Comparison of MRS acquisition methods for separation of overlapping signals at 3 T

Journal of Neuroscience Methods 423 (2025) 110523

Contents lists available at ScienceDirect

Journal of Neuroscience Methods

journal homepage: www.elsevier.com/locate/jneumeth

Comparison of MRS acquisition methods for separation of overlapping signals at 3 T

Tiffany K. Bell [a,b,c,g,*], Dana Goerzen [d], Jamie Near [e,f], Ashley D. Harris [a,b,c]

[a] Deparment of Radiology, University of Calgary, Calgary, Canada
[b] Hotchkiss Brain Institute. Calgary, Canada
[c] Alberta Children's Hospital Research Institute, Calgary, Canada
[d] Weill Cornell Medicine, Cornell University
[e] Physical Studies Research Platform, Sunnybrook Research Institute, Toronto, Canada
[f] Department of Medical Biophysics, University of Toronto, Toronto, Canada
[g] School of Psychology and Clinical Language Sciences, University of Reading, Reading, UK

ARTICLE INFO

Keywords:
Magnetic resonance spectroscopy
7 T
High-field
STEAM
Semi-laser
PRESS

ABSTRACT

Background: Proton magnetic resonance spectroscopy (MRS) can be used to quantify multiple neurometabolites. However, due to the difficulty of separating overlapping signals at the commonly used field strength of 3 T, the quantified values are often composites of metabolically related chemicals. This can complicate interpretation and mask effects of interest. Therefore, it is important to determine the ability to accurately separate these signals at 3 T. Data acquired at 7 T can provide a benchmark, as higher field strength facilitates spectral resolution and reduces the signal overlap.

New Methods: This study assessed the ability of multiple 3 T MRS sequences to separate the commonly acquired neurochemicals (Glutamate (Glu) and Glutamine (Gln); N-Acetyl aspartate (NAA) and N-acetylaspartylglutamate (NAAG); Creatine (Cr) and Phosphocreatine (PCr); Choline (Cho), Phosphocholine (PC) and Glycerophosphocholine (GPC)). We compared metabolites quantified at 3 T from 6 acquisitions (PRESS, TE= 20, 30, 40, 80 ms, semi-LASER, TE=28 ms and STEAM TE=6 ms) with those quantified at 7 T using STEAM (TE=8 ms).

Results: Sequences with short echo times (STEAM-6, PRESS-20) generally performed better at separating most metabolites when using correlative and difference analyses with 7 T reference data. The exceptions were NAAG, which was best quantified with PRESS-80, and Cr and PCr, which were not well separated by any sequence.

Comparison with existing methods and conclusion: When wanting to specifically separate composite metabolite signals using single voxel MRS, shorter echo times generally perform better. Researchers should be mindful of the effects of acquisition parameters on the metabolite measures.

1. Introduction

Proton magnetic resonance spectroscopy (MRS) is used to measure neurometabolite levels non-invasively *in vivo*. The main metabolites of interest that can be measured with MRS are Glutamate (Glu), N-acetyl aspartate (NAA), creatine (Cr), and choline (Cho). However, interpretation of these measures is complicated due to overlapping signals. At 3 T, signal from each of these metabolites is overlapped by signal of a metabolite with similar functional groups. Subsequently, these metabolites are often reported as a "sum" of multiple metabolites (e.g. Glx = Glu + glutamine (Gln), tNAA = NAA + N-acetylaspartylglutamate (NAAG), tCr = Cr + phosphocreatine (PCr), tCho = Cho + phosphocholine (PC) + glycerophosphocholine (GPC)).

Some studies do not report the composite signal, as there is often a desire to be specific to one metabolite (e.g. Duncan et al., 2013; Reyngoudt et al., 2011; Sarchielli et al., 2005; Siniatchkin et al., 2012). This may result in inaccurate quantification, for example reporting NAA without consideration that some NAA signal may be quantified as NAAG, or vice versa. A more comprehensive approach could be to report both the individual signals and the composite signal, which has been seen for Glu and Glx (e.g. Joyce et al., 2022) but is not common for other metabolites. A related challenge is attempting to interpret changes in the composite signal as changes in one metabolite only, without considering changes in the other, for example changes in Glx interpreted as changes

* Correspondence to: ACHRI 4th floor offices, Alberta Children's Hospital, 28 Oki Drive, Calgary AB T3B 6A8, Canada.
E-mail address: t.bell@reading.ac.uk (T.K. Bell).

https://doi.org/10.1016/j.jneumeth.2025.110523
Received 7 February 2025; Received in revised form 10 June 2025; Accepted 29 June 2025
Available online 30 June 2025

in Glu only. These assumptions can result in inaccurate interpretations, for example changes in Glu are generally interpreted as changes in glutamatergic processing, however changes in Gln have been associated with increased GABAergic activity (Rae, 2014). Additionally, reporting the metabolites as a single sum could mask changes in one metabolite due to altered levels of the other. Indeed, several functional studies have shown that reporting the composite signal masks changes in individual metabolites (e.g. tCho (Bell et al., 2018), tNAA (Castellano et al., 2012)).

Previous work has attempted to identify a preferred approach to separate signals at 3 T by comparing Cramer-Rao Lower Bounds (CRLBs) and the intra-subject coefficient of variation (CVs) of data acquired using various sequences. CRLBs are a measure of the minimum error associated with a measurement and are generally reported as a percentage of the quantified value (Kreis, 2016). In the context of comparing MRS sequences, measures with lower CRLBs are interpreted to be more accurate. Intra-subject CVs measure the variation in data dispersion between two measures and are used as a measure of reliability or consistency between two repeated scans, sequences with lower CVs are interpreted to be more reliable/consistent. For example, Wijtenburg and Knight-Scott (2011) compared CRLBs and CVs for four sequences: (1) phase-rotated short echo time STEAM (TE=6.5 ms), (2) PRESS (TE=40 ms), (3) STEAM (TE=72 ms) and (4) TE-averaging. They found short echo STEAM (TE=6.5 ms) to have the lowest CRLB (5.3) and the lowest CV (7.1 %) for Glu. Deelchand et al. (2018) showed that semi-LASER (sLASER, TE=28 ms) produced data with lower CRLBs and better repeatability for most metabolites compared to data acquired using PRESS (TE=30 ms).

CRLBs reported as a percentage have been critisized as a measure of error as they heavily depend on the estimated metabolite level, larger metabolite levels will inherently have smaller CRLBs (Kreis, 2016). The CV suffers from similar limitations, smaller measures will produce larger CVs than larger measures. Therefore interpreting these measures as measures of accuracy is misleading. To determine accuracy, a measure of ground truth is neeeded, which is absent in in vivo data. Instead, simulated data can be used to provide a known value for comparison. For example, Hancu (2009) used simulated data to compare the ability of multiple sequences to separate Glu & Gln (PRESS, 4 different echo times, STEAM, 2 different echo times, JPRESS, 2 different echo times and Carr Purcell echo train, 2 different echo times). They concluded JPRESS (TE=35–190 ms) to be the most accurate (had the lowest difference between measured value and simulated value), however this sequence had lower repeatability. PRESS (TE=15 ms) showed high accuracy and strong repeatability for Glu measures, however was unable to be tested in vivo due to RF peak power limitations, highlighting a limitation of using simulated data. Additionally, there are numerous variables affecting in vivo data that are often unaccounted for in simulations, creating a gap between simulated and in vivo data. For example, although Hancu found a lower CV for PRESS (TE=30 ms) compared to PRESS (TE=45 ms) using simulated data, using in vivo data Mullins et al. (2008) found PRESS (TE=40 ms) to have a lower CV for Glu than PRESS (TE=30 ms), though this discrepancy may in part be due to the slightly longer echo time in Hancu's study. Mullins et al. chose TE= 40 ms over TE= 45 ms as this echo time minimizes contribution from overlapping Gln signal, whilst still maintaining signal intensity on a similar level to that of shorter echo time sequences, a critique of using a TE of 80 ms, which suppresses the Gln signal. Both Mullins et al. and Hancu were in agreement that PRESS (TE=30 ms) produced lower CRLBs, however Hancu was able to show that PRESS (TE=30 ms) overestimates Glu levels, highlighting the limitations of using CRLBs as a measure of accuracy.

An approach to bridge the gap between simulated and in vivo data is to use data acquired at a higher field strength, i.e. 7 T, as a reference for comparison. 7 T has a higher spectral resolution, meaning overlapping signals are more easily resolved at 7 T than at 3 T (Pradhan et al., 2015). This manuscript assesses the agreement between multiple sequences acquired at 3 T and 7 T data acquired using STEAM (TE=8 ms) as a reference comparison. 3 T sequences that show strong agreement will be assumed to be better at separating metabolite signals. 7 T STEAM (TE=8 ms) was chosen based on Bell et al. (2025), who showed this sequence to be optimal for the separation of metabolites at 7 T. Data acquired at 7 T using sLASER (TE=34 ms) will also be included to check if the choice of 7 T sequence for reference biased findings.

2. Methods

2.1. Participants

14 healthy participants (18–40 y) were recruited and written informed consent was obtained from all participants. Data were collected with approval from the Conjoint Health Research Ethics Board, University of Calgary (REB19–1461), the Research Ethics Board at The Douglas Research Centre (IUSMD-19–12), the McGill Faculty of Medicine Institutional Review Board (A06-M49–22B) and the Montreal Neurological Institute Institutional Review Board (2020–6510).

2.2. Data acquisition

Participants underwent two imaging sessions, one at 3 T (Siemens Prisma) and one at 7 T (Siemens Terra). The sessions took place within 3 days at approximately the same time of day and the session order was counterbalanced so that equal numbers of participants underwent the 3 T scan vs the 7 T scan first. A 32-channel head coil was used for each scanner. Each session began with a T1-weighted image for voxel placement and segmentation (3 T: MPRAGE, TE/TR= 2.01/2300 ms, isotropic resolution=1 mm^3; 7 T: MPRAGE, TE/TR= 1.84/4300 ms, isotropic resolution=1 mm^3). 2.5 × 2.5 × 2.5 cm voxels were placed in the parietal cortex centred on the midline (Fig. 1).

At 3 T, the following acquisitions were performed:

1. PRESS-20 (TE=20 ms)
2. PRESS-30 (TE=30 ms)
3. PRESS-40 (TE=40 ms)
4. PRESS-80 (TE=80 ms)
5. STEAM-6 (TE/TM=6/32 ms)
6. sLASER-28 (TE=28.24 ms)

All 3 T acquisitions had the following parameters: TR= 2000 ms, number of points= 4128, spectral width= 4000 Hz, 128 water-suppressed averages, VAPOR water suppression, and 8 water unsuppressed averages. Sequence order was counterbalanced across subjects. At 7 T, a phase-rotated STEAM acquisition was performed (TE/TR/TM=8/6000/40 ms, 64 water-suppressed averages using VAPOR, 8 water unsuppressed averages) which was originally developed to improve glutamate sensitivity (Wijtenburg and Knight-Scott, 2011). As a supplemental reference data set, sLASER data was also acquired at 7 T (TE/TR=34/5000 ms, 64 water-suppressed averages using VAPOR, 8 water unsuppressed averages).

2.3. Data processing and quantification

MRS data were preprocessed using FID-A (Simpson et al., 2017) with the following steps: coil combination, removal of motion corrupted averages, frequency drift correction, and zero order phase correction. LCModel (Provencher, 2001) was used to apply eddy current correction and for quantification relative to water. The ATT_{H2O} parameter was set to 1 to allow for correction of T_2 decay of water at each echo time external to LCModel.

Basis sets for quantification were simulated using FID-A based on exact timings and RF pulses for each sequence. Basis sets included the following metabolites: Alanine, Aspartate, Choline (Cho), Glycerophosphocholine (GPC), Phosphocholine (PC), Creatine (Cr), Phosphocreatine (PCr), GABA, Glu, Gln, Lactate, Myo-Inisitol, NAA, N-

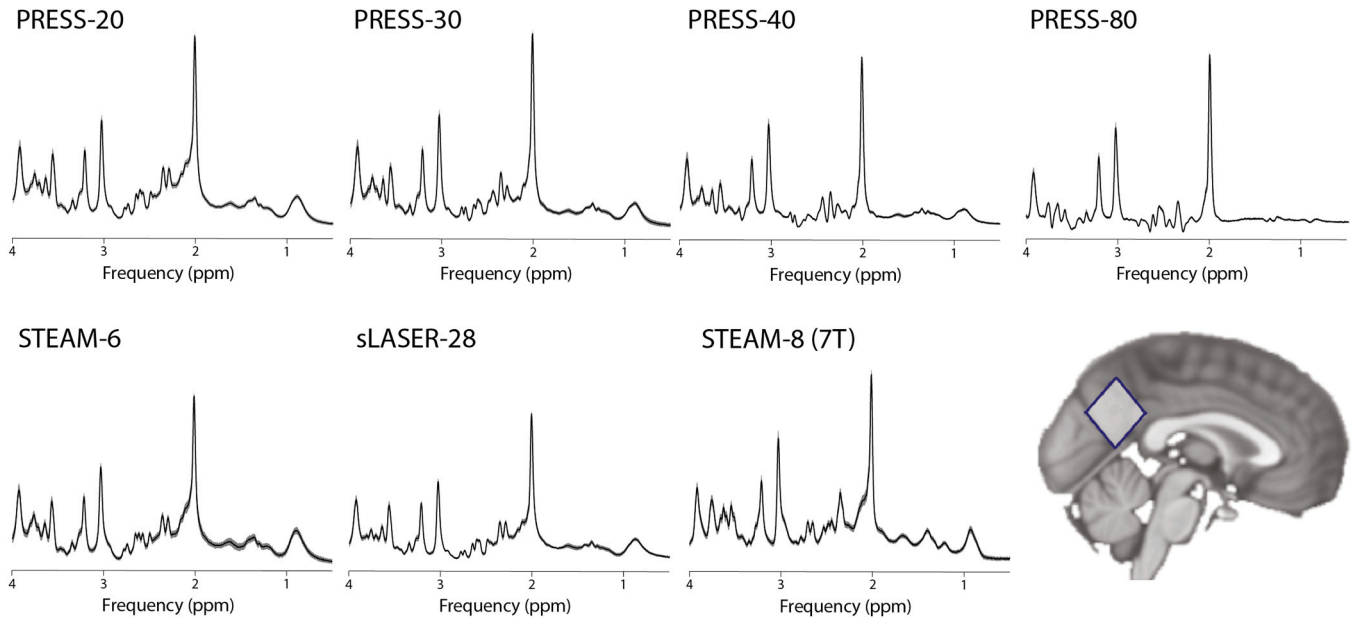

Fig. 1. Mean (black) and standard deviation (shaded grey) of spectra acquired for each sequence, and location of voxel in parietal cortex.

acetylaspartylglutamate (NAAG), Scyllo-Inositol, Glutathione (GSH), Glucose, Taurine, Glycine and PE (basis sets, including simulated macromolecules, are available at https://github.com/HarrisBrainLab/BasisSets).

Macromolecules for the basis sets were simulated using information obtained by parameterizing metabolite suppressed data. Metabolite suppressed data were acquired from five participants at 3 T from a region of interest in the parietal cortex using an inversion recovery SPECIAL sequence with TE = 8.5 ms and TI = 800 ms in five healthy volunteers. The five metabolite nulled spectra were then aligned and combined into a single "average" macromolecule spectrum. Nine macromolecule resonances were then identified in the average macromolecule spectrum, and each resonance was fit to a Lorentzian lineshape using in house custom MATLAB scripts to estimate the frequency, linewidth, and relative amplitude of each component. For 7 T macromolecule basis sets, freely available 7 T metabolite suppressed data from six healthy individuals (Považan et al., 2015, https://zenodo.org/records/3906754#.XvOWVvJ7n7e) was used to identify nine macromolecule resonances, and each macromolecule resonance was fit to a Gaussian lineshape using in house custom MATLAB scripts.

As described in Bell et al. (2025), a single-proton FID-A spin system was created for each identified macromolecule resonance, with the "scaleFactor" parameter adjusted to account for the number of protons associated with each macromolecule resonance (see Table 1 below, FID-A spin systems are available at https://github.com/HarrisBrainLab/MMSim). Macromolecules were then simulated with the appropriate simulation function in FID-A based on the parameters in Table 1, using exact timings and RF pulses. The simulated macromolecule linewidth was 6 Hz smaller than their corresponding linewidths in the parameterized spectra to account for the line broadening typically applied by LCModel to match the narrow-simulated metabolite basis functions to the typically broader linewidths observed in *in vivo* data. A soft constraint on the relative concentration of each macromolecule relative to $MM_{0.89}$ was applied in LCModel as specified in Table 1 (Bell et al., 2025).

T1 weighted images were segmented using the CoRegStandAlone function (Harris et al., 2015) from the Gannet toolbox (Edden et al., 2014). Tissue correction was performed using the Gasparovic method, including correcting for echo time, repetition time and T_1 and T_2 decay using field specific values (Gasparovic et al., 2006). The op_getLW function in FID-A was used to estimate the linewidth of the NAA peak (1.8–2.2 ppm) and the water peak (4.4–4.8 ppm). SNR was estimated using the op_getSNR function in FID-A, which measures the height of the NAA resonance between 1.8 and 2.2 ppm and normalizes this to the standard deviation of the signal in a peak free region.

Table 1
Values used for simulation of macromolecules.

Macromolecule	Frequency (ppm)	No. Protons[a]	Simulated Linewidth (Hz)	LCModel Constraint[b]
$M_{0.89}$	0.89	3	17.3	n/a
$M_{1.19}$	1.19	3	13.9	0.41 ± 0.16
$M_{1.37}$	1.37	3	18.5	0.98 ± 0.67
$M_{1.62}$	1.62	2	35.6	1.185 ± 1.10
$M_{2.00}$	2.00	2	29.8	2.53 ± 0.81
$M_{2.25}$	2.25	2	45.0	2.00 ± 0.17
$M_{2.97}$	2.97	2	37.1	1.30 ± 0.32
$M_{3.19}$	3.19	2	4.0	0.10 ± 0.22
$M_{3.79}$	3.79	2	55.9	2.33 ± 0.72

[a] Values obtained from Cudalbu et al., (2021)
[b] Concentration relative to $M_{0.89}$

2.4. Statistical analyses

Statistical analyses were performed using R (version 4.1.2, R Core Team 2020 https://www.R-project.org/). Data acquired at 7 T were considered the reference standard. To determine the consistency of voxel placement, the dice coefficient of the overlap between the 3 T and 7 T voxels was calculated. Pearson's correlation analyses was used to assess the relationship between data acquired at 3 T and 7 T. Absolute magnitudes between 0 and 0.3 are considered weak correlations, absolute magnitudes between 0.4 and 0.6 are considered moderate correlations and absolute magnitudes between 0.7 and 1 are considered strong correlations (Akoglu, 2018). A mixed effects intraclass correlation coefficient (ICC) model with 95 % confidence intervals was calculated using the irr package (v 0.84.1) (Gamer et al., 2019) to test the level of agreement. Values below 0.5 are considered poor agreement, values between 0.5 and 0.7 are considered moderate agreement and values above 0.7 are considered good agreement (Koo and Li, 2016). As a secondary analysis, the pairwise correlation coefficient from LCModel was used to assess the relationship between overlapping metabolites (for example, the correlation between NAA and NAAG). A correlation closer

to zero implies better separation of overlapping signal, whereas positive or negative correlations with a larger magnitude indicate poor separation (Dou et al., 2015), and it is recommended in the LCModel manual that signal has not been adequately separated if the correlation between a pair is consistently less than −0.3.

2.5. Sample size calculations

Calculations were performed for each individual metabolite to estimate the number of participants per group that would be needed to detect a 10 % change between two groups, based on the formula described by Wong et al. (2018)

$$n = (u + v)^2 \quad \times \quad \frac{(2 \quad \times \quad \sigma^2)}{(\Delta \times \mu)^2}$$

Where $u = 1.28$ % for 90 % power, $v = 1.96$ % for 95 % confidence level, $\Delta = 0.1$ for a 10 % change, σ = the standard deviation of each metabolite across all subjects and μ = the mean of each metabolite across all subjects.

3. Results

3.1. Quality metrics

Table 2 shows the linewidth and SNR for each sequence. The mean of the CRLBs for each metabolite for each sequence can be found in supplementary table 1. CRLBs were relatively consistent across methods. The mean CRLBs for Glu and Glx were all below 5, and typically the mean CRLBs for Gln were between 13 and 18, with the exception of sLASER-28, the only method that was substantially higher (CRLB = 27). Similarly for NAA and tNAA, mean CRLBs were below 2 for all sequences. For NAAG, the PRESS sequences had mean CRLBs between 8 and 12, whereas STEAM-6 and sLASER-28 had mean CRLBs of 15 and 14, respectively. Mean CRLBs for Cr metabolites were all around 5 or below, except for STEAM-6 PCr values, which had a mean of 7.

The mean dice coefficient was 0.68 (standard deviation = 0.12).

3.2. 7 T STEAM comparisons

3.2.1. Glutamate

GLU: All correlations between 3 T and 7 T data were significant. The 7 T data correlated most strongly with Glu values acquired using sLASER-28, followed by STEAM-6 and PRESS-20, which performed similarly (Fig. 3A). The smallest difference between 7 T and 3 T values was obtained using STEAM-6, followed by PRESS-40 and PRESS-20 (Fig. 2, Fig. 3B). All sequences resulted in ICC values above 0.5. sLASER-28 produced the largest ICC value, followed by STEAM-6 and PRESS-20 (Fig. 3C). The number of participants required to detect a 10 % difference ranged from 6 to 12, with STEAM-6 and PRESS-20 requiring equally low numbers (Table 3).

GLN: All correlations between 3 T and 7 T data were significant. The 7 T data correlated most strongly with Gln values acquired using PRESS-20, followed by PRESS-30 and sLASER-28, which performed equally well (Fig. 3A). The smallest difference between 7 T and 3 T values was obtained using PRESS-40, followed by STEAM-6 and PRESS-80. Gln was consistently underestimated by all 3 T sequences (Fig. 2A, Fig. 3B). All sequences resulted in ICC values above 0.5. PRESS-20 and PRESS-30 produced the largest ICC values, followed by sLASER-28 (Fig. 3C). The number of participants required to detect a 10 % difference ranged from 44 to 393, with PRESS-40 requiring the lowest number (Table 3).

GLX: All correlations between 3 T and 7 T data were significant. The 7 T data correlated most strongly with Glx values acquired using PRESS-20, followed by sLASER-28 (Fig. 3A). The smallest difference between 7 T and 3 T values was obtained using PRESS-30 and PRESS-40 (Fig. 2, Fig. 3B). All sequences resulted in ICC values above 0.5. PRESS-20 and sLASER-28 produced the largest ICC value (Fig. 3C). The number of participants required to detect a 10 % difference ranged from 8 to 14, with PRESS-20 and STEAM-6 requiring equally low numbers (Table 3).

The magnitude of the LCModel pairwise correlation coefficients was lower than 0.3 for all sequences (Table 4).

3.2.2. N-acetyl aspartate

NAA: The 7 T data correlated most strongly with NAA values acquired using STEAM-6, followed by PRESS-80. The only other significant correlation was with PRESS-40 (Fig. 3A). Generally, the 3 T sequences overestimated NAA. The smallest difference between 7 T and 3 T values was obtained using sLASER-28, followed by PRESS-20. (Fig. 2B, Fig. 3B). STEAM-6 produced the largest ICC value. PRESS-80 and PRESS-40 were the only other sequences with ICC values above 0.5 (Fig. 3C). The number of participants required to detect a 10 % difference ranged from 2 to 12, with STEAM-6 and sLASER-28 requiring equally low numbers (Table 3).

NAAG: The 7 T data correlated most strongly with NAAG values acquired using PRESS-80, with the only other significant correlation being STEAM-8 (Fig. 3A). The smallest difference between 7 T and 3 T values was obtained using PRESS-80, followed by PRESS-40 (Fig. 2, Fig. 3B). STEAM-6 produced the largest ICC value, followed by PRESS-80. No other sequences produced ICC values above 0.5 (Fig. 3C). The number of participants required to detect a 10 % difference ranged from 12 to 72, with PRESS-30 requiring the lowest number (Table 3).

tNAA: All correlations between 3 T and 7 T data were significant except for PRESS-30 data. The 7 T data correlated most strongly with NAA values acquired using STEAM-6, followed by PRESS-80 (Fig. 3A). Generally, the 3 T sequences overestimated tNAA. The smallest difference between 7 T and 3 T values was obtained using sLASER-28, followed by PRESS-20 (Fig. 2, Fig. 3B). All ICC values were above 0.5 except for PRESS-30 data. STEAM-6 produced the largest ICC value (Fig. 3C). The number of participants required to detect a 10 % difference ranged from 1 to 11, with STEAM-6 requiring the lowest number (Table 3).

The magnitude of the LCModel pairwise correlation coefficients was higher than 0.3 for all sequences, in contrast to 7 T STEAM, which had a pairwise correlation coefficient of −0.04 (Table 4).

3.2.3. Creatine

Cr: There were no significant correlations between 7 T Cr values and 3 T Cr values from any sequence. The highest correlation was with PRESS-20, but this did not reach statistical significance ($r = 0.24$, $p = 0.44$) (Fig. 3A). The smallest difference between 7 T and 3 T values was obtained using PRESS-20 followed by sLASER-28 (Fig. 2C, Fig. 3B). None of the ICC values were above 0.5. sLASER-28 and PRESS-20 produced the highest ICC values, but this did not meet the threshold of agreement (Fig. 3C). The number of participants required to detect a 10 % difference ranged from 7 to 26, with sLASER-28 requiring the lowest number (Table 3).

Table 2
Mean and standard deviation of quality metrics for each sequence.

	PRESS-20	PRESS-30	PRESS-40	PRESS-80	sLASER-28	STEAM-6	7 T STEAM
NAA LW (Hz)	6.22 ± 2.52	4.74 ± 0.64	4.34 ± 0.49	5.30 ± 5.11	5.04 ± 0.69	7.35 ± 3.75	13.18 ± 1.65
H_2O LW (Hz)	7.28 ± 3.17	6.24 ± 1.03	6.45 ± 1.80	7.27 ± 4.73	5.61 ± 0.84	20.10 ± 5.17	12.26 ± 2.79
SNR	76.58 ± 41.60	73.37 ± 39.28	60.53 ± 32.27	45.48 ± 28.56	71.75 ± 39.71	43.12 ± 23.16	235.38 ± 76.93

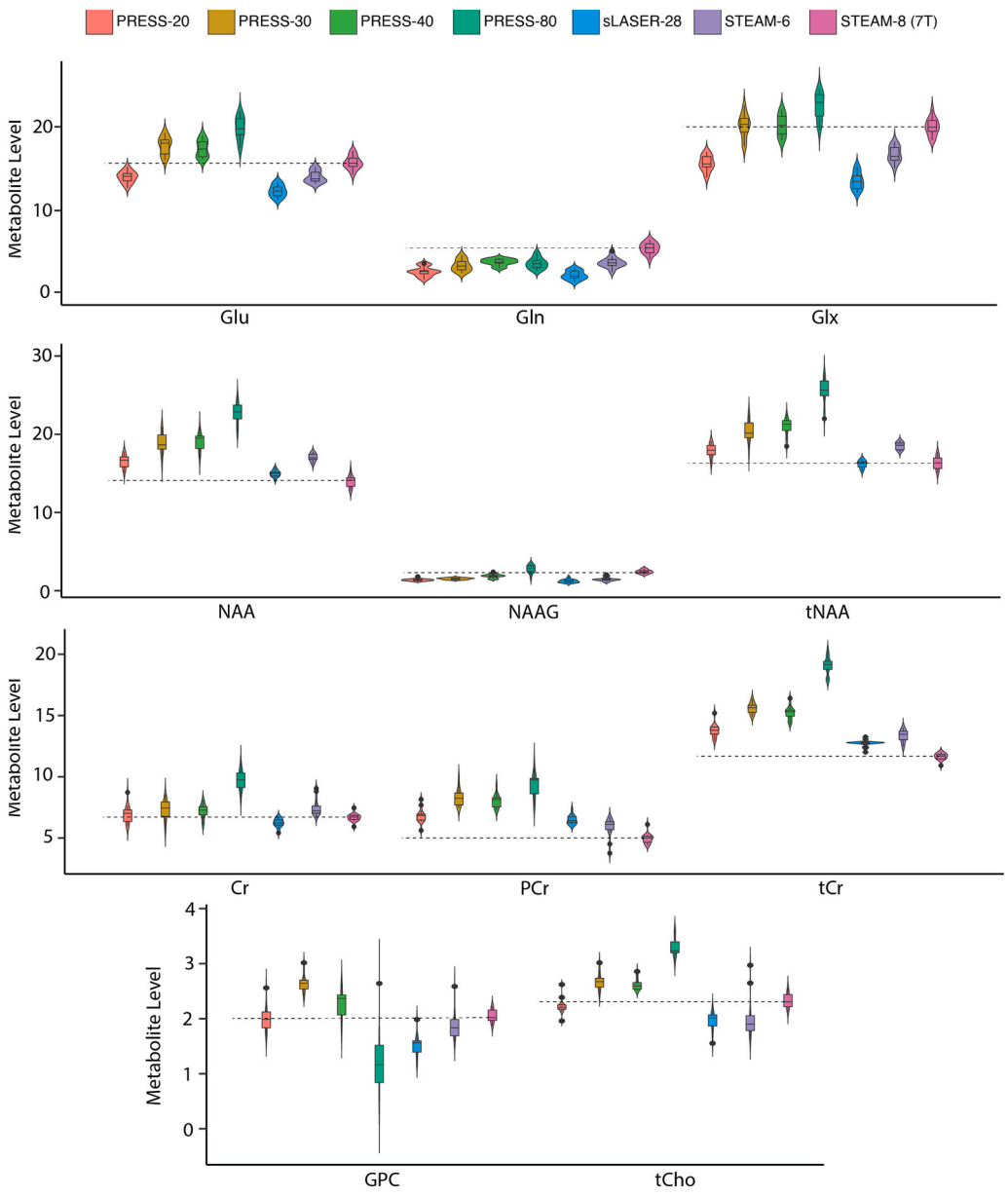

Fig. 2. Spread of data for each metabolite quantified by each sequence. Dashed line represents mean of 7 T data. Cho and PC were not well quantified and therefore were not included for analyses.

Table 3
Number of subjects needed to detect a 10 % difference between groups.

Met	PRESS-20	PRESS-30	PRESS-40	PRESS-80	STEAM-6	sLASER-28	7 T STEAM
Glu	6	9	8	12	6	8	6
Gln	266	173	44	131	114	393	47
Glx	8	12	10	12	8	14	6
NAA	6	12	8	9	2	2	6
NAAG	28	12	42	72	72	65	21
tNAA	5	10	7	11	1	2	6
Cr	25	26	13	15	20	7	6
PCr	17	14	10	26	44	7	15
tCr	4	2	3	3	3	1	1
GPC	31	8	34	562	48	40	7
tCho	9	7	4	7	80	17	7

PCr: Only PCr data acquired using PRESS-20 significantly correlated with 7 T data. (Fig. 3A). Generally, the 3 T sequences overestimated PCr. The smallest difference between 7 T and 3 T values was obtained using STEAM-6 followed by sLASER-28 (Fig. 2C, Fig. 3B). Only PCr data acquired using PRESS-20 produced an ICC above 0.5 (Fig. 3C). The number of participants required to detect a 10 % difference ranged from

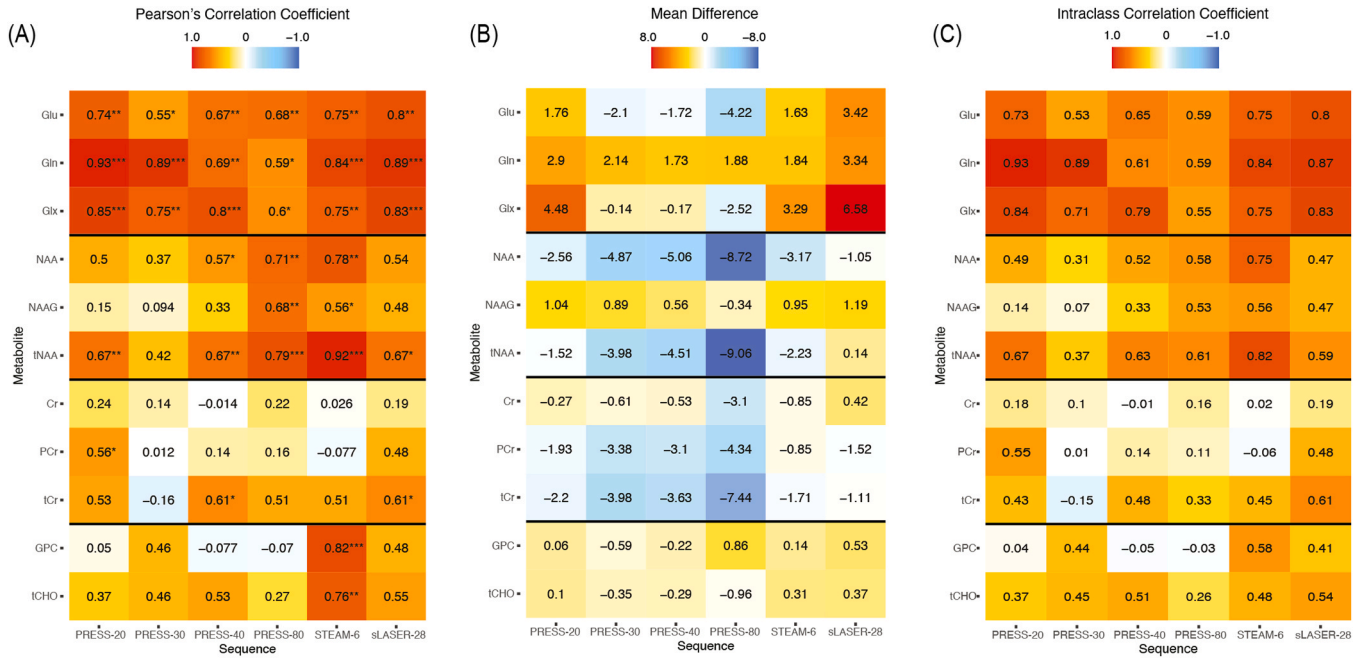

Fig. 3. Heatmaps showing (A) correlation coefficient of data from each 3 T sequence compared to 7 T (higher correlations shown in red), (B) mean difference between metabolite values from each 3 T sequence compared to 7 T (i.e. 7 T measure – 3 T measure, smaller differences shown in white), (C) intraclass correlation coefficient for each metabolite from each 3 T sequence compared to 7 T (higher coefficients shown in red). *p < 0.05, **p < 0.01, ***p < 0.001.

Table 4
Mean and standard deviation of the LCModel pairwise correlation coefficients.

	Glx	tNAA	tCr
PRESS−20	0.16 ± 0.04	−0.47 ± 0.06	−0.81 ± 0.30
PRESS−30	−0.03 ± 0.03	−0.55 ± 0.04	−0.90 ± 0.01
PRESS−40	−0.11 ± 0.02	−0.59 ± 0.09	−0.89 ± 0.01
PRESS−80	0.26 ± 0.03	−0.74 ± 0.08	−0.92 ± 0.01
sLASER−28	0.16 ± 0.02	−0.46 ± 0.08	−0.87 ± 0.003
STEAM−6	0.07 ± 0.03	−0.35 ± 0.06	−0.83 ± 0.01
7 T STEAM	0.02 ± 0.02	−0.04 ± 0.07	−0.68 ± 0.02

7 to 44, with sLASER-28 requiring the lowest number (Table 3). Interestingly, PRESS-30, PRESS-40 and sLASER-28 required lower numbers than 7 T STEAM data.

tCr: Only tCr data acquired using PRESS-40 and sLASER-28 significantly correlated with 7 T data. (Fig. 3A). Generally, the 3 T sequences overestimated tCr. The smallest difference between 7 T and 3 T values was obtained using sLASER-28 (Fig. 2C, Fig. 3B). Only tCr data acquired using sLASER-28 produced an ICC above 0.5 (Fig. 3C). The number of participants required to detect a 10 % difference ranged from 1 to 4 with sLASER-28 requiring the lowest number (Table 3).

The magnitude of the LCModel pairwise correlation coefficients was higher than 0.3 for all sequences. This was also the case for 7 T STEAM data (Table 4).

3.2.4. Choline

Cho: Cho was not well quantified at 3 T and therefore was not compared.

PC: PC was not well quantified at 7 T and therefore 3 T data was not compared.

GPC: Only GPC data acquired using STEAM-6 significantly correlated with 7 T data (Fig. 3A). The smallest difference between 7 T and 3 T values was obtained using PRESS-20, followed by STEAM-6 (Fig. 2D, Fig. 3B). Only GPC data acquired using STEAM-6 produced an ICC above 0.5 (Fig. 3C). The number of participants required to detect a 10 % difference ranged from 8 to 562, with PRESS-30 requiring the lowest number (Table 3).

tCho: Only tCho data acquired using STEAM-6 significantly correlated with 7 T data. (Fig. 3A). The smallest difference between 7 T and 3 T values was obtained using PRESS-20, followed by STEAM-6 and sLASER-28 (Fig. 2D, Fig. 3B). Only tCho data acquired using PRESS-40 and sLASER-28 produced an ICC above 0.5 (Fig. 3C). The number of participants required to detect a 10 % difference ranged from 4 to 80, with PRESS-40 requiring the lowest number (Table 3).

As Cho and PC were not well quantified, the magnitude of the LCModel pairwise coefficients was not compared.

3.2.5. 7 T sLASER comparisons

To determine the effect of using a 7 T STEAM sequence for comparison, we also repeated the above comparisons with a sLASER sequence (TE=34 ms) acquired at 7 T. Results from these comparisons can be found in the supplementary material.

Glu acquired using PRESS-20, sLASER-28 and STEAM-6 at 3 T agreed equally well with 7 T sLASER (TE=34 ms) data. For Gln data, PRESS-20 showed the strongest agreement followed by sLASER-28. For Glx data, sLASER-28 showed the strongest agreement, followed by PRESS-20. NAA and tNAA data from any sequence did not agree with 7 T sLASER data. NAAG data acquired using PRESS-20 showed the strongest agreement. Cr data also showed poor agreement. PCr data acquired from PRESS-20 and tCr data acquired from sLASER-28 showed the strongest agreement. 3 T GPC data did not agree with 7 T sLASER data. tCho data from STEAM-6 showed the strongest agreement with 7 T sLASER data.

4. Discussion

A consensus of a preferred approach to separate signals at 3 T has not been reached, partially due to the gap between simulated and in vivo data. To bridge this gap, this manuscript used in vivo data acquired at 7 T as a reference point for comparison. Higher field strength data provides higher SNR and improved spectral resolution, and thus better separation of overlapping signals. Data acquired at 3 T using multiple sequences and echo times were compared to data acquired at 7 T. 3 T sequences that show strong agreement will be assumed to be better at separating metabolite signals.

Data acquired using PRESS-40 had the lowest mean NAA linewidth, and data acquired using sLASER-28 had the lowest water linewidth, however generally linewidth was similarly low across sequences. As linewidth increases with field strength (Pradhan et al., 2015; Tkáč et al., 2009), data acquired at 7 T generally had a higher linewidth than 3 T data, apart from data acquired using STEAM-6, which had the highest water linewidth. This may be due to the increased susceptibility of STEAM to motion effects due to the separation of crushers by the TM time period, though this effect is usually minimal (Barker and Lin, 2006). Additionally, the slice profile of STEAM pulses generally results in a larger effective voxel size than PRESS. Therefore there is likely better homogeneity in the "smaller" PRESS voxel (Moonen et al., 1989). As expected, 7 T had a substantially higher SNR than 3 T data. At 3 T, SNR decreased with increasing echo time, likely due to increased signal decay, and STEAM-6 had the lowest SNR, which is due to the nature of the STEAM acquisition.

Although Glu acquired with sLASER-28 at 3 T showed the strongest correlation (0.8) and the highest ICC value (0.8), it also had the highest mean difference (3.42). STEAM-6 and PRESS-20 also showed strong correlations (0.75 and 0.74, respectively) and high ICC values (0.75 and 0.73, respectively) and a lower mean difference (1.63 and 1.76, respectively). While these coefficients are slightly lower than sLASER, the difference is marginal, therefore, based on the lower mean difference, we conclude STEAM-6 to show the strongest agreement with 7 T data followed by PRESS-20 and sLASER-28. However, it is unclear if there is substantial benefit of one over the other for Glu. Based on the three metrics, Gln acquired with PRESS-20 showed the strongest agreement with 7 T data. Though it should be noted that Gln acquired with PRESS-20 also required one of the largest number of subjects to detect a 10 % difference. This number is a function of the variability of the metabolite and it's concentration level, therefore metabolites with low concentration and low standard deviations (such as Gln quantified using PRESS-20) will require a larger number of subjects to detect a difference. Whether the inter-subject variability is truly this low or whether the sequence cannot accurately quantify variability in concentrations this low is unknown.

Overall, we interpret shorter echo times to be more successful in the quantification of Glu and Gln. Although previous work has shown longer echo times can quantify glutamate well due to suppression of the Gln signal (Schubert et al., 2004; Wong et al., 2018), longer echo times were not directly compared to ultra-short echo times as in this study. Here we show that ultra-short echo times out-perform TEs specifically chosen to suppress the Gln signal (PRESS-40, Mullins et al. 2008 and PRESS-80, Schubert et al. 2004). This is in line with previous work by Shanbhag et al. (2006), who used simulations of both STEAM and PRESS with TEs 10–35 ms to show that the overlap between Glu and Gln signal increases with increasing TE (Shanbhag et al., 2006). Additionally, Wijtenburg and Knight-Scott (2011) showed that Glu could be well quantified at 3 T using STEAM with TE= 6.5 ms (Wijtenburg and Knight-Scott, 2011).

Hancu (2009) used simulations to compare Glu quantification from several sequences similar to the ones used in the present study (PRESS TE=15 ms, 35 ms, 45 ms, 80 ms, STEAM TE=5 ms, CV = 2.9, 4.0, 5.9, 5.1 and 4.1, respectively). They found PRESS (TE=15 ms) to produce the smallest CV, CRLB and absolute error, followed by PRESS (TE=35 ms) and STEAM (TE=5 ms), which performed similarly (Hancu, 2009). Based on our in vivo comparisons, we found STEAM-6 and PRESS-20 to perform similarly, and PRESS-30 ranked much lower for quantification of Glu. This difference may in part be due to the slight difference in echo times but is also likely influenced by the inherent differences between simulated and in vivo data. Thus, the in vivo study presented here builds on previous work using simulations to show the benefit of low echo times for metabolite quantification whilst providing support for the use of these sequences in vivo.

NAA data acquired using STEAM-6 showed the strongest agreement with 7 T data. NAAG data acquired using PRESS-80 showed the strongest agreement with 7 T data, followed by STEAM-6. With the stronger ICC, we therefore conclude STEAM-6 performed the best at separation of NAA metabolites at 3 T, though PRESS-80 may be better if NAAG is of specific interest. This is surprising, as simulations have shown that the 2.67 ppm aspartyl moiety of NAA is nulled/inverted at TE= 70–100 ms (Chan et al., 2017; Wilman and Allen, 1996), therefore it is likely that this is the case for NAAG as well. Further, caution should be taken when interpreting separated values, as LCModel pairwise correlation coefficients were above the cutoff for all sequences (though STEAM-6 was close to the cut-off value at −0.35). NAA levels were generally higher at 3 T compared to 7 T, in line with previous work from Pradhan et al. (2015). This could be due to Glu signal being inappropriately assigned as NAA, however if this was the case it would be expected that Glu values would then be lower at 3 T compared to 7 T. Alternatively, the shorter metabolite T_2 values seen at 7 T have been estimated to cause 2–8 % more signal loss than at 3 T (Pradhan et al., 2015), and therefore NAA levels may have been underestimated at 7 T. Though we have attempted to account for field specific T_2 values in our quantification, it may still be a factor. NAA was also generally higher at longer echo times, which may be a result of a less complicated baseline as macromolecular signals have decayed. For quantification, we included a simulated macromolecule baseline, but when macromolecules are present, NAA signal may be inappropriately assigned as macromolecules or vice versa. A measured macromolecule baseline may be able to better account for this, and inclusion of this may alter the findings.

Of note, NAA and tNAA from all 3 T sequences showed poor agreement with 7 T sLASER data. Bell et al. (2025) previously showed 7 T sLASER (TE=34 ms) produced significantly lower values for NAA and tNAA and significantly higher CVs for these metabolites compared to 7 T STEAM (TE=8 ms). We suggest this is a function of the ultra-short echo time reducing T_2 decay. This is consistent with the results of Wijtenburg and Knight-Scott (2011) who similarly found STEAM with TE= 8 ms produced more reliable NAA and tNAA measurements (evidenced by lower CVs) compared to PRESS with TE= 40 ms. Therefore this lack of agreement may be due to reduced reliability of the 7 T comparison sequence. This provides support for the choice of STEAM (TE=8 ms) as our primary comparison sequence, whilst highlighting the impact of the comparison sequence on potential conclusions.

Only PCr measured using PRESS-20 and tCr acquired using sLASER-28 or PRESS-40 correlated with 7 T data, there were no other correlations between Cr, PCr and tCr acquired at 3 T and 7 T. We therefore conclude that Cr and PCr cannot be accurately seperated at 3 T. This is supported by the LCModel pairwise correlation coefficients, which were all well above the cut-off value. It should also be noted that, based on the findings of Bell et al. (2025), Cr and PCr were also not well separated at 7 T (Bell et al., 2025). If PCr is the sole metabolite of interest, PRESS-20 may be used, however researchers should be cautious that Cr is not well quantified with this sequence. There is evidence from phantom studies that the T_2 of PCr is shorter than that of Cr, and the difference between the two in different environements is not linear. Subsequenctly, Cr and PCr signal will change in different ways as a function of TE (Ke et al., 2002), which may explain why there was little agreement between 3 T and 7 T values. Indeed, these differences have been exploited to separate Cr and PCr at 1.5 T (Ke et al., 2003). Additionally, levels of Cr and PCr are related to metabolic demand, and have been shown to be altered in the visual cortex with photic stimulation (Dorst et al., 2022; Ke et al., 2002). Though our measures were from the parietal cortex, it is possible that there were slight alterations in metabolic requirements between the scans. Indeed, a limitation of this study is that the scans were acquired on different days, contributing to variation between the sessions.

For choline-containing metabolites, only GPC and tCho measured using STEAM-6 at 3 T significantly correlated with 7 T measures. At 7 T, Cho and GPC were able to be quantified separately, however PC was not quantified at 7 T. This is likely due to the strong overlap between the PC (3.23 ppm) and GPC (3.22 ppm) resonances. At 3 T, Cho was not quantified. This is suprising due to the the slight separation of the Cho (3.19 ppm) peak compared to the close overlap of the GPC (3.23 ppm)

and the PC (3.22 ppm) peaks, though this may be due to the much lower concentration of Cho compared to GPC and PC. Lindner et al. (2017) used a combination of simulations and in vivo data to show that Cho could be accurately separated from GPC+PC at 3 T (Lindner et al., 2017), however they used the analysis software jMRUI, which operates in the time domain, as opposed to the frequency domain as with LCModel. This difference may be the source of these discrepancies. Both GPC and tCho levels acquired using PRESS-30 showed the strongest agreement with 7 T data. This may be due to the relatively long T_2 values of GPC/tCho, particularly compared with Cr (Cianfoni et al., 2011), resulting in a strong GPC/tCho signal at at 3.9 ppm when using TE= 30 ms, with less contamination by Cr signal.

To determine the effect of the 7 T STEAM sequence as the reference for comparison, we repeated the above comparisons with 7 T sLASER (TE=34 ms) data. In this case, data acquired using PRESS-20 and sLASER-28 showed the strongest agreement with the 7 T data. This could imply that the conclusion of the use of STEAM-6 at 3 T is driven by the comparison data. However, it should be noted that for many metabolites the agreement between 3 T and 7 T data was stronger when using 7 T STEAM data than 7 T sLASER data. Additionally, PRESS-20 agreed well with 7 T data in both cases, highlighting the utility of short echo times in signal separation. It could also be hypothesized that 7 T STEAM and sLASER showed higher correlations with the low echo time data due to the similarities in echo time, and perhaps long echo time 7 T data would correlate better with long echo time 3 T data. To assess this, we also looked at the relationship between 3 T data and 7 T sLASER (TE=105 ms) and found poor agreement with all sequences (results not shown). This may be a function of the longer echo time resulting in more extensive T2 decay (the longest echo time assessed at 3 T was 80 ms).

In this manuscript, we interpret 3 T sequences that show strong agreement with 7 T data to be better at separating metabolite signals. However, there are other factors that will contribute to agreement between sequences, therefore this interpretation should be considered with caution and combined with other literature sources to choose an appropriate sequence. For example, macromolecule content will vary based on echo time. We attempted to account for this by creating a parameterized macromolecule baseline, but it is possible that longer echo time 3 T sequences showed poor agreement with the short TE 7 T sequences due to lower contribution of the macromolecular baseline. However, longer echo times result in lower SNR and dephasing of multiplets, reducing quantitation accuracy of lower concentration or multiplet metabolites. For example, Birch et al. (2017) found TE= 80 ms generated the lowest CVs for high SNR and low complexity metabolites such as tNAA, but TE= 35 ms generated the lowest CVs for Glu, regardless of the method for measuring the macromolecule baseline. The fact that agreement was similar using both 7 T sequences (STEAM TE=8 ms and sLASER TE=34 ms) despite the difference in echo times supports our conclusions of using short echo times to separate overlapping signal.

Variability in voxel placement will also contribute to metabolite concentration agreement. To assess the overlap between 3 T and 7 T voxel position, we calculated the dice coefficient. The average overlap was 68 %, which equates to a 10 % (or 2.5 mm displacement) along the three axes. In addition to the differences in the imaging field-of-view (ie orientation and location of the T1-weighted image) and voxel placement, some of this displacement error also originates from imperfections in registration.

As is the nature of in vivo MRS, a limitation of this study is a lack of ground truth. Though simulations can be used to create a known value for comparison, there are numerous factors affecting in vivo data which are often unaccounted for, which motivated this study. To bridge the gap between simulated and in vivo data, we used data acquired at 7 T as an estimation of the ground truth, due to it's increased SNR and spectral resolution which allows for better separation of overlapping signals. However, the accuracy of the 7 T data is still unknown, and there will likely still be some spectral overlap. Additionally, though we attempted to keep conditions as similar as possible for both scans, including time of day, natural fluctuations in metabolite levels between the scans and voxel repositioning effects likely reduce the agreement seen between the two measures.

In summary we show that 3 T sequences with short echo times (STEAM-6, PRESS-20) generally showed stronger agreement with 7 T data. We interpret this to mean they performed better at separating most metabolites. The exceptions were NAAG, which showed the strongest agreement when using PRESS-80, and Cr and PCr, which did not show agreement when using any sequence. When wanting to specifically separate composite metabolite signals, researchers should be mindful of the effects of acquisition parameters on the metabolite measures and can choose sequences to best target metabolites of interest.

CRediT authorship contribution statement

Jamie Near: Supervision, Funding acquisition, Writing – review & editing, Methodology, Conceptualization. **Dana Goerzen:** Investigation, Project administration. **Ashley D Harris:** Supervision, Funding acquisition, Writing – review & editing, Methodology, Conceptualization. **Tiffany Bell:** Writing – review & editing, Visualization, Formal analysis, Writing – original draft, Methodology.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was undertaken thanks in part to funding from the Natural Sciences and Engineering Research Council of Canada (NSERC, RGPIN-2020–05917, RGPIN-2017–03875), a Canadian Foundation for Innovation John R Evans Leaders Fund (CFI-JELF), and support by the Hotchkiss Brain Institute and the Alberta Children's Hospital Research Institute, University of Calgary. T.K.B is supported by a Canadian Institute of Health Research (CIHR) Postdoctoral Fellowship. J.N. holds funding from the CIHR (PJT-183715) and the New Frontiers Research Fund of Canada (NFRF, NFRFT 2022–00327). A.D.H. holds a Canada Research Chair in Magnetic Resonance Spectroscopy in Brain Injury and holds funding from the CIHR (PJ-T175085) and The Arthritis Society.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.jneumeth.2025.110523.

Data availability

Data will be made available on request.

References

Akoglu, H., 2018. User's guide to correlation coefficients. Turk. J. Emerg. Med. 18, 91–93. https://doi.org/10.1016/j.tjem.2018.08.001.

Barker, P.B., Lin, D.D.M., 2006. In vivo proton MR spectroscopy of the human brain. Prog. Nucl. Magn. Reson. Spectrosc. 49, 99–128. https://doi.org/10.1016/j.pnmrs.2006.06.002.

Bell, T., Lindner, M., Mullins, P.G., Christakou, A., 2018. Functional neurochemical imaging of the human striatal cholinergic system in reversal learning. Eur. J. Neurosci. 47, 1184–1193. https://doi.org/10.1111/ejn.13803.

Bell, T.K., Goerzen, D., Near, J., Harris, A.D., 2025. Examination of methods to separate overlapping metabolites at 7T. Magn. Reson. Med. 93, 470–480. https://doi.org/10.1002/mrm.30293.

Birch, R., Peet, A.C., Dehghani, H., Wilson, M., 2017. Influence of macromolecule baseline on 1H MR spectroscopic imaging reproducibility. Magn. Reson. Med. 77, 34–43. https://doi.org/10.1002/mrm.26103.

Castellano, G., Dias, C.S.B., Foerster, B., Li, L.M., Covolan, R.J.M., 2012. NAA and NAAG variation in neuronal activation during visual stimulation. Braz. J. Med. Biol. Res. 45, 1031–1036. https://doi.org/10.1590/S0100-879X2012007500128.

Chan, K.L., Saleh, M.G., Oeltzschner, G., Barker, P.B., Edden, R.A.E., 2017. Simultaneous measurement of Aspartate, NAA, and NAAG using HERMES spectral editing at 3 Tesla. Neuroimage 155, 587–593. https://doi.org/10.1016/j.neuroimage.2017.04.043.

Cianfoni, A., Law, M., Re, T.J., Dubowitz, D.J., Rumboldt, Z., Imbesi, S.G., 2011. Clinical pitfalls related to short and long echo times in cerebral MR spectroscopy. J. Neuroradiol. 38, 69–75. https://doi.org/10.1016/j.neurad.2010.10.001.

R. Core Team (2020) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL ⟨https://www.R-project.org/⟩, n.d.

Cudalbu, C., Behar, K.L., Bhattacharyya, P.K., Bogner, W., Borbath, T., de Graaf, R.A., Gruetter, R., Henning, A., Juchem, C., Kreis, R., Lee, P., Lei, H., Marjańska, M., Mekle, R., Murali-Manohar, S., Považan, M., Rackayová, V., Simicic, D., Slotboom, J., Soher, B.J., Starčuk, Z., Starčuková, J., Tkáč, I., Williams, S., Wilson, M., Wright, A.M., Xin, L., Mlynárik, V., 2021. Contribution of macromolecules to brain 1H MR spectra: Experts' consensus recommendations. NMR Biomed. 34. https://doi.org/10.1002/nbm.4393.

Deelchand, D., Kantarci, K., Öz, G., 2018. Improved localization, spectral quality, and repeatability with advanced MRS methodology in the clinical setting. Magn. Reson. Med 79 (3), 1241–1250. https://doi.org/10.1002/mrm.26788.

Dorst, J., Borbath, T., Landheer, K., Avdievich, N., Henning, A., 2022. Simultaneous detection of metabolite concentration changes, water BOLD signal and pH changes during visual stimulation in the human brain at 9.4T. J. Cereb. Blood Flow. Metab. 42, 1104–1119. https://doi.org/10.1177/0271678X221075892.

Dou, W., Kaufmann, J., Li, M., Zhong, K., Walter, M., Speck, O., 2015. The separation of Gln and Glu in STEAM: a comparison study using short and long TEs/TMs at 3 and 7 T. Magn. Reson. Mater. Phys. Biol. Med. 28, 395–405. https://doi.org/10.1007/s10334-014-0479-7.

Duncan, N.W., Wiebking, C., Tiret, B., Marjańska, M., Hayes, D.J., Lyttleton, O., Doyon, J., Northoff, G., 2013. Glutamate Concentration in the Medial Prefrontal Cortex Predicts Resting-State Cortical-Subcortical Functional Connectivity in Humans. PLoS One 8. https://doi.org/10.1371/journal.pone.0060312.

Edden, R.A.E., Puts, N.A.J., Harris, A.D., Barker, P.B., Evans, C.J., 2014. Gannet: A Batch-Processing Tool for the Quantitative Analysis of Gamma-Aminobutyric Acid–Edited MR Spectroscopy Spectra. J Magn Reson Imaging 40, 1445–1452. https://doi.org/10.1038/jid.2014.371.

Gamer, M., Lemon, J., Fellows, I., Singh, P., 2019. irr: Various Coefficients of Interrater Reliability and Agreement.

Gasparovic, C., Song, T., Devier, D., Bockholt, H.J., Caprihan, A., Mullins, P.G., Posse, S., Jung, R.E., Morrison, L.A., 2006. Use of tissue water as a concentration reference for proton spectroscopic imaging. Magn. Reson. Med. 55, 1219–1226. https://doi.org/10.1002/mrm.20901.

Hancu, I., 2009. Optimized glutamate detection at 3T. J. Magn. Reson. Imaging 30, 1155–1162. https://doi.org/10.1002/jmri.21936.

Harris, A.D., Puts, N.A.J., Edden, R.A.E., 2015. Tissue correction for GABA-edited MRS: Considerations of voxel composition, tissue segmentation, and tissue relaxations. J. Magn. Reson. Imaging 42, 1431–1440. https://doi.org/10.1002/jmri.24903.

Joyce, J.M., Mercier, L.J., Stokoe, M., La, P.L., Bell, T., Batycky, J.M., Debert, C.T., Harris, A.D., 2022. Glutamate, GABA and glutathione in adults with persistent post-concussive symptoms. NeuroImage Clin. 36, 103152. https://doi.org/10.1016/j.nicl.2022.103152.

Ke, Y., Cohen, B.M., Lowen, S., Hirashima, F., Nassar, L., Renshaw, P.F., 2002. Biexponential transverse relaxation (T2) of the proton MRS creatine resonance in human brain. Magn. Reson. Med. 47, 232–238. https://doi.org/10.1002/mrm.10063.

Ke, Y., Streeter, C., Lowen, S., Nassar, L.E., Parow, A.M., Hennen, J., Yurglun-Todd, D.A., Sarid-Segal, O., Awad, L.A., Rendall, M., Gruber, S.A., Nason, A., Mudrick, M.J., Blank, S.R., Ciraulo, D.A., Renshaw, P.F., 2003. Increased frontal lobe phosphocreatine levels observed in heavy cocaine users after treatment for cocaine dependence - An 1H MRS T2 relaxometry study. Spectroscopy 17, 231–239. https://doi.org/10.1155/2003/859107.

Koo, T.K., Li, M.Y., 2016. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J. Chiropr. Med. 15, 155–163. https://doi.org/10.1016/j.jcm.2016.02.012.

Kreis, R., 2016. The trouble with quality filtering based on relative Cram??r-Rao lower bounds. Magn. Reson. Med. 75, 15–18. https://doi.org/10.1002/mrm.25568.

Lindner, M., Bell, T., Iqbal, S., Mullins, P.G., Christakou, A., 2017. In vivo functional neurochemistry of human cortical cholinergic function during visuospatial attention. PLoS One 12, e0171338. https://doi.org/10.1371/journal.pone.0171338.

Moonen, C.T., von, Kienlin, M., van Zijl, P.C., Cohen, J., Gillen, J., Daly, P., Wolf, G., 1989. Comparison of single-shot localization methods (STEAM and PRESS) for in vivo proton NMR spectroscopy. NMR Biomed. 2, 201–208.

Mullins, P.G., Chen, H., Xu, J., Caprihan, A., Gasparovic, C., 2008. Comparative reliability of proton spectroscopy techniques designed to improve detection of J-coupled metabolites. Magn. Reson. Med. 60, 964–969. https://doi.org/10.1002/mrm.21696.

Považan, M., Hangel, G., Strasser, B., Gruber, S., Chmelík, M., Trattnig, S., Bogner, W., 2015. Mapping of brain macromolecules and their use for spectral processing of 1H-MRSI data with an ultra-short acquisition delay at 7T. Neuroimage 121, 126–135. https://doi.org/10.1016/j.neuroimage.2015.07.042.

Pradhan, S., Bonekamp, S., Gillen, J.S., Rowland, L.M., Wijtenburg, S.A., Edden, R.A.E., Barker, P.B., 2015. Comparison of single voxel brain MRS AT 3T and 7T using 32-channel head coils. Magn. Reson. Imaging 33, 1013–1018. https://doi.org/10.1016/j.mri.2015.06.003.

Provencher, S.W., 2001. Automatic quantitation of localized in vivo 1H spectra with LCModel. NMR Biomed. 14, 260–264.

Rae, C.D., 2014. A guide to the metabolic pathways and function of metabolites observed in human brain 1H magnetic resonance spectra. Neurochem. Res. 39, 1–36. https://doi.org/10.1007/s11064-013-1199-5.

Reyngoudt, H., Paemeleire, K., Descamps, B., De Deene, Y., Achten, E., 2011. P-MRS demonstrates a reduction in high-energy phosphates in the occipital lobe of migraine without aura patients. Cephalalgia 31 31, 1243–1253. https://doi.org/10.1177/0333102410394675.

Sarchielli, P., Tarducci, R., Presciutti, O., Gobbi, G., Pelliccioli, G.P., Stipa, G., Alberti, A., Capocchi, G., 2005. Functional 1H-MRS findings in migraine patients with and without aura assessed interictally. Neuroimage 24, 1025–1031. https://doi.org/10.1016/j.neuroimage.2004.11.005.

Schubert, F., Gallinat, J., Seifert, F., Rinneberg, H., 2004. Glutamate concentrations in human brain using single voxel proton magnetic resonance spectroscopy at 3 Tesla. Neuroimage 21, 1762–1771. https://doi.org/10.1016/j.neuroimage.2003.11.014.

Shanbhag, D., Dunham, S., Knight-Scott, J., 2006. Very Short Echo Time STEAM Improves Detection of Glutamate at 3-T. Proc. 14th Sci. Meet. Int. Soc. Magn. Reson. Med. 14, 3235.

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

Siniatchkin, M., Sendacki, M., Moeller, F., Wolff, S., Jansen, O., Siebner, H., Stephani, U., 2012. Abnormal changes of synaptic excitability in migraine with aura. Cereb. Cortex 22, 2207–2216. https://doi.org/10.1093/cercor/bhr248.

Tkáč, I., Öz, G., Adriany, G., Uğurbil, K., Gruetter, R., 2009. In vivo ^{1}H NMR spectroscopy of the human brain at high magnetic fields: Metabolite quantification at 4T vs. 7T. Magn. Reson. Med. 62, 868–879. https://doi.org/10.1002/mrm.22086.

Wijtenburg, S.A., Knight-Scott, J., 2011. Very short echo time improves the precision of glutamate detection at 3T in 1H magnetic resonance spectroscopy. J. Magn. Reson. Imaging 34, 645–652. https://doi.org/10.1002/jmri.22638.

Wilman, A.H., Allen, P.S., 1996. Observing N-acetyl aspartate via both its N-acetyl and its strongly coupled aspartate groups in in vivo proton magnetic resonance spectroscopy. J. Magn. Reson. - Ser. B, 113, 203–213. https://doi.org/10.1006/jmrb.1996.0178.

Wong, D., Schranz, A.L., Bartha, R., 2018. Optimized in vivo brain glutamate measurement using long-echo-time semi-LASER at 7 T. NMR Biomed. 31, 1–13. https://doi.org/10.1002/nbm.4002.