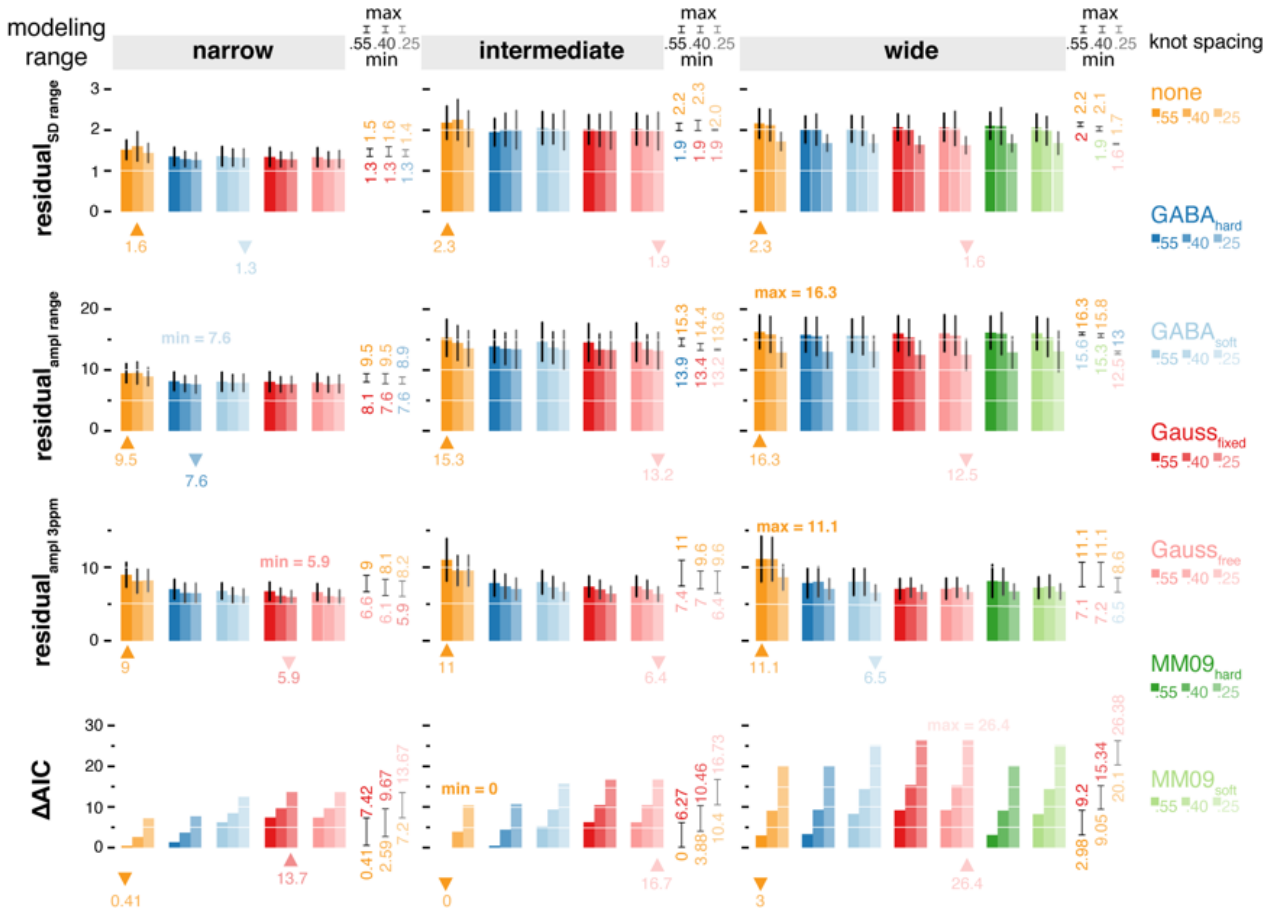


Figure 4 – Evaluation of all modeling strategies. Comparably low residual$_{ampl\ ange}$ are related to the high data quality without artifacts between 0.5 and 2 ppm. Including an MM$_{3co}$ model reduces the 3-ppm residual by ~30% without significant impact on the ΔAIC. All three modeling ranges (column) and three spline knot spacings (within each column), and co-edited MM models (color-coded) are presented. Bar plots represent mean values; SD is indicated by whiskers where appropriate. Minimum/maximum values for each modeling range are indicated as downwards/upwards triangle in the color corresponding to the MM$_{3co}$ model.

Minimum/maximum values for each baseline knot spacing within a specific modeling range are reported on the right side of each column. Global minimum and maximum values across all models are added as text.

Figure 4 summarizes the metrics used for model evaluation. The residual over the modeled frequency range ($residual_{SD\ range}$ and $residual_{ampl\ range}$) is lowest for the narrow modeling range. For the intermediate and wide modeling ranges, $residual_{ampl\ range}$ is substantially higher, largely driven by the 2-ppm region (see also **Figure 2**). Consequentially, $residual_{ampl\ range}$ is comparable between MM modeling strategies for a given knot spacing (see Minimum/Maximum column of **Figure 4**).

The residual around the GABA+ peak ($residual_{ampl\ 3ppm}$) is consistently reduced by up to 30% if a MM_{3co} model is included, in line with the reduction of structured residual in **Figure 2**. This effect is less pronounced for the dense knot spacing (0.25 ppm), indicating that a flexible baseline is to some degree capable of accounting for otherwise unmodeled MM signal. Together, these findings again support the notion that omitting an explicit MM_{3co} model does not capture the whole edited 3-ppm signal, which remains unmodeled (in the residual) or gets partially absorbed by the baseline or interpreted incorrectly as GABA signal.

The strategy with the lowest AIC is the 'none' model with the intermediate modeling range and sparse knot spacing, reflecting the low number of model parameters: there is no separate basis function for MM, and the low number of splines. The ΔAIC (the difference between the lowest AIC and the individual model's AIC) consequently increases for larger modeling ranges, as more splines are included. Similarly, ΔAIC increases for denser knot spacings, and in fact, this increase is much stronger compared to the resulting reduction in both residual measures, suggesting that the increased flexibility and reduction of the residual does not justify the greater number of model parameters.

For any given knot spacing and modeling range, ΔAIC values are comparable between MM_{3co} models, with moderate increases when more parameters are estimated. Together with its low CV (9.8% compared to the minimum CV value 9.7% for the $GABA_{soft}$ value with a narrow fit range) for GABA+, the ΔAIC for the $MM09_{hard}$ model over the wide modeling range with sparse knot spacing (ΔAIC = 3.1) indicates a good performance of this particular model without introducing

overfitting. Despite the slightly higher ΔAIC, it is beneficial to opt for the $MM09_{hard}$ model, since the $MM_{0.94}$ peak provides an 'external', non-overlapped reference anchor point for the amplitude of the expected MM_{30} peak – the MM landscape is thought to be relatively stable across healthy subjects in a narrow age range, at least in the absence of pathology [8]. Furthermore, the $MM09_{hard}$ model does not impose any amplitude assumptions or constraints on the target metabolite GABA.

Separation of GABA and MM_{3co}

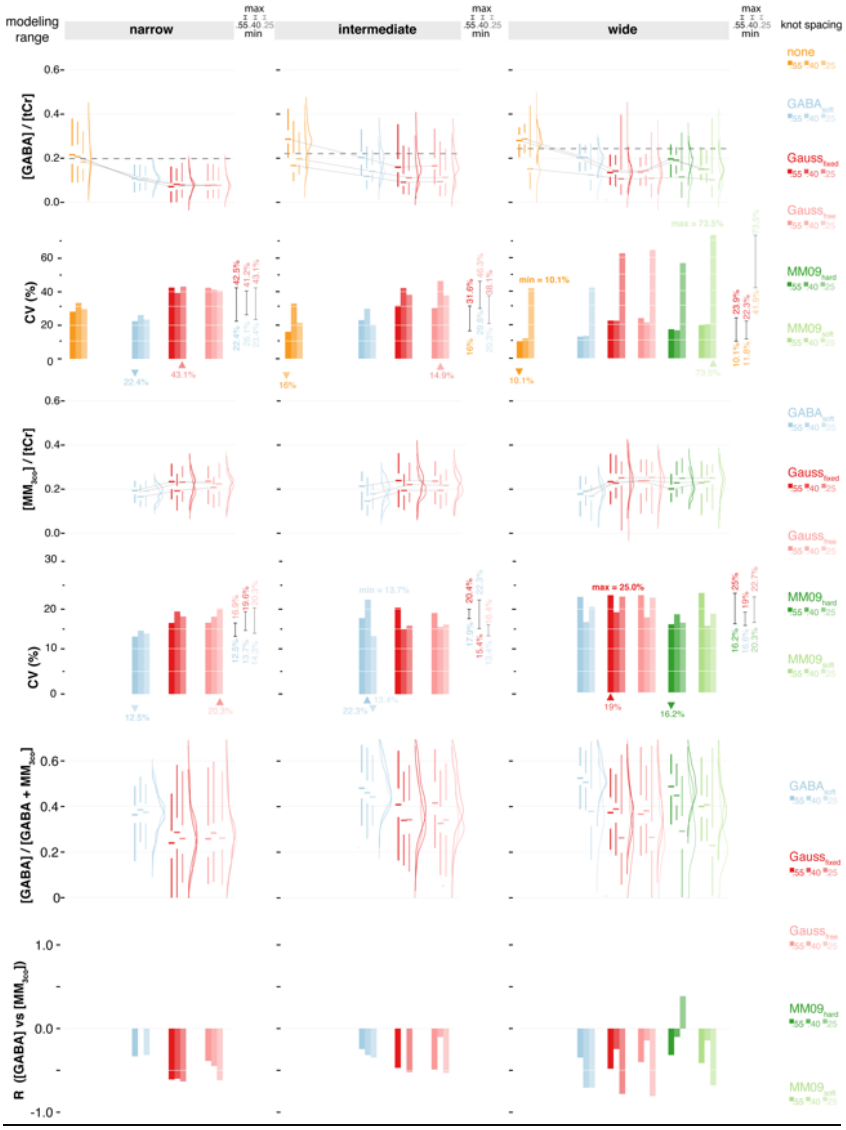
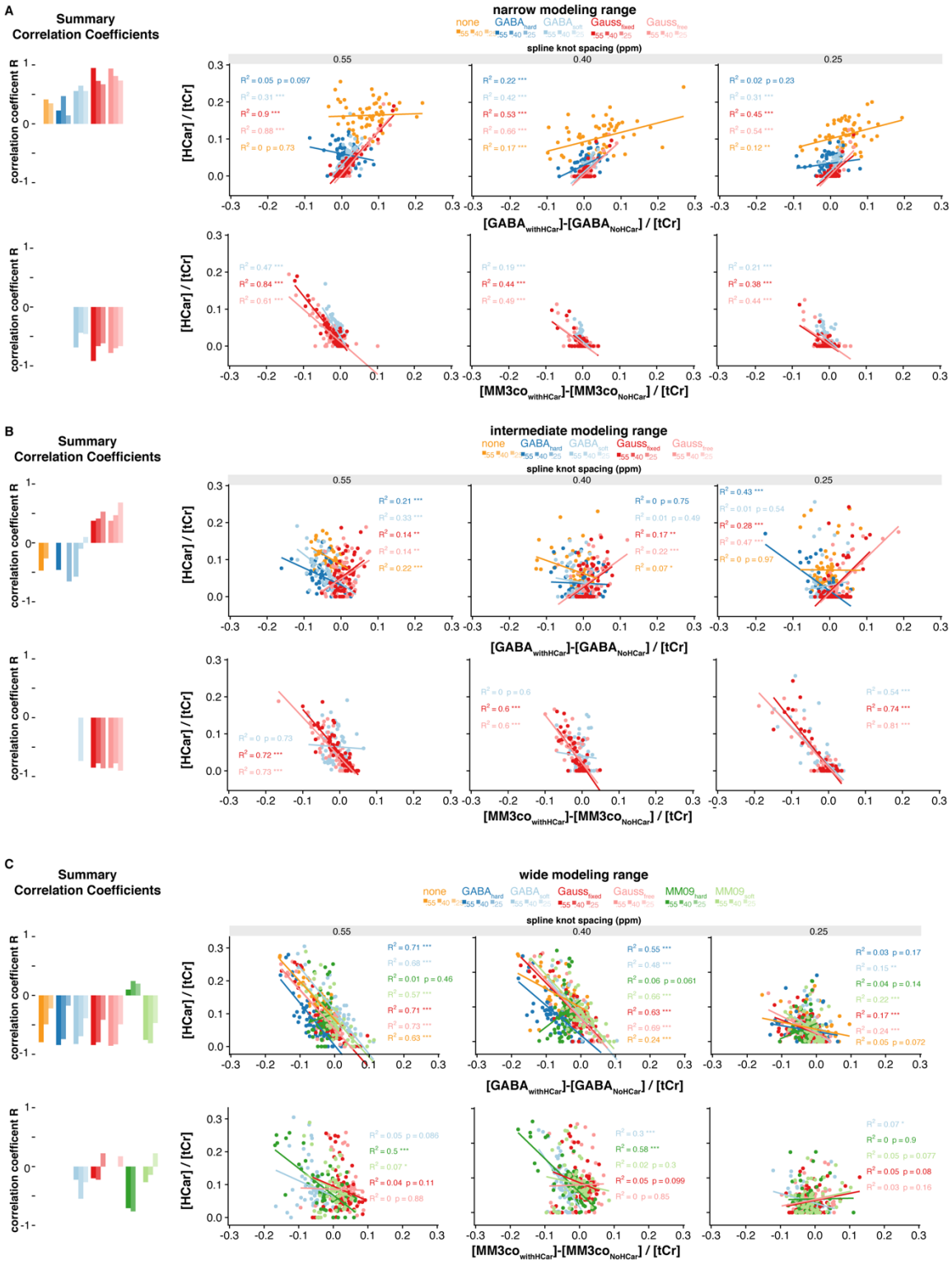


Figure 5 - Distribution of GABA and MM_{3co} estimates, relative contribution of GABA to GABA+ and Pearson's R between GABA and MM_{3co} for all modeling strategies. All three modeling ranges (column) and three spline knot spacings (within each column), and MM_{3co} models (color-coded) are presented. Distributions are shown as half-violins (smoothed distribution), box plots with median, interquartile range, and $25^{th}/75^{th}$ quartile. The median lines of the box plots are connected to visualize trends within a specific baseline knot spacing. The mean estimates across the three spline knot spacings of the 'none' approach are indicated as a dashed line for each modeling range. Across subject CVs are summarized as bar plots. Minimum/maximum CVs for each modeling range are indicated as downwards/upwards triangle in the color corresponding to the MM_{3co} model. Minimum/maximum CVs for each baseline knot spacing within a specific modeling range are reported on the right side of each column. Global minimum and maximum CVs across all models are added as text.

Figure 5 shows the distributions and CVs of the separate GABA and MM_{3co} estimates of all modeling strategies. Including a separate MM_{3co} basis function significantly decreases GABA estimates, suggesting that not doing so may lead to GABA overestimation, as MM signal is mistakenly modeled as GABA. As was seen for the composite GABA+ estimates in **Figure 3**, sparser knot spacing appears to stabilize modeling, leading to lower CVs of GABA. This becomes especially obvious for the wide modeling range, where GABA CVs exceed 50% for dense knot spacing.

MM_{3co} estimates are stable across the different knot spacings, suggesting that the different parametrizations accurately account for most of the co-edited MM signal at 3 ppm.

The GABA model, in combination with a wide modeling range and 0.55 ppm knot spacing, exhibits the lowest CV for GABA (10.4%). However, the $MM09_{hard}$ model in combination with the same knot spacing and modeling range has only slightly higher GABA CVs (17.3%) with the corresponding MM_{3co} CVs being 16.2% ($MM09_{hard}$). Again, despite slightly higher CV values it is beneficial to use a modeling strategy with a constraint to an 'external' reference peak ($MM09_{hard}$) instead of the highly overlapped MM_{3co} peak of $GABA_{soft}$ (CV = 12.8%), or entirely omitting MM_{3co}. Additionally, the correlation between the GABA and the MM_{3co} estimates is lower for the $MM09_{hard}$ model, potentially implying a better separation of GABA and MM_{3co}. However, a separation of GABA and co-edited macromolecules remains difficult with a low-to-moderate correlation between GABA and MM_{3co} estimates for all but one modeling strategy ($MM09_{hard}$ for the wide fit range and 0.4 ppm baseline knot spacing). Supplementary Material **5** reports the mean and SDs of the GABA and MM_{3co} estimates as well as the statistics. Significant differences between the mean or the SD compared to the corresponding model omitting co-edited MMs are indicated as shaded box.



494

Figure 6 – Impact of including homocarnosine in the basis set. The directionality of the correlation indicates that HCar absorbs GABA signal specifically for the intermediate and wide modeling range and absorbs MM_{3co} signal for all modeling ranges. Correlation analysis between the differences between GABA/ MM_{3co} estimates with and without HCar in the basis set and the

*HCar estimates. All three modeling ranges (A-C) and three spline knot spacings (within each subplot) were investigated. A summary bar plot with the correlation coefficient R is shown in the beginning of each row. Pearson's correlation was calculated for each MM_{3co} model (color-coded). Asterisks indicate significant correlations with $p < 0.05$ = *, $p < 0.01$ = **, and $p < 0.001$ = ***.*

Finally, **Figure 6** shows the impact of including HCar into the basis set with the difference in GABA and MM_{3co} estimates between the modeling strategies with and without HCar (ΔGABA and ΔMM_{3co}, respectively). Interestingly, clear differences in the systematic effects of HCar are evident between the modeling ranges:

For the narrow modeling range (**Figure 6 A**), HCar estimates correlate positively with ΔGABA, but the correlation is *only* substantial ($R > 0.25$) for strategies with a *separate* MM basis function. For precisely these strategies, HCar estimates correlate negatively with ΔMM_{3co}. These observations suggest that HCar is likely to account for MM_{3co} in the narrow modeling range. In contrast, HCar and ΔGABA correlate negatively for most strategies in the intermediate and wide modeling ranges (**Figure 6 B** and **C**). The negative correlations between HCar and ΔMM_{3co} are notably weaker for these modeling ranges, indicating that HCar is more likely to substitute for GABA signal instead of MM.

This behavior can possibly be explained by the HCar signal shape for each modeling range (**Supplementary Material 6**). For the narrow modeling range, the HCar basis function offers the model an additional degree of freedom to account for deviations of the actual edited 3-ppm signal from pure GABA and the symmetric Gaussian MM3co component, as no resonances below 2.78 ppm are considered. As a result, HCar shows a high correlation with the difference in MM_{3co}. For the intermediate and wide range, the HCar difference spectrum basis function more nearly resembles its GABA counterpart since other resonances are included, thereby more effectively coupling GABA and HCar estimates to each other. Perhaps unsurprisingly, HCar estimates are significantly higher for 'none' modeling strategy, and are substantially lower for more flexible baselines, supporting the notion that HCar rather serves as a substitute for an explicit MM signal, in particular if the baseline cannot absorb the latter (**Supplementary Material 6**). Within a given knot spacing and modeling range, HCar estimates are comparable between different MM_{3co} models, a behavior observed for GABA estimates as well.

The GABA+ plus homocarnosine estimates show a slight increase compared to the GABA+ estimates without HCar (**Supplementary Material 7**). For the 'none' model, stronger changes

522 occur as HCar accounts for MM signal (see also Figure 6). There was no improvement in the
523 CVs observed when including HCar in the model. The relative contribution of HCar to GABA+
524 ranged between 2.2% and 19.1% for modeling strategies with an MM_{3co} basis function and
525 between 18% and 36% for the 'none' model.
526

Discussion

The application of linear combination modeling to edited difference spectra is neither straightforward nor intuitive. The conceptual advantage of spectral editing arises from isolating a resolved target resonance, i.e. reducing the overlap of the target metabolite with other signals, as well as the number of signals in the spectrum in general[1]. LCM, on the other hand, benefits from maximizing the use of prior knowledge to solve the spectral modeling problem, i.e. using all available information for meaningful constraint, including from overlapping signals. The specific case of GABA-edited MRS at 3T poses unique and unresolved challenges. Firstly, a compromise must be drawn between maximizing the prior knowledge by increasing the modeling range and reducing the impact of co-edited and unwanted signals. Secondly, an appropriate parametrization of poorly characterized co-edited signals must be found, and possible interactions with the target metabolite GABA must be evaluated. Thirdly, effects of baseline modeling must be studied, again a consequence of the macromolecular background signal in the GABA-edited difference spectrum not being determined to this date. In this study, a total of 102 linear combination modeling strategies were compared for GABA-edited difference spectra, each with different modeling ranges, parametrizations of co-edited signals, and baseline model flexibility. The key findings are:

- Including a dedicated basis function for co-edited MM improves fit residuals, reduces CVs of GABA and GABA+ estimates, and avoids overestimation of GABA.
- Reducing the modeling range does not substantially stabilize or destabilize modeling, while removing potentially valuable information ($MM_{0.93}$ and 2-ppm NAA peak) from the optimization.
- Sparser baseline spline knot spacing leads, on average, to the lowest CV across all modeling ranges.

There is surprisingly little systematic investigation into linear combination modeling of GABA-edited difference spectra. To the best of our knowledge, there is only one conference abstract studying MM parametrization in GABA-edited MRS with the LCModel software[27]. The results from this preliminary investigation indicate that including a specific MM basis function

significantly reduces GABA estimates and were confirmed in the subsequent study [28], which is substantiated by our findings.

Although the substantial contribution of broad MM signals to the 3-ppm peak in the GABA-edited spectrum is widely known[1,35], it is rarely explicitly addressed in linear combination modeling. Instead, it is assumed that either an incomplete model (without explicit MM term) will still provide an accurate GABA estimate, or that baseline modeling will account for the MM signal. The current results provide evidence that including an appropriately parametrized MM model is a preferrable and easily implemented strategy, reducing the residual over the 3-ppm signal range by up to 30%, with similar or lower CVs for GABA+. In contrast, not including an MM model likely causes systematic overestimation of GABA, as the least-squares optimization attempts to minimize the model-data difference with an inadequate set of basis functions (only GABA), particularly when a rigid baseline is chosen. Including MM_{3co} is a justified and reasonable measure without overfitting (reflected by AIC), and stable mean estimates and CVs of MM_{3co} suggest an adequately parametrized model.

The different MM models in this study were based on certain assumptions, including the relative contribution of MM_{3co} to the 3-ppm GABA peak to be around 50%[1,6,8,19]. Levels of $MM_{0.93}$ have been found to be stable across the whole brain[36] and are thought to be stable across healthy subjects. Under these assumptions, the $MM09_{hard}$ model with a rigid amplitude coupling between MM_{3co} and the non-overlapped $MM_{0.93}$ peak is a suitable strategy, supported by favorable CVs and ΔAIC. Further studies need to be performed to investigate the distribution and correlation between $MM_{0.93}$ and MM_{3co} in the brain.

Unedited MRSI data measured at 7T indicates significant differences between white and gray matter for several macromolecules in the healthy brain[36]. Changes in the MM concentrations during disease may also affect the relative contribution to the 3-ppm peak, and therefore render models with prior amplitude assumptions inaccurate. If there is reason to expect strong fluctuations of MM_{3co}, a modeling strategy with fewer assumptions about amplitude ratios between the metabolite of interest GABA or the $MM_{0.93}$ signal and the MM_{3co} signal is preferable to the $MM09_{hard}$ strategy. Here, the $Gauss_{free}$ and $Gauss_{fixed}$ strategies could be used to account for

587 changes in the MM_{3co} contribution more freely, as their mean estimates of GABA and GABA+
588 were in good agreement with the more constrained approaches, although they led to increased
589 CVs and ΔAICs. In addition, the less-constrained models might be more appropriate for investi-
590 gating changes in MM_{3co} due to age[37] or disease, or for exploring frequency-drift-related effects
591 on the co-edited MM signal[1,8,38]. Another potential way to model the co-edited MM signal is to
592 include lysine in the simulated basis set, as it has been identified as the potential source of the
593 signal[6], although this approach would require appropriate broadening and incorporation of chem-
594 ical shift and coupling values from protein databases[39].
595
596 Overall, results did not differ drastically between modeling ranges, although it is noteworthy that
597 the effects of baseline flexibility were less pronounced for the narrow modeling range, likely be-
598 cause the complex interaction of the overlapping 2.25 ppm GABA and Glx signals with the un-
599 derlying baselines is omitted. Furthermore, there was no evidence that the intermediate modeling
600 range, which is proposed in the LCModel manual[18] to avoid frequently occurring co-edited lipid
601 signals, improved quantification substantially compared to both other modeling ranges, although
602 it should be mentioned that this particular dataset did not suffer from severe lipid contamination.
603 Taken together, the choice of modeling range does not impact quantitative results as substantially
604 as the inclusion of an MM model.
605
606 Baseline models are included in most LCM algorithms to account for signals not otherwise mod-
607 eled, e.g. residual water tails or unparametrized macromolecules and lipids. Compared to con-
608 ventional short-TE spectra, water and non-co-edited MMs are removed upon subtraction in the
609 GABA-edited spectrum, which is therefore frequently modeled with a stiffer baseline[4,18]. Our re-
610 sults show that sparser knot spacing (0.55 ppm) leads to lower CVs in metabolite estimates. A
611 more flexible baseline (0.25 ppm) improves local and global residuals, but not enough to justify
612 the additional model parameters (as per the AICs). More importantly, an overly flexible baseline
613 may absorb edited signal, although it appeared that it did not do so excessively even for the 0.25-
614 ppm strategies. The exception was the ‘none’ model, where the baseline was the only available
615 part of the model to take up signal, underlining the inadequacy of the default LCModel approach.
616 Taken together, a relatively rigid baseline with a parametrized MM basis function is preferable
617 for LCM of GABA-edited spectra. A caveat to this recommendation is the observation of

618 structural baseline fluctuations underneath the 2.25 ppm signals from GABA, Glx, GSH, NAA
619 and NAAG, particularly for the 0.25 ppm knot spacing and a relatively broad increase in the
620 baseline between 2.7 and 3 ppm. These were observed previously[27], and are likely signals from
621 un-parametrized MMs directly and indirectly affected by the editing pulse. Rigid baselines may
622 force a wrong metabolite model in that region and interfere with accurate estimation of GABA
623 and Glx. In fact, the structural Glx residual at 3.75 ppm suggests a systematic misestimation of
624 the Glx phase, likely driven by the 2.25 ppm signals. While beyond the scope of this investiga-
625 tion, it is conceivable that more informed parametrization (or, ideally, direct measurement) of
626 this unexplored MM background may benefit the modeling of the entire difference spectrum. Al-
627 ternatively, hitherto unexplored approaches with variable baseline knot spacing may be worth
628 investigating.
629
630 The HCar molecule has a GABA moiety with similar chemical shifts and is therefore co-edited.
631 Evidence regarding in-vivo HCar levels in the human brain is inconclusive – early work deter-
632 mined HCar levels to be 0.5 mM[5] (compared to ~1 mM for GABA), while a recent hybrid up-
633 field/downfield inversion-recovery method determined the HCar/GABA ratio as 17%[40]. There-
634 fore, we tested the impact of adding HCar to the basis set without additional constraints. Includ-
635 ing HCar systematically affected GABA and MM_{3co} estimates, in a way that strongly depended
636 on the choice of modeling range. HCar estimates themselves ranged from 2.2% to 19.1% of the
637 GABA+ signal, depending strongly on the degree of baseline flexibility. The results suggest that
638 the overlap between the three model terms (HCar, GABA, MM_{3co}) is too substantial for reliable
639 three-way separation, particularly in the presence of a highly flexible baseline. A minor increase
640 in "GABA+ plus HCar" estimates compared to GABA+ estimates was observed and the inclu-
641 sion of HCar did not substantially improve the CVs. Additionally, the disagreement between the
642 model and the data at 2.9 ppm indicates that a simple unconstrained addition of HCar to the mod-
643 eling is not justified.
644
645 Symmetric GABA-editing (edit-ON frequency at 1.9 ppm and edit-OFF frequency at 1.7 ppm) is
646 commonly used eliminate the MM_{3co} contamination of the 3-ppm GABA+ signal. In practice, B_0
647 instabilities lead to residual MM_{3co} components with variable polarity [11]. The $Gauss_{free}$ and
648 $Gauss_{fixed}$ MM_{3co} models could potentially be used to account for those variable MM_{3co}

649 contributions in those spectra. Modeling of those spectra with the current strategies that do have
650 a non-negative model component as constraint would be challenging. Those modeling strategies
651 could potentially be adapted by using the B_0 history during the experiment [38] to predict the polar-
652 ity and relative amplitude of the MM_{3co} signal, and include those as a soft constraint relative to
653 the MM_{09} signal ($MM09_{hard}$ or $MM09_{soft}$) or the GABA signal ($GABA_{hard}$ or $GABA_{soft}$).

654 Limitations

655 A limitation of this study is the high spectral quality (SNR, linewidth, no apparent subtraction
656 artefacts, or lipid contaminations) of the dataset analyzed. We did not investigate model para-
657 metrizations of movement or drift, which may introduce systematic changes to the co-edited MM
658 signal. While our results suggest that using the wide modeling range with a rigid baseline is ben-
659 eficial, strong co-edited lipid signals are likely to not be modeled appropriately, and the interme-
660 diate modeling range may be more suitable. Further studies of the possible impact of changes in
661 spectral quality need to be performed to validate the modeling strategies under suboptimal condi-
662 tions.
663
664 Another limitation is that there is no 'gold standard' of metabolite level estimation in GABA-ed-
665 ited MRS to validate the results against. The performance of different algorithms or in this study
666 modeling strategy is often judged by the level of variance [26]. A lower variance does, of course,
667 not necessarily reflect greater modeling accuracy, but under the assumption that the homogene-
668 ous study population and data acquisition contribute comparably little biological and instrumen-
669 tal variance, CVs will predominantly reflect variance introduced by the modeling approach. Re-
670 cently, the field is witnessing increasing efforts to generate simulated spectra with known ground
671 truth as a gold standard, although these approaches can only be successful to the extent that those
672 spectra are truly representative of in-vivo data. Further, such gold standard studies with a known
673 ground truth could be used to validate whether a correct separation of GABA and MM_{3co} is
674 achievable by advanced LCM. This study indicates a low-to-moderate correlation between the
675 GABA and MM_{3co} estimates, suggesting that the two components are not reliably separated.
676 However, some of the modeling strategies appeared to have a lower association between both es-
677 timates and could possibly be validated further on a synthetic dataset with known GABA and
678 MM_{3co} concentrations.

AIC as a measure of the goodness of fit can be used for linear and non-linear approaches if the log-likelihood is obtained similarly. However, there are two potential limitations for the for the application of AIC in this study. First, for linear-combination modeling of MRS data, as implemented in Osprey, a non-linear optimization is followed by a linear optimization during each iteration. Parameters are treated equally in the calculation of the AIC regardless of whether they are non-linear (e.g., a phase parameter) or linear (an amplitude parameter). Second, the AIC penalizes complex models, but does not measure effects of soft constraints and is likely to prefer models without a soft constraint as those should have a reduced likelihood [41]. Here, we introduced a rather arbitrary correction term of 0.5 per soft constraint for those models to reduce this effect. Therefore, the resulting ΔAIC values in this study should be interpreted with care and considered as only one among several metrics to evaluate model performance.

Conclusion

This study proposed and compared different modeling strategies for LCM of GABA+-edited difference spectra from a multi-site MEGA-PRESS dataset. Introducing a parametrized model for co-edited macromolecules reduces fit residuals, while maintaining low coefficients of variation of GABA+ estimates. A rigid baseline was found to be beneficial, while using a narrower modeling range did not significantly improve the modeling. The overall modeling results suggest that GABA-edited data are reliably modeled with an adequately parametrized MM_{3co} model, constrained by the non-overlapped 0.93-ppm MM resonance, in combination with a full modeling range and sparse knot spacing. Incorporating homocarnosine into the modeling did not significantly improve the GABA+ estimates and did not allow for a stable separation of GABA and HCar.

References

1. Choi I-Y, Andronesi OC, Barker P, et al. Spectral editing in 1H magnetic resonance spectroscopy: Experts' consensus recommendations. *NMR Biomed*. 2021;n/a(n/a):e4411. doi:https://doi.org/10.1002/nbm.4411

2. Edden RAE, Puts NAJ, Harris AD, Barker PB, Evans CJ. Gannet: A batch-processing tool for the quantitative analysis of gamma-aminobutyric acid-edited MR spectroscopy spectra. *J Magn Reson Imaging*. 2014;40(6):1445-1452. doi:10.1002/jmri.24478

3. Wilson M, Reynolds G, Kauppinen RA, Arvanitis TN, Peet AC. A constrained least-squares approach to the automated quantitation of in vivo 1H magnetic resonance spectroscopy data. *Magn Reson Med*. 2011;65(1):1-12.

4. Provencher SW. Automatic quantitation of localized in vivo1H spectra with LCModel. *NMR Biomed*. 2001;14(4):260-264. doi:10.1002/nbm.698

5. Petroff OAC, Hyder F, Rothman DL, Mattson RH. Topiramate Rapidly Raises Brain GABA in Epilepsy Patients. *Epilepsia*. 2001;42(4):543-548. doi:https://doi.org/10.1046/j.1528-1157.2001.18800.x

6. Deelchand DK, Marjańska M, Henry P-G, Terpstra M. MEGA-PRESS of GABA+: Influences of acquisition parameters. *NMR Biomed*. 2019;n/a(n/a):e4199. doi:https://doi.org/10.1002/nbm.4199

7. Marjańska M, Terpstra M. Influence of fitting approaches in LCModel on MRS quantification focusing on age-specific macromolecules and the spline baseline. *NMR Biomed*. November 2019. doi:10.1002/nbm.4197

8. Shungu DC, Mao X, Gonzales R, et al. Brain γ-aminobutyric acid (GABA) detection in vivo with the J-editing (1) H MRS technique: a comprehensive methodological evaluation of sensitivity enhancement, macromolecule contamination and test-retest reliability. *NMR Biomed*. 2016;29(7):932-942. doi:10.1002/nbm.3539

9. Bhagwagar Z, Wylezinska M, Jezzard P, et al. Reduction in Occipital Cortex γ-Aminobutyric Acid Concentrations in Medication-Free Recovered Unipolar Depressed and Bipolar Subjects. *Biol Psychiatry*. 2007;61(6):806-812. doi:10.1016/j.biopsych.2006.08.048

10. Mikkelsen M, Barker PB, Bhattacharyya PK, et al. Big GABA: Edited MR spectroscopy at 24 research sites. *NeuroImage*. 2017;159:32-45. doi:10.1016/j.neuroimage.2017.07.021

11. Edden RAE, Oeltzschner G, Harris AD, et al. Prospective frequency correction for macromolecule-suppressed GABA editing at 3T. *J Magn Reson Imaging JMRI*. 2016;44(6):1474-1482. doi:10.1002/jmri.25304

12. Oeltzschner G, Zöllner HJ, Hui SCN, et al. Osprey: Open-source processing, reconstruction & estimation of magnetic resonance spectroscopy data. *J Neurosci Methods*. 2020;343:108827. doi:10.1016/j.jneumeth.2020.108827

13. Osprey GitHub repository. Osprey GitHub repository. https://github.com/schorschinho/osprey. Published 2020. Accessed May 27, 2020.

14. Klose U. In vivo proton spectroscopy in presence of eddy currents. *Magn Reson Med*. 1990;14(1):26-30. doi:10.1002/mrm.1910140104

15. Mikkelsen M, Tapper S, Near J, Mostofsky SH, Puts NAJ, Edden RAE. Correcting frequency and phase offsets in MRS data using robust spectral registration. *NMR Biomed*. July 2020:e4368. doi:10.1002/nbm.4368

16. Barkhuijsen H, de Beer R, van Ormondt D. Improved algorithm for noniterative time-domain model fitting to exponentially damped magnetic resonance signals. *J Magn Reson 1969*. 1987;73(3):553-557. doi:10.1016/0022-2364(87)90023-0

17. Simpson R, Devenyi GA, Jezzard P, Hennessy TJ, Near J. Advanced processing and simulation of MRS data using the FID appliance (FID-A)—An open source, MATLAB-based toolkit. *Magn Reson Med*. 2017;77(1):23-33. doi:10.1002/mrm.26091

18. Provencher S. LCModel & LCMgui User's Manual. LCModel & LCMgui User's Manual. http://s-provencher.com/pub/LCModel/manual/manual.pdf. Published 2020. Accessed April 28, 2020.

19. Henry PG, Dautry C, Hantraye P, Bloch G. Brain gaba editing without macromolecule contamination. *Magn Reson Med*. 2001;45(3):517-520. doi:10.1002/1522-2594(200103)45:3<517::AID-MRM1068>3.0.CO;2-6

20. Levenberg K. A method for the solution of certain non-linear problems in least squares. *Q Appl Math*. 1944;2(2):164-168. doi:10.1090/qam/10666

21. Marquardt DW. An Algorithm for Least-Squares Estimation of Nonlinear Parameters. *J Soc Ind Appl Math*. 1963;11(2):431-441. doi:10.1137/0111030

22. Becker S. LBFGSB (L-BFGS-B) mex wrapper - File Exchange - MATLAB Central. LBFGSB (L-BFGS-B) mex wrapper - File Exchange - MATLAB Central. https://www.mathworks.com/matlabcentral/fileexchange/35104-lbfgsb-l-bfgs-b-mex-wrapper. Published February 23, 2015. Accessed March 3, 2021.

23. Byrd RH, Lu P, Nocedal J, Zhu C. A Limited Memory Algorithm for Bound Constrained Optimization. *SIAM J Sci Comput*. 1995;16(5):1190-1208. doi:10.1137/0916069

24. Zhu C, Byrd RH, Lu P, Nocedal J. Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization. *ACM Trans Math Softw*. 1997;23(4):550-560. doi:10.1145/279232.279236

25. Wilson M, Andronesi O, Barker PB, et al. *Methodological Consensus on Clinical Proton MRS of the Brain: Review and Recommendations*. Vol 82. John Wiley and Sons Inc.; 2019. doi:10.1002/mrm.27742

773 26. Zöllner HJ, Považan M, Hui SCN, Tapper S, Edden RAE, Oeltzschner G. Comparison of
774 different linear-combination modeling algorithms for short-TE proton spectra. *NMR Bio-*
775 *med*. 2021;n/a(n/a):e4482. doi:https://doi.org/10.1002/nbm.4482

776 27. Murdoch JB, Dydak U. Modeling MEGA-PRESS macromolecules for a better grasp of
777 GABA. In: *19th Annual Meeting of the International Society for Magnetic Resonance in*
778 *Medicine (ISMRM)*. ; 2011. https://scholar.google.com/scholar_lookup?title=Model-
779 ing%20MEGA-PRESS%20macromolecules%20for%20a%20bet-
780 ter%20grasp%20of%20GABA&publication_year=2011&author=J.B.%20Murdoch&au-
781 thor=U.%20Dydak. Accessed July 1, 2020.

782 28. Dydak U, Jiang YM, Long LL, et al. In vivo measurement of brain GABA concentrations
783 by magnetic resonance spectroscopy in smelters occupationally exposed to manganese. *En-*
784 *viron Health Perspect*. 2011;119(2):219-224. doi:10.1289/ehp.1002192

785 29. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria:
786 R Foundation for Statistical Computing; 2017. https://www.R-project.org/.

787 30. SpecVis GitHub repository. SpecVis GitHub repository.
788 https://github.com/hezoe100/SpecVis. Published 2020. Accessed May 27, 2020.

789 31. Wickham H. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York;
790 2009. http://ggplot2.org.

791 32. Zöllner HJ. Comparison of linear-combination modeling strategies for GABA-edited MRS
792 at 3T. https://osf.io/aqm8f/. Published April 30, 2021. Accessed April 30, 2021.

793 33. Akaike H. A new look at the statistical model identification. *IEEE Trans Autom Control*.
794 1974;19(6):716-723. doi:10.1109/TAC.1974.1100705

795 34. Mikkelsen M, Rimbault DL, Barker PB, et al. Big GABA II: Water-referenced edited MR
796 spectroscopy at 25 research sites. *NeuroImage*. 2019;191:537-548. doi:10.1016/J.NEU-
797 ROIMAGE.2019.02.059

798 35. Cudalbu C, Behar KL, Bhattacharyya PK, et al. Contribution of macromolecules to brain
799 1H MR spectra: Experts' consensus recommendations. *NMR Biomed Revis*. 2020.

800 36. Považan M, Strasser B, Hangel G, et al. Simultaneous mapping of metabolites and individ-
801 ual macromolecular components via ultra-short acquisition delay 1H MRSI in the brain at
802 7T. *Magn Reson Med*. 2018;79(3):1231-1240. doi:10.1002/mrm.26778

803 37. Marjańska M, Deelchand DK, Hodges JS, et al. Altered macromolecular pattern and con-
804 tent in the aging human brain. *NMR Biomed*. 2018;31(2):e3865. doi:10.1002/nbm.3865

805 38. Veen JW van der, Marenco S, Berman KF, Shen J. Retrospective correction of frequency
806 drift in spectral editing: The GABA editing example. *NMR Biomed*. 2017;30(8):e3725.
807 doi:https://doi.org/10.1002/nbm.3725

808 39. Borbath T, Manohar SM, Henning A. Towards a Fitting Model of Macromolecular Spectra:
809 Amino Acids. In: *27th Annual Meeting of the International Society for Magnetic Resonance*
810 *in Medicine (ISMRM)*. Montreal, Canada; 2019.

811 40. Landheer K, Prinsen H, Petroff OA, Rothman DL, Juchem C. Elevated homocarnosine and
812 GABA in subject on isoniazid as assessed through 1H MRS at 7T. *Anal Biochem*.
813 2020;599:113738. doi:10.1016/j.ab.2020.113738

814 41. Endres DM, Chiovetto E, Giese M. Model selection for the extraction of movement primi-
815 tives. *Front Comput Neurosci*. 2013;7. doi:10.3389/fncom.2013.00185

816

Supplementary Material

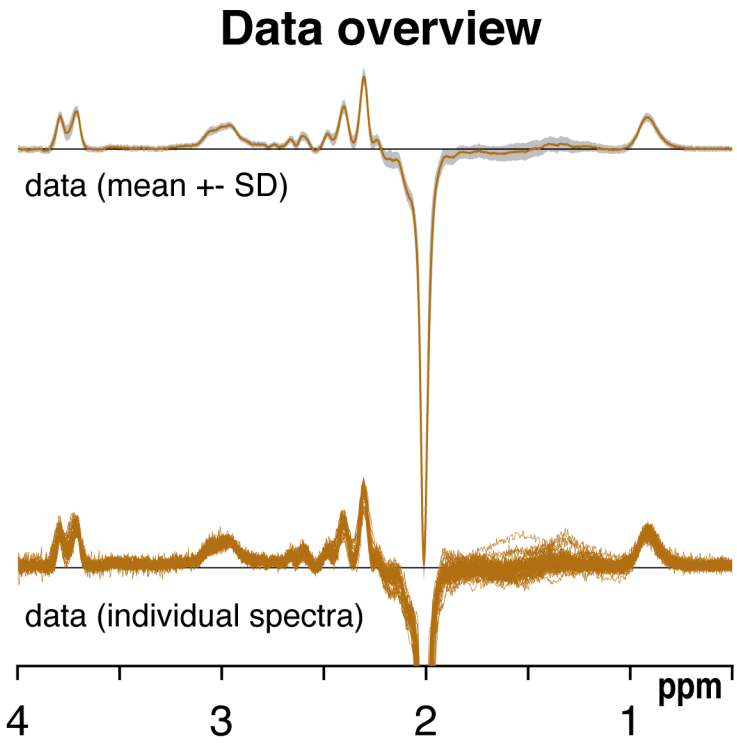
Table of Contents

Supplementary Material 1 – List of included subjects. All datasets are available at https://www.nitrc.org/projects/biggaba/

site	**subjects**	**Σ**
P01	S01,S03,S04,S05,S08	5
P03	S02,S03,S04,S07,S08,S09,S10,S11,S12	9
P05	S01,S02,S03,S05,S06,S07,S08	7
P06	S01,S02,S03,S04,S05,S06,S07,S08,S09	9
P07	S02,S03,S04,S09,S10,S11,S12	7
P08	S01,S02,S03,S04,S05,S06,S07,S08,S09,S10,S11,S12	12
P09	S01,S02,S03,S04,S05,S06,S07,S08,S09,S10,S11,S12	12
Σ = 7		**Σ** = 61

Supplementary Material 2. *Properties of the Gaussian functions of the broad macromolecule and lipid resonances included in the basis sets, taken from section 11.7 of the LCModel manual. The amplitude values are scaled relative to the CH_3 singlet of creatine with amplitude 3.*

Name	*Frequencies [ppm]*	*FWHM [ppm]*	*Amplitude*
edit-OFF spectrum basis set			
$MM_{0.94}$	0.915	0.14	3.00
$MM_{1.22}$	1.22	0.15	2.00
$MM_{1.43}$	1.43	0.17	2.00
$MM_{1.70}$	1.67	0.15	0.20
$MM_{2.05}$	2.08	0.15	1.33
	2.25	0.20	0.33
	1.95	0.15	0.33
	3.00	0.20	0.40
Lip09	0.89	0.14	3.00
Lip13	1.28	0.15	2.00
	1.28	0.089	2.00
Lip20	2.04	0.15	1.33
	2.25	0.15	0.67
	2.80	0.20	0.87
Difference spectrum basis set			
$MM_{0.94}$	0.915	0.14	3
MM_{3co}	3	**14 Hz**	2



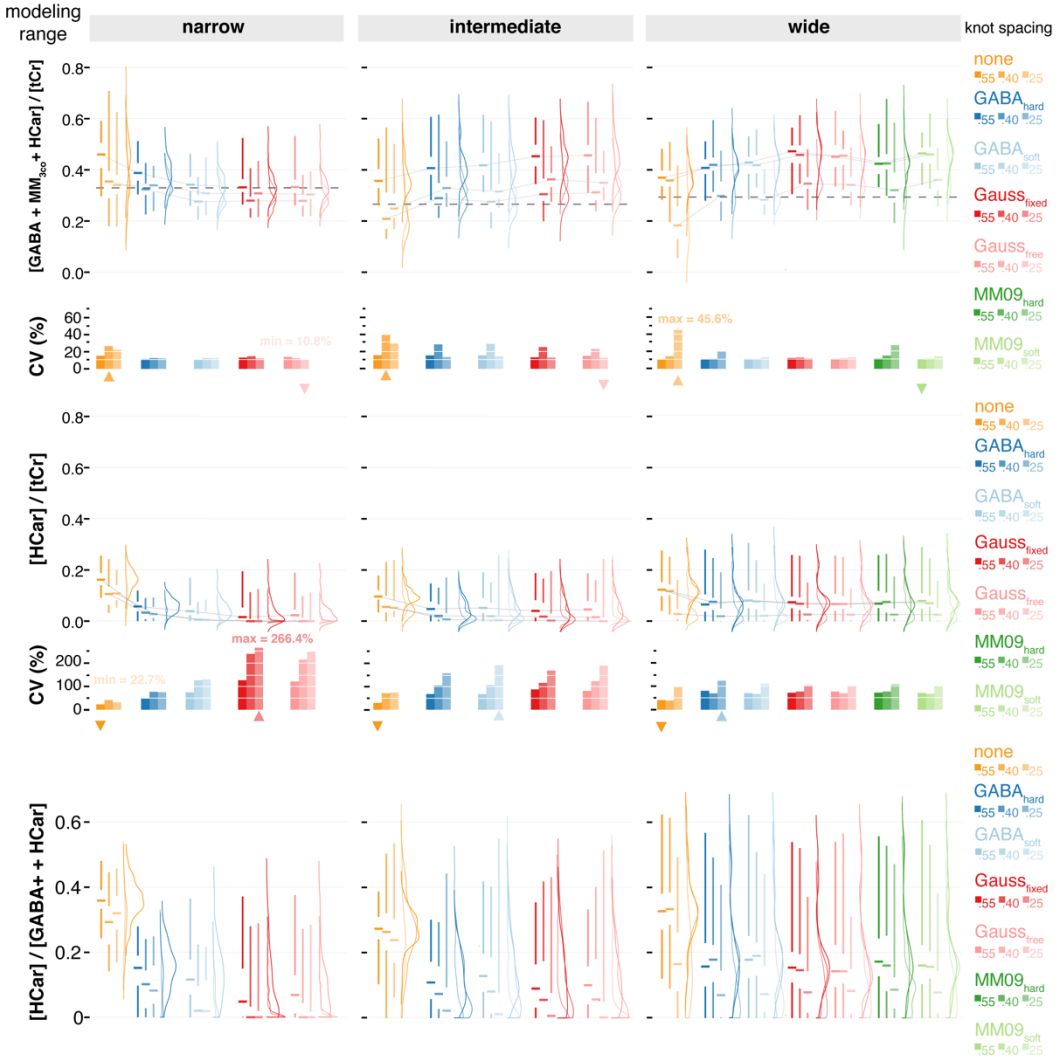
Supplementary Material 4 – Overview of the processed data including the mean ± SD and individual data.

Supplementary Material 5 – GABA and MM_{3co} mean and SDs for all modeling strategies (ratios to tCr). Significant differences ($p < .05$) between the corresponding model and the 'none' are shaded in gray. Differences in MM_{3co} are compared between the corresponding model and the $GABA_{soft}$ model and significant differences are also shaded in gray .

modeling range		**narrow**			**intermediate**			**wide**		
knot spacing (ppm)		0.55	0.4	0.25	0.55	0.4	0.25	0.55	0.4	0.25
none	[GABA]	0.22 *.062*	0.213 *.071*	0.193 *.057*	0.297 *.048*	0.186 *.061*	0.204 *.044*	0.284 *.029*	0.288 *.034*	0.158 *.066*
GABA$_{hard}$	[GABA]	0.164 *.021*	0.146 *.02*	0.156 *.023*	0.207 *.024*	0.148 *.031*	0.165 *.021*	0.194 *.021*	0.207 *.023*	0.15 *.034*
	[MM$_{3co}$]	-	-	-	-	-	-	-	-	-
GABA$_{soft}$	[GABA]	0.108 *.024*	0.102 *.027*	0.108 *.025*	0.207 *.048*	0.13 *.039*	0.148 *.03*	0.203 *.026*	0.204 *.027*	0.115 *.049*
	[MM$_{3co}$]	0.19 *.026*	0.165 *.025*	0.186 *.027*	0.21 *.038*	0.148 *.033*	0.175 *.024*	0.172 *.039*	0.194 *.032*	0.156 *.032*
Gauss$_{fixed}$	[GABA]	0.074 *.031*	0.08 *.031*	0.081 *.035*	0.174 *.055*	0.106 *.045*	0.118 *.045*	0.136 *.031*	0.145 *.032*	0.108 *.068*
	[MM$_{3co}$]	0.233 *.039*	0.192 *.038*	0.223 *.041*	0.237 *.048*	0.192 *.03*	0.215 *.035*	0.233 *.054*	0.237 *.045*	0.235 *.053*
Gauss$_{free}$	[GABA]	0.079 *.034*	0.078 *.032*	0.081 *.033*	0.18 *.054*	0.107 *.05*	0.119 *.045*	0.137 *.033*	0.145 *.031*	0.107 *.069*
	[MM$_{3co}$]	0.226 *.038*	0.198 *.036*	0.22 *.045*	0.229 *.044*	0.193 *.031*	0.213 *.035*	0.227 *.057*	0.237 *.042*	0.234 *.053*
MM09$_{hard}$	[GABA]	-	-	-	-	-	-	0.195 *.034*	0.189 *.032*	0.106 *.06*
	[MM$_{3co}$]	-	-	-	-	-	-	0.201 *.032*	0.232 *.043*	0.24 *.04*
MM09$_{soft}$	[GABA]	-	-	-	-	-	-	0.152 *.03*	0.154 *.031*	0.098 *.072*
	[MM$_{3co}$]	-	-	-	-	-	-	0.216 *.051*	0.232 *.037*	0.24 *.045*

Supplementary Material 6 – Mean modeling results and homocarnosine estimates for all modeling strategies with homocarnosine. A substantial structured residual is visible at 3 ppm if for all modeling strategies and for the narrow and intermediate modeling range the homocarnosine concentrations are significantly lower compared to omitting the co-edited MM, especially for knot spacings <= 0.4 ppm . All

three modeling ranges (columns), three spline knot spacings (rows), and MM_{3co} model (color-coded) are presented with mean residuals and fits, as well as the GABA, MM_{3co}, homocarnosine (HCar) and spline baseline models. The mean data is included in black. The arrows indicate the range of values for a specific modeling range and spline knot spacing with the color corresponding to the MM_{3co} model with minimum/maximum value.



Supplementary Material 7 - Distribution of GABA+ plus HCar and HCar estimates and the relative contribution of HCar to GABA+ plus HCar for all modeling strategies. The mean estimates of GABA+ plus HCar across the three spline knot spacings of the 'none' approach are indicated as an dashed line for each modeling range. All three modeling ranges (column) and three spline knot spacings (within each column), and MM_{3co} models (color-coded) are presented. Distributions are shown as half-violins

(smoothed distribution), box plots with median, interquartile range, and $25^{th}/75^{th}$ quartile. The median lines of the box plots are connected to visualize trends within a specific baseline knot spacing. CVs are summarized as bar plots. Minimum/maximum CVs for each spline knot spacing are indicated as downwards/upwards triangle in the color corresponding to the MM_{3co} model. Global minimum and maximum CVs across all models are added as text.

Declaration of competing interests

The authors have nothing to declare.

Acknowledgement

This work is supported by NIH grants P41 EB031771, R01 EB016089, R01 EB023963, R01 EB028259, R21 AG060245, and K99/R00 AG062230.

CRediT authorship contribution statement

Helge J. Zöllner: Software, Formal Analysis, Investigation, Writing – Original Draft, Writing – Review & Editing, Visualization. **Sofie Tapper**: Investigation, Writing – Review & Editing. **Steve C. N. Hui**: Investigation, Writing – Review & Editing. **Richard A. E. Edden**: Conceptualization, Formal Analysis, Writing – Review & Editing, Supervision, Project administration, Funding acquisition. **Peter B. Barker**: Writing – Review & Editing, Supervision, Funding acquisition. **Georg Oeltzschner**: Conceptualization, Methodology, Software, Investigation, Formal Analysis, Writing – Review & Editing, Supervision, Funding acquisition.