RESEARCH ARTICLE

WILEY

The intra-individual reliability of ^{1}H-MRS measurement in the anterior cingulate cortex across 1 year

Meng-Yun Wang[1,2] | Max Korbmacher[2,3,4] [iD] | Rune Eikeland[1,2] |
Alexander R. Craven[1,5] | Karsten Specht[1,2,6] [iD]

[1]Department of Biological and Medical Psychology, University of Bergen, Bergen, Norway

[2]Mohn Medical Imaging and Visualization Centre (MMIV), Haukeland University Hospital, Bergen, Norway

[3]Department of Health and Functioning, Western Norway University of Applied Sciences, Bergen, Norway

[4]NORMENT Centre for Psychosis Research, Division of Mental Health and Addiction, University of Oslo and Oslo University Hospital, Oslo, Norway

[5]Department of Clinical Engineering, Haukeland University Hospital, Bergen, Norway

[6]Department of Education, UiT The Arctic University of Norway, Tromsø, Norway

Correspondence
Karsten Specht, Department of Biological and Medical Psychology, University of Bergen, Bergen, Norway.
Email: karsten.specht@uib.no

Meng-Yun Wang, Department of Biological and Medical Psychology, University of Bergen, Bergen, Norway.
Email: mengyun.wang@uib.no

Abstract

Magnetic resonance spectroscopy (MRS) is the primary method that can measure the levels of metabolites in the brain in vivo. To achieve its potential in clinical usage, the reliability of the measurement requires further articulation. Although there are many studies that investigate the reliability of gamma-aminobutyric acid (GABA), comparatively few studies have investigated the reliability of other brain metabolites, such as glutamate (Glu), N-acetyl-aspartate (NAA), creatine (Cr), phosphocreatine (PCr), or myo-inositol (mI), which all play a significant role in brain development and functions. In addition, previous studies which predominately used only two measurements (two data points) failed to provide the details of the time effect (e.g., time-of-day) on MRS measurement within subjects. Therefore, in this study, MRS data located in the anterior cingulate cortex (ACC) were repeatedly recorded across 1 year leading to at least 25 sessions for each subject with the aim of exploring the variability of other metabolites by using the index coefficient of variability (CV); the smaller the CV, the more reliable the measurements. We found that the metabolites of NAA, tNAA, and tCr showed the smallest CVs (between 1.43% and 4.90%), and the metabolites of Glu, Glx, mI, and tCho showed modest CVs (between 4.26% and 7.89%). Furthermore, we found that the concentration reference of the ratio to water results in smaller CVs compared to the ratio to tCr. In addition, we did not find any time-of-day effect on the MRS measurements. Collectively, the results of this study indicate that the MRS measurement is reasonably reliable in quantifying the levels of metabolites.

KEYWORDS
creatine, glutamate, magnetic resonance spectroscopy, myo-inositol, reliability, tCho, tNAA

1 | INTRODUCTION

Magnetic resonance spectroscopy (MRS) can non-invasively quantify the levels of metabolites in the central neural system by measuring different resonance frequencies of the proton hydrogen (^{1}H) embedded in them (Oz et al., 2014; Wilson et al., 2019). In typical applications, a 5-min acquisition time can provide a good-quality dataset, where the spectra of a single voxel can be obtained (Wilson et al., 2019). With MRS, one can distinguish different brain lesions with similar MRI appearance (Oz et al., 2014). In addition, it can be used to investigate the neurometabolic responses to external stimuli in vivo among healthy and patient populations (Pasanta et al., 2023).

Therefore, it is an important means and can be used to assist clinical diagnosis, monitor treatment effects, and facilitate patient management (Oz et al., 2014). However, to fully unleash its potential in investigating neurometabolic responses and to advance its clinical usage, the reliability of the MRS measurement should be articulated.

Indeed, there are several studies focusing on the reliability of MRS with the main focus on the reliability of gamma-aminobutyric acid (GABA), where test–retest analysis was used between sessions (Baeshen et al., 2020; Brix et al., 2017; Duda et al., 2021; Mikkelsen et al., 2016; Near et al., 2014) or within sessions (Brix et al., 2017; O'Gorman et al., 2011). GABA is the principal inhibitory neurotransmitter in the human brain, which is crucial for normal neurological function and plays a significant role in learning, memory, and other cognitive functions (Pasanta et al., 2023). It is shown that the reliability of GABA in the brain, which is indicated by the coefficient of variability (CV), varies between 4% and 15% depending on region, quantification techniques, and the study design (Baeshen et al., 2020; Brix et al., 2017; Duda et al., 2021; Mikkelsen et al., 2016; Near et al., 2014; O'Gorman et al., 2011). Even in the same brain region, however, the CV can be different. For example, one study showed that the CV of GABA among the anterior cingulate cortex (ACC) was 8% using the ratio to total creatine (tCr) while 7.5% with the ratio to water (Duda et al., 2021). To be noticed, all the aforementioned studies used a 3 T MRI scanner and MEGA-PRESS to detect GABA concentration levels. Accordingly, the wide range of the CV is not only because of different brain locations but also due to different quantification indices used in these studies. There are two commonly used quantification indices: the ratio to tCr, and the ratio to water with tissue correction (Gasparovic et al., 2006; Near et al., 2021), which will be referred to as the ratio to WaterTC. We also reported values scaled to raw water (without tissue correction), as WaterRaw. It is suggested that the ratio to water performs better regarding the reliability of the GABA (Duda et al., 2021).

Although much insight has been made regarding GABA, surprisingly, very few studies investigated the reliability of other critical brain metabolites (Kirov et al., 2012; Mullins et al., 2003; Schirmer & Auer, 2000; van Veenendaal et al., 2018), which also play significant roles in brain function. For example, glutamate (Glu) is the most abundant excitatory neurotransmitter, which plays a significant role in brain cognition and neurological development by countering balance with GABA; N-acetyl-aspartate (NAA), and N-acetyl-aspartyl-glutamate (NAAG) are neuromodulators which inhibit the synaptic release of GABA, glutamate, and dopamine, and regulate GABA receptor expression; myo-inositol (mI) is a key precursor of membrane phospho-inositides and phospholipids, and is considered as a glial marker by involving in the cell membrane and myelin sheet structures (Haris et al., 2011; Harris et al., 2017).

Besides the limited numbers, the aforementioned studies also possess their own limitations. For example, two studies focused on the white matter (WM) instead of the gray matter (GM) (Mullins et al., 2003; Schirmer & Auer, 2000), of which the participants of one study were only schizophrenia patients (Mullins et al., 2003). One study only focused on Glu (van Veenendaal et al., 2018) while

the remaining one covered almost a third of the brain (Kirov et al., 2012) which is a deviation from the most widely used method in MRS named single-voxel spectroscopy (Wilson et al., 2019). Furthermore, none of them investigates the different performance between the ratio to tCr and water on the reliability of the MRS measurement. In addition, previous studies generally have very low sampling rates within subjects, where only two measurements were predominantly collected within a time period ranging from the same day to several months (Baeshen et al., 2020; Brix et al., 2017; Duda et al., 2021; Mikkelsen et al., 2016; Mullins et al., 2003; Near et al., 2014; O'Gorman et al., 2011; Schirmer & Auer, 2000; van Veenendaal et al., 2018). The low sampling rate fails to illustrate the details of the time effect on MRS measurements, such as the time-of-day effect manifesting on the functional brain organization (Orban et al., 2020; Vaisvilaite et al., 2022).

Therefore, in this study, three subjects were repeatedly scanned in an MRI scanner across 1 year, resulting in at least 25 MRS sessions for each subject. To be noticed, this study is part of a project named Bergen breakfast scanning club (BBSC) where we also collected structural and functional MRI data (Korbmacher et al., 2023; Wang et al., 2022, 2023). The aims of the study are threefold. First, to investigate the reliability of other brain metabolites other than GABA. GABA is a weak signal that often requires dedicated spectral editing techniques to be resolved and often would be compromised by artifacts and scanner drift. Given that, we posit that the reliability of other metabolites is better than that of GABA. Second, to explore whether the ratio to water can generate more reliable results compared to the ratio to tCr. Third, to explore the time-of-day effect on the levels of the metabolites.

2 | METHODS

2.1 | Participants

Three participants (Table 1) were scanned in this study, which is part of a precision brain mapping project titled BBSC Project (Korbmacher et al., 2023; Wang et al., 2022, 2023). In this project, we aim to chart the individual brain organizations and explore the reliability of the MRI data by repeatedly scanning three subjects for a year including resting state fMRI data, MRS data, and structural brain imaging data (Wang et al., 2022, 2023). The participants were repeatedly scanned twice a week between February 2021 and February 2022 except for

TABLE 1 Basic demographic information of subjects.

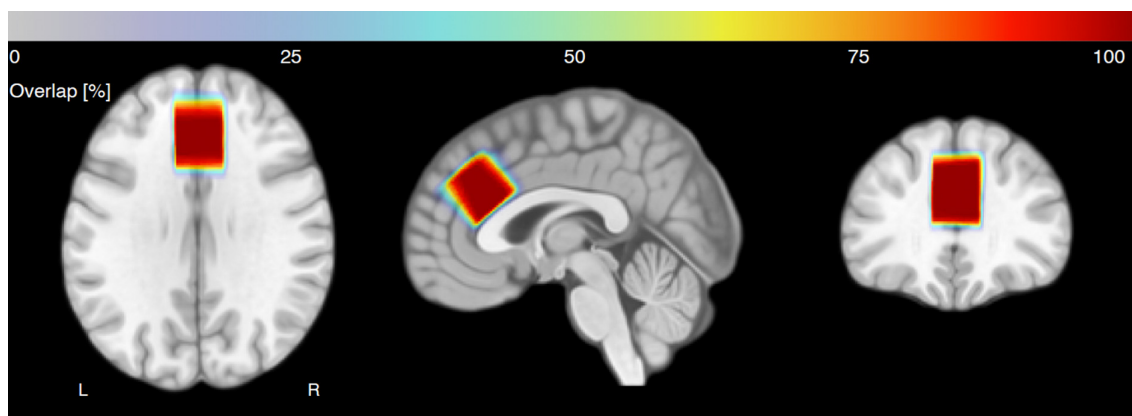|	Sub1	Sub2	Sub3
Gender	Male	Male	Male
Age	31	27	40
Laterality	Right	Right	Right
Regular caffeine consumption?	No	Yes	Yes
Regular nicotine consumption?	No	No	No

FIGURE 1   The location of the MRS voxel. The center of the MNI coordinate is [2, 32, 29]. The color spectrum represents the position overlap percentage across all sessions and all subjects. The figure was generated with Osprey 2.4.0.

the two breaks (June to October 2021, and January 2022). In total, there are 38, 39, and 25 MRS sessions for subjects 1, 2, and 3, respectively. The timing of the individual sessions collection is available in Appendix S1.

They are all male and can speak at least two languages (their native languages and English). To be noticed, subject 1 has been learning another language since January 2021 and did not get COVID-19 during data collection, subject 2 got COVID-19 around December 2021 while subject 3 got COVID-19 around August 2021.

2.2 | Data recording

MRS data collection was embedded in a functional protocol of the BBSC project (Wang et al., 2022, 2023). This protocol includes collecting MRS data, resting state fMRI (rs-fMRI) data, and their anatomical reference T1-weighted (T1w) MRI data, which lasts around 25 min in total. MRS data were collected directly after the T1w MRI and prior to the rs-fMRI scanning to avoid gradient heating effects caused by the demanding fMRI scans. MRI data were collected with a 3 T MR scanner (GE Discovery MR750) with a 32-channel head coil at the Haukeland University Hospital in Bergen, Norway. The minimum reporting standards in MRS (MRSinMRS) (Lin et al., 2021) are provided in Table S1, and the technical details are as follows.

Seven-minute structural T1w image data were acquired using a 3D Fast Spoiled Gradient-Recalled Echo (FSPGR) sequence with the following parameters: 188 contiguous slices acquired, with repetition time (TR) = 6.88 ms, echo time (TE) = 2.95 ms, FA (flip angle) = 12°, slice thickness = 1 mm, in-plane resolution = 1 mm × 1 mm, and field of view (FOV) = 256 mm, with an isotropic voxel size of 1 mm^3.

After that, around 4-min ^{1}H-MRS-spectra data were obtained from the ACC (voxel size 25 × 25 × 25 mm^3) by using a single-voxel point-resolved spectroscopy (PRESS) sequence (TE/TR = 35 ms/1500 ms, 128 repetitions). Unsuppressed water reference spectra (eight repetitions) were acquired automatically after the water-suppressed

metabolite spectra. The MRS voxel is located in the ACC as illustrated in Figure 1, where the center of the MNI coordinate is [2, 32, 29].

2.3 | Data processing

The MRS data were analyzed with Osprey 2.4.0 (Oeltzschner et al., 2020) with the integrated LCModel (linear-combination model) fitting algorithm (Provencher, 1993) based on the MATLAB® (R2022b) platform, which provides an automated and uniform processing pipeline including pre-processing, linear combination modeling, tissue correction, and quantification.

A concise description of the processing pipeline is as follows. The raw data were first aligned and averaged, then fitted using the LCModel which is embedded in the Osprey with the default settings. The spectra were analyzed between 0.5 and 4.0 parts per million (ppm) and the baseline knot spacing parameter DKMNT was set to 0.15 ppm. 17 metabolites and 9 macromolecular/lipids (MM/Lip) were included in the model: ascorbate (Asc), aspartate (Asp), Cr, GABA, glycerophosphocholine (GPC), glutathione (GSH), glutamine (Gln), Glu, mI, lactate (Lac), NAA, NAAG, phosphocholine (PCh), phosphocreatine (PCr), phosphorylethanolamine (PE), scyllo-inositol (sI), taurine (Tau), and –CrCH2; Lip 09, Lip 13a, Lip 13b, Lip 20, MM 09, MM 12, MM 14, MM 17, and MM 20. Individual spectra, mean spectra, and mean fit for each subject are presented in Figure 2.

Before quantification, the brain was segmented into GM, WM, and cerebrospinal fluid (CSF) after coregistration to the structural image with functions from SPM12 (Friston et al., 1994) invoked by Osprey. Quantification of the metabolites was calculated using two different references: as the ratio to tCr and as water-scaled metabolite estimates including WaterRaw and WaterTC according to the *Gasparovic* method (Gasparovic et al., 2006) with default parameters in Osprey 2.4.0. For more details about the processing pipeline, we recommend interested readers refer to the original article about the Osprey (Oeltzschner et al., 2020).
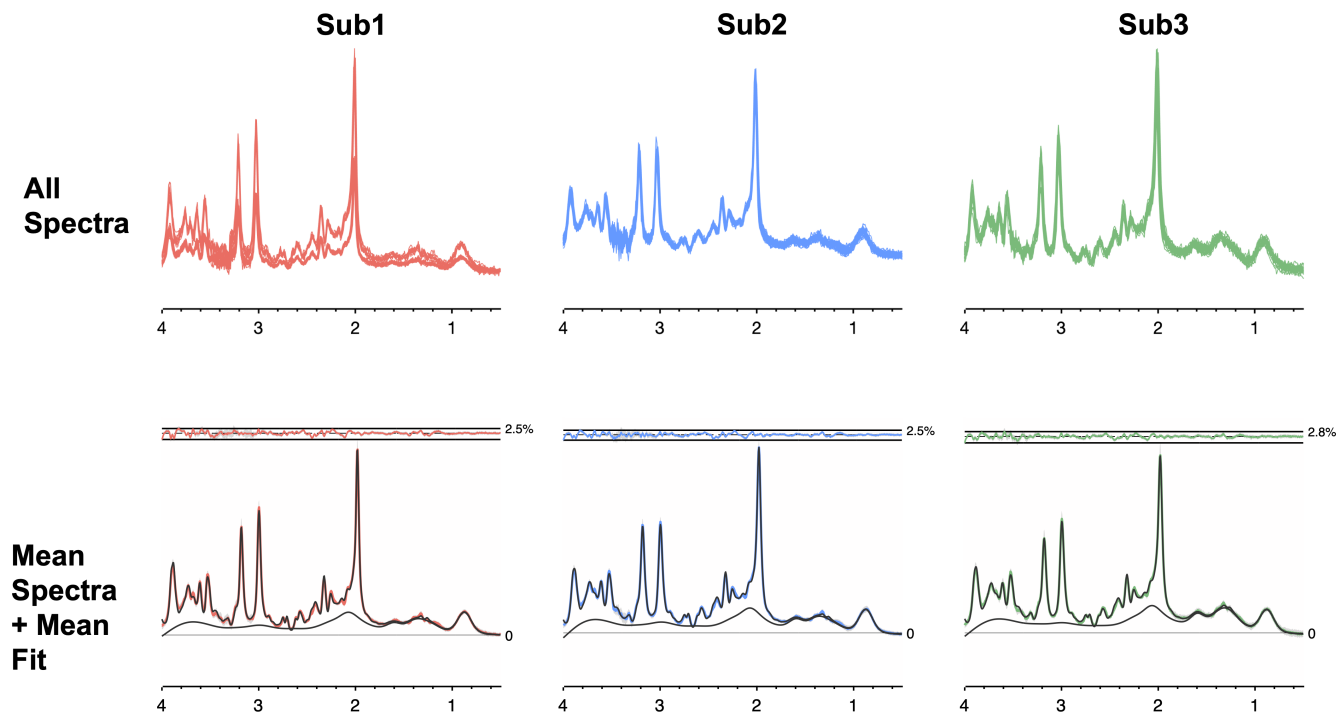
FIGURE 2 The all spectra, mean spectra, and mean fit across all sessions. The upper three panels illustrate the spectra of all sessions for each subject where red, blue, and green represent Sub1, Sub2, and Sub3, respectively. The bottom three panels demonstrate the mean spectra and mean fit for each subject: mean spectra (red/blue/green) ± SD (gray ribbons); mean fit (black), and mean residual (above).

TABLE 2 The %SD (CRLB) values of the selected metabolites across sessions.

	Sub1 (mean ± std)	Sub2 (mean ± std)	Sub3 (mean ± std)
tNAA	2.00 ± 0.00	1.97 ± 0.16	1.96 ± 0.20
NAA	2.42 ± 0.50	2.56 ± 0.50	2.96 ± 0.54
tCr	2.00 ± 0.00	2.03 ± 0.16	2.00 ± 0.00
PCr	7.13 ± 1.02	7.92 ± 1.40	8.00 ± 1.61
Cr	7.03 ± 2.14	9.77 ± 3.64	9.48 ± 4.34
Glx	5.47 ± 0.56	5.36 ± 0.54	5.64 ± 0.49
Glu	5.39 ± 0.59	5.05 ± 0.46	5.2 ± 0.50
ml	4.39 ± 0.55	4.54 ± 0.55	4.56 ± 0.65
tCho	2.66 ± 0.53	2.87 ± 0.41	3.08 ± 0.57

Note: All values are in percentage (%). The smaller the better.

The criteria of %SD < 15% (CRLB: Cramér–Rao lower bounds), which was generated from the LCModel, was set to quality check the data (the lower the better). Only those that consistently under 15% across all sessions will be reported in this study, which are tNAA, NAA; tCr, PCr, Cr; Glx, Glu; ml, tCho as listed in Table 2.

2.4 | Statistical analysis

To evaluate the within-subject reliability of MRS measurements, the coefficient of variance (CV) is used:

$$CV\,(\%) = \left(\frac{std}{mean}\right) \times 100,$$

where std represents the standard deviation of the sample between sessions within each subject while the mean is the average value of the sample between sessions within each subject.

The coefficient of variation can provide the comparison of the variability of data sets with different units of measurement or scales, which has been ubiquitously used in evaluating the reliability of the MRS measurement (Baeshen et al., 2020; Brix et al., 2017; Duda et al., 2021; Kirov et al., 2012; Mikkelsen et al., 2016; Mullins

et al., 2003; Near et al., 2014; O'Gorman et al., 2011; Schirmer & Auer, 2000; van Veenendaal et al., 2018). A lower CV represents a lower degree of variability relative to the mean, which indicates lower relative variability and thus greater precision or consistency in the data; while a higher CV represents a higher degree of variability relative to the mean, which indicates higher relative variability and thus lower precision or consistency.

To compare the reliability levels qualified by the ratio to tCr and water, a one-tailed paired t-test was executed since we expected that the ratio to water would generate smaller CV values. In addition, a two-tailed paired t-test was executed to explore the difference between two water measurements, WaterRaw and WaterTC.

To assess the time-of-day effect, we first separate the dataset into morning and afternoon sessions according to the recording time. In total, there are 22 morning sessions and 16 afternoon sessions in Sub1 (Sub2: 23, 16; Sub3: 15, 10). Then t-test was used to assess the significant difference between these two sessions for each subject. To control multiple comparisons, FDR correction was used. All statistical analyses were done on R 4.2.1 (R Core Team, 2022).

3 | RESULTS

The signal-to-noise ratio (SNR) and the linewidth of the dataset measured from tCr at 3 ppm, which are extracted from Osprey, are 82 ± 11 Hz and 5.85 ± 0.82 Hz, respectively.

3.1 | Coefficients of variability

The scatter plots of the metabolites measured by the ratio of WaterTC are illustrated in Figure 3, and the plots of the measurements of WaterRaw and tCr are demonstrated in Figures S1 and S2. The CVs of the metabolites in the ACC of subjects are shown in Table 3.

It is shown that the variabilities of NAA, tNAA, and tCr are low with CVs ranging from 1.43 to 4.90. In addition, the variabilities of Glu, Glx, mI, and tCho are modest with CVs ranging from 4.26% to 7.89%. Furthermore, the variabilities of Cr and PCr are high with CVs between 14.92% and 23.16%. Lastly, compared with the measurement of the ratio to tCr, the CVs of both the WaterRaw ($t_{23} = -3.19$, $p = .002$, Cohen's $d = 0.65$) and the WaterTC measurement ($t_{23} = -2.9$, $p = .004$, Cohen's $d = 0.59$) are smaller. However, there is no significant difference ($t_{26} = 1.89$, $p = .07$, Cohen's $d = 0.36$) between these two water measurements.

The percentage of GM, WM, and CSF in the MRS voxel, ACC in this study, are described in Table 4.

3.2 | Time of day

The levels of metabolites quantified with the ratio to water (WaterTC, specifically) were used to examine the time-of-day effect since it showed smaller CVs. It is shown that the majority of the metabolites analyzed in this study manifested higher levels in afternoon sessions than that of morning while only PCr and NAA showed an opposite
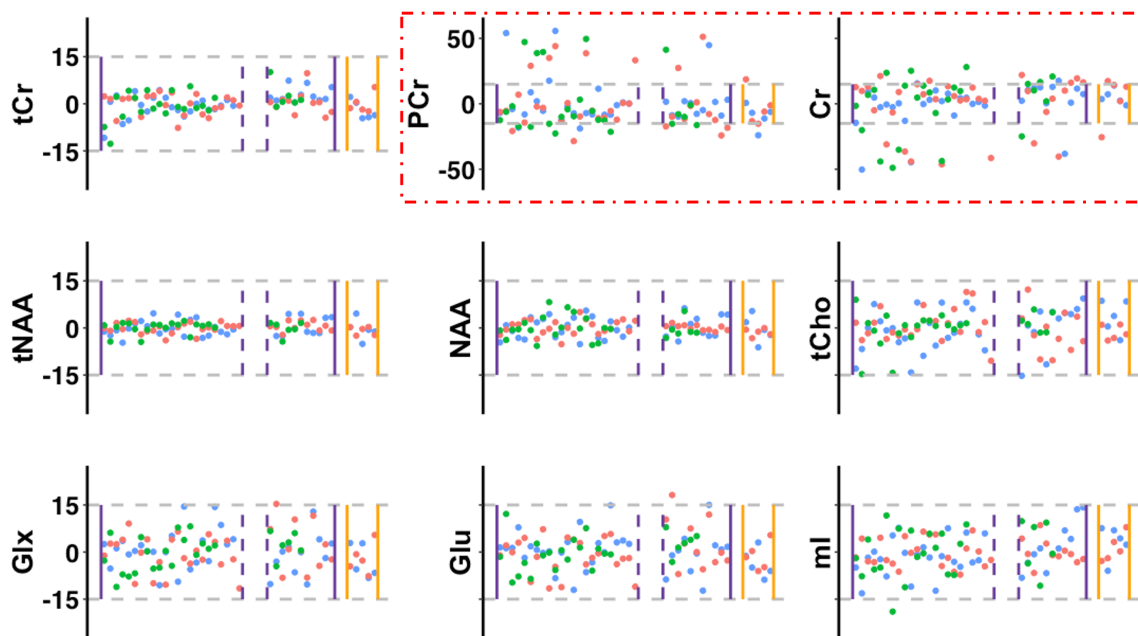

FIGURE 3 The fluctuation of the metabolites across a year. The y-axis represents the percentage change relative to the corresponding mean values while the x-axis represents sessions collected across a year. The unit of the y-axis is percentage (%), and the gray dashed lines represent ±15%. The dots with the colors red, blue, and green represent Sub1, Sub2, and Sub3, respectively. The purple vertical lines represent that the data were collected in the same year where the first and second parts were collected from January to May and from November to December, respectively. The orange vertical lines denote that the data were collected in February of the next year. The fluctuation of PCr and Cr are highlighted in the red box, where the y limits are between ±60%.

TABLE 3 The CVs of metabolites levels in the anterior cingulate cortex.

	Sub1			Sub2			Sub3		
	tCr	WaterRaw	WaterTC	tCr	WaterRaw	WaterTC	tCr	WaterRaw	WaterTC
tNAA	3.74	2.47	2.64	3.79	1.43	1.61	4.46	1.78	1.96
NAA	3.84	3.25	3.38	3.71	1.88	1.99	4.90	3.47	3.61
tCr	N/A	3.45	3.63	N/A	3.40	3.30	N/A	3.96	4.43
PCr	15.89	17.32	17.25	17.78	19.65	19.68	20.76	22.85	23.00
Cr	15.13	14.92	15.06	21.35	20.49	20.47	23.16	22.76	22.83
Glx	7.81	6.42	6.31	7.72	6.32	6.39	7.24	5.62	5.37
Glu	7.47	6.21	6.08	7.61	6.32	6.41	7.66	5.58	5.29
ml	7.20	6.46	6.49	5.10	4.31	4.26	7.12	7.46	7.89
tCho	7.27	6.97	6.78	6.81	5.31	5.39	5.65	4.85	5.24

Note: All values are in percentage (%). The smaller the better.

	Sub1		Sub2		Sub3	
	(Mean ± std)	CV	(Mean ± std)	CV	(Mean ± std)	CV
GM	60.09 ± 0.81	1.35	60.63 ± 1.13	1.86	59.28 ± 1.36	2.29
WM	23.65 ± 1.05	4.44	28.14 ± 1.15	4.09	26.09 ± 1.05	4.02
CSF	16.26 ± 0.81	4.98	11.23 ± 0.59	5.25	14.64 ± 1.02	6.97

TABLE 4 The percentage of different brain tissues in the MRS voxel across sessions.

Note: All values are in percentage (%).

pattern (Tables 5–7). However, we did not find any significant difference between morning and afternoon sessions within the subject after FDR correction (Tables 5–7). The results are similar to WaterRaw as demonstrated in Tables S3–S5. The distributions of the metabolites of the morning and afternoon sessions are illustrated in Figure 4, which were generated with the R package *raincloudplots* v0.2 (Allen et al., 2019).

4 | DISCUSSION

In this study, a longitudinal MRS dataset is constructed, where three participants were repeatedly scanned for a year with an even interval. With the aim of exploring the reliability of MRI measurements, this study specifically for the MRS measurements belongs to the BBSC Project (Korbmacher et al., 2023; Wang et al., 2022, 2023). There are three major findings. First, it is found that MRS measurements showed decent reliability with CVs ranging from 1.43 to 7.89% except for Cr and PCr, which showed lower reliability with CVs around 18%. Second, it is shown that using water as an internal concentration reference resulted in smaller CVs compared to the ratio to tCr. Third, we did not find any significant results of the time-of-day effect on MRS measurements. The potential applications of the results are discussed as follows.

4.1 | Variability of metabolites level

As indicated by Table 3, we found that tNAA, NAA, Glx, Glu, tCho, and ml in the ACC showed good reliability across a year with CVs ranging from around 1.5% to 8%, which is comparable to the CVs generated from previous studies (Baeshen et al., 2020; Kirov et al., 2012; O'Gorman et al., 2011; van Veenendaal et al., 2018). To be noticed, the data collected in previous studies vary on the collection duration, MRI scanner, and brain regions. For example, the data were collected on the same day (O'Gorman et al., 2011; van Veenendaal et al., 2018), within a week (Baeshen et al., 2020), or across 3 years (Kirov et al., 2012). In conjunction with our results, it indicates that the level of metabolites is reasonably reliable across different timescales.

Additionally, the CV values of different metabolites reported in this study agree well with the %SD (CRLB) generated by the LCModel. Generally, smaller %SD (CRLB) values generate more reliable results as demonstrated with smaller CV values except for Cr and PCr. However, it is prudent to consider delineating between the modeling of uncertainty, as indicated by the percentage standard deviation of the Cramér–Rao lower bound (%SD CRLB), and genuine physiological variability. The task of distinguishing the latter from modeling uncertainty can pose challenges, particularly when the magnitude of physiological variability approaches or aligns with that of modeling uncertainty.

As expected, the level of tCr, which is the combination of Cr and PCr, showed greater reliability than Cr and PCr alone, since the constituent components are highly overlapping and difficult to disentangle from a regular PRESS sequence. Accordingly, we suggest that tCr should be emphasized instead of focusing on Cr and PCr individually.

4.2 | The ratio to water

There are several ways to quantify the level of metabolites, of which the ratio to water and tCr are the most widely used ones. Previous

TABLE 5 The different levels of metabolites during morning and afternoon with the ratio to WaterTC in Sub1.

	Morning (mean ± std)	Afternoon (mean ± std)	t	P_fdr
tNAA	18.01 ± 0.44	18.14 ± 0.53	−0.78	0.95
NAA	15.81 ± 0.50	15.76 ± 0.59	0.27	0.96
tCr	13.82 ± 0.49	13.89 ± 0.54	−0.40	0.96
PCr	7.33 ± 1.44	6.52 ± 0.53	2.43	0.11
Cr	6.95 ± 1.27	7.83 ± 0.50	−2.94	0.09
Glx	23.97 ± 1.44	24.34 ± 1.64	−0.72	0.95
Glu	20.40 ± 1.14	20.51 ± 1.41	−0.25	0.96
mI	10.65 ± 0.72	10.99 ± 0.64	−1.56	0.43
tCho	3.62 ± 0.27	3.76 ± 0.19	−1.88	0.26

Note: The unit of the concentration values is a.u.

TABLE 6 The different levels of metabolites during morning and afternoon with the ratio to WaterTC in Sub2.

	Morning (mean ± std)	Afternoon (mean ± std)	t	P_fdr
tNAA	18.83 ± 0.30	18.81 ± 0.32	0.20	0.96
NAA	17.03 ± 0.35	16.98 ± 0.33	0.48	0.95
tCr	12.34 ± 0.40	12.22 ± 0.42	0.96	0.95
PCr	7.31 ± 1.63	6.42 ± 0.59	2.40	0.11
Cr	5.44 ± 1.42	6.20 ± 0.43	−2.39	0.11
Glx	25.02 ± 1.65	26.33 ± 1.31	−2.74	0.09
Glu	21.60 ± 1.28	22.83 ± 1.32	−2.88	0.09
mI	10.87 ± 0.42	10.90 ± 0.53	−0.14	0.96
tCho	3.82 ± 0.19	3.83 ± 0.24	−0.18	0.96

Note: The unit of the concentration values is a.u.

TABLE 7 The different levels of metabolites during morning and afternoon with the ratio to WaterTC in Sub3.

	Morning (mean ± std)	Afternoon (mean ± std)	t	P_fdr
tNAA	18.55 ± 0.38	18.64 ± 0.35	−0.58	0.95
NAA	16.50 ± 0.56	16.48 ± 0.68	0.06	0.98
tCr	13.55 ± 0.73	13.73 ± 0.34	−0.80	0.95
PCr	7.51 ± 1.78	7.34 ± 1.70	0.24	0.96
Cr	6.49 ± 1.39	6.84 ± 1.73	−0.54	0.95
Glx	24.95 ± 1.60	24.94 ± 0.90	0.03	0.98
Glu	21.41 ± 1.32	21.17 ± 0.79	0.58	0.95
mI	10.90 ± 0.79	10.59 ± 0.94	0.84	0.95
tCho	3.58 ± 0.19	3.52 ± 0.19	0.69	0.95

Note: The unit of the concentration values is a.u.

studies have shown that the ratio to water performs better than that of the ratio to tCr regarding the reliability of GABA (Duda et al., 2021). Our results are in line with this finding, where the ratio to water (both WaterRaw and WaterTC) is also better performed for other metabolites. Indeed, we found that a dominant number of previous studies, which investigate the reliability of the level of MRS measurement, used the ratio to water to quantify the metabolites level (Baeshen et al., 2020; Bogner et al., 2010; Brix et al., 2017; Duda et al., 2021; Mikkelsen et al., 2016; O'Gorman et al., 2011; van Veenendaal et al., 2018) instead of the ratio to tCr as the internal concentration reference (Near et al., 2014; Schirmer & Auer, 2000). More importantly, a recent consensus recommendation suggests using the ratio of water (Near et al., 2021). Regarding WaterRaw and WaterTC, as relaxation rates differ between the tissue classes and CSF typically contains no observable metabolites (hence, metabolite signal only originates from the GM and WM tissue portions within the voxel), it is advantageous to employ the WaterTC ratio. Besides, our findings indicate that the CV values for WaterTC do not exhibit significant
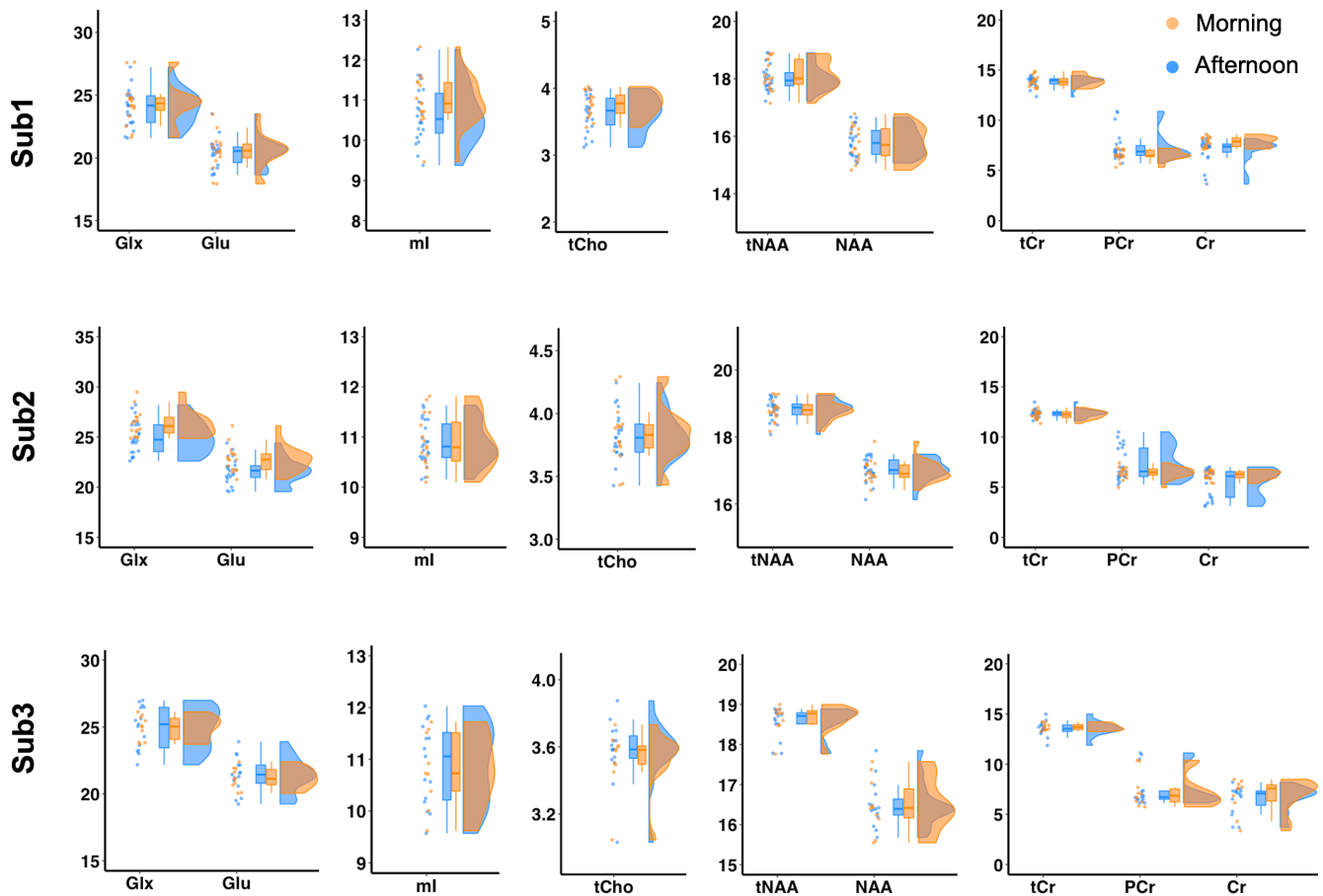
FIGURE 4 The distributions of the metabolites concentration. The unit of the *y*-axis is an arbitrary unit (a.u.) and the values were calculated as the ratio to WaterTC.

disparities when compared to WaterRaw. Accordingly, we advocate the ratio to water with full tissue correction as the internal concentration reference in MRS studies.

However, we noticed that the CV values of WaterTC are slightly larger than the WaterRaw although not significantly different. This is expected or reasonable since the method of WaterTC depends on the variability of the fractions of different brain tissues in the MRS voxel (Gasparovic et al., 2006). This underscores the significance of ensuring proper alignment and overlap of MRS localization across various sessions or subjects. Additionally, therefore, we recommend including information on the proportions of different brain tissues when applying the WaterTC method.

4.3 | The time-of-day effect

Despite there being evidence showing that different data collection time points could affect the functional brain organizations (Orban et al., 2020; Vaisvilaite et al., 2022), we did not find this effect on the levels of metabolites. Correspondingly, it corroborates the indication that the MRS measurement is reasonably stable.

4.4 | Limitations

Before making any conclusions, some limitations should be articulated. First, phantom data can be used to assess the absolute estimate of metabolite concentrations and indicate the reliable performance of the MRI scanner (Brix et al., 2017; van Veenendaal et al., 2018). Since we found that the MRS measurement is reasonably reliable, we believe it will be much more stable when ruling out the variance contributions from the measurement itself. Second, only the default LCM algorithm was utilized which could affect the conclusion, since different LCM algorithms could provide different quantifications (Zollner et al., 2021). Third, the sample size is relatively small compared to other conventional studies. However, the present new data collection method, where the same participants are scanned repeatedly for a period of time, does not necessarily need a very large sample. But it can provide invaluable insights that conventional studies cannot offer (Korbmacher et al., 2023; Wang et al., 2022, 2023). Lastly, the absence of female participants may affect the conclusions of this study, since there is evidence suggesting that the menstrual cycle could influence the GABA level (Harada et al., 2011) and other metabolites level (Hjelmervik et al., 2018).

5 | CONCLUSION

In summary, it is shown that the MRS measurement is quite reliable in detecting the concentration of the metabolites even across a year. In addition, the ratio to water with full tissue correction is suggested for use as the internal concentration reference.

CONFLICT OF INTEREST STATEMENT

No declared conflicts of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in Github at https://github.com/MengYunWang/BBSC/tree/main/MRS/Data.

ORCID

Max Korbmacher 🔟 https://orcid.org/0000-0002-8113-2560
Karsten Specht 🔟 https://orcid.org/0000-0002-9946-3704

REFERENCES

Allen, M., Poggiali, D., Whitaker, K., Marshall, T. R., & Kievit, R. A. (2019). Raincloud plots: A multi-platform tool for robust data visualization. *Wellcome Open Research*, 4, 63. https://doi.org/10.12688/wellcomeopenres.15191.1

Baeshen, A., Wyss, P. O., Henning, A., O'Gorman, R. L., Piccirelli, M., Kollias, S., & Michels, L. (2020). Test–retest reliability of the brain metabolites GABA and Glx with JPRESS, PRESS, and MEGA-PRESS MRS sequences in vivo at 3T. *Journal of Magnetic Resonance Imaging*, 51(4), 1181–1191. https://doi.org/10.1002/jmri.26921

Bogner, W., Gruber, S., Doelken, M., Stadlbauer, A., Ganslandt, O., Boettcher, U., Trattnig, S., Doerfler, A., Stefan, H., & Hammen, T. (2010). In vivo quantification of intracerebral GABA by single-voxel 1H-MRS—How reproducible are the results? *European Journal of Radiology*, 73(3), 526–531. https://doi.org/10.1016/j.ejrad.2009.01.014

Brix, M. K., Ersland, L., Hugdahl, K., Dwyer, G. E., Gruner, R., Noeske, R., Beyer, M. K., & Craven, A. R. (2017). Within- and between-session reproducibility of GABA measurements with MR spectroscopy. *Journal of Magnetic Resonance Imaging*, 46(2), 421–430. https://doi.org/10.1002/jmri.25588

Duda, J. M., Moser, A. D., Zuo, C. S., Du, F., Chen, X., Perlo, S., Richards, C. E., Nascimento, N., Ironside, M., Crowley, D. J., Holsen, L. M., Misra, M., Hudson, J. I., Goldstein, J. M., & Pizzagalli, D. A. (2021). Repeatability and reliability of GABA measurements with magnetic resonance spectroscopy in healthy young adults. *Magnetic Resonance in Medicine*, 85(5), 2359–2369. https://doi.org/10.1002/mrm.28587

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J. P., Frith, C. D., & Frackowiak, R. S. (1994). Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping*, 2(4), 189–210.

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. *Magnetic Resonance in Medicine*, 55(6), 1219–1226. https://doi.org/10.1002/mrm.20901

Harada, M., Kubo, H., Nose, A., Nishitani, H., & Matsuda, T. (2011). Measurement of variation in the human cerebral GABA level by in vivo MEGA-editing proton MR spectroscopy using a clinical 3 T instrument and its dependence on brain region and the female menstrual cycle. *Human Brain Mapping*, 32(5), 828–833. https://doi.org/10.1002/hbm.21086

Haris, M., Cai, K. J., Singh, A., Hariharan, H., & Reddy, R. (2011). In vivo mapping of brain myo-inositol. *NeuroImage*, 54(3), 2079–2085. https://doi.org/10.1016/j.neuroimage.2010.10.017

Harris, A. D., Saleh, M. G., & Edden, R. A. E. (2017). Edited H-1 magnetic resonance spectroscopy in vivo: Methods and metabolites. *Magnetic Resonance in Medicine*, 77(4), 1377–1389. https://doi.org/10.1002/mrm.26619

Hjelmervik, H., Hausmann, M., Craven, A. R., Hirnstein, M., Hugdahl, K., & Specht, K. (2018). Sex- and sex hormone-related variations in energy-metabolic frontal brain asymmetries: A magnetic resonance spectroscopy study. *NeuroImage*, 172, 817–825. https://doi.org/10.1016/j.neuroimage.2018.01.043

Kirov, I. I., George, I. C., Jayawickrama, N., Babb, J. S., Perry, N. N., & Gonen, O. (2012). Longitudinal inter- and intra-individual human brain metabolic quantification over 3 years with proton MR spectroscopy at 3 T. *Magnetic Resonance in Medicine*, 67(1), 27–33. https://doi.org/10.1002/mrm.23001

Korbmacher, M., Wang, M. Y., Eikeland, R., Buchert, R., Andreassen, O. A., Espeseth, T., Leonardsen, E., Westlye, L. T., Maximov, I. I., & Specht, K. (2023). Considerations on brain age predictions from repeatedly sampled data across time. *Brain and Behavior*, 13, e3219. https://doi.org/10.1002/brb3.3219

Lin, A. D., Andronesi, O., Bogner, W., Choi, I. Y., Coello, E., Cudalbu, C., Juchem, C., Kemp, G. J., Kreis, R., Krššák, M., Lee, P., Maudsley, A. A., Meyerspeer, M., Mlynarik, V., Near, J., Öz, G., Peek, A. L., Puts, N. A., Ratai, E.-M., ... Experts' Working Group on Reporting Standards for MR Spectroscopy. (2021). Minimum reporting standards for in vivo magnetic resonance spectroscopy (MRSinMRS): Experts' consensus recommendations. *NMR in Biomedicine*, 34(5), e4484. https://doi.org/10.1002/nbm.4484

Mikkelsen, M., Singh, K. D., Sumner, P., & Evans, C. J. (2016). Comparison of the repeatability of GABA-edited magnetic resonance spectroscopy with and without macromolecule suppression. *Magnetic Resonance in Medicine*, 75(3), 946–953. https://doi.org/10.1002/mrm.25699

Mullins, P. G., Rowland, L., Bustillo, J., Bedrick, E. J., Lauriello, J., & Brooks, W. M. (2003). Reproducibility of 1H-MRS measurements in schizophrenic patients. *Magnetic Resonance in Medicine*, 50(4), 704–707. https://doi.org/10.1002/mrm.10598

Near, J., Harris, A. D., Juchem, C., Kreis, R., Marjanska, M., Oz, G., Slotboom, J., Wilson, M., & Gasparovic, C. (2021). Preprocessing, analysis and quantification in single-voxel magnetic resonance spectroscopy: experts' consensus recommendations. *NMR in Biomedicine*, 34(5), e4257. https://doi.org/10.1002/nbm.4257

Near, J., Ho, Y. C. L., Sandberg, K., Kumaragamage, C., & Blicher, J. U. (2014). Long-term reproducibility of GABA magnetic resonance spectroscopy. *NeuroImage*, 99, 191–196. https://doi.org/10.1016/j.neuroimage.2014.05.059

Oeltzschner, G., Zollner, H. J., Hui, S. C. N., Mikkelsen, M., Saleh, M. G., Tapper, S., & Edden, R. A. E. (2020). Osprey: Open-source processing, reconstruction & estimation of magnetic resonance spectroscopy data. *Journal of Neuroscience Methods*, 343, 108827. https://doi.org/10.1016/j.jneumeth.2020.108827

O'Gorman, R. L., Michels, L., Edden, R. A., Murdoch, J. B., & Martin, E. (2011). In vivo detection of GABA and glutamate with MEGA-PRESS: Reproducibility and gender effects. *Journal of Magnetic Resonance Imaging*, *33*(5), 1262–1267. https://doi.org/10.1002/jmri.22520

Orban, C., Kong, R., Li, J. W., Chee, M. W. L., & Yeo, B. T. T. (2020). Time of day is associated with paradoxical reductions in global signal fluctuation and functional connectivity. *PLoS Biology*, *18*(2), e3000602. https://doi.org/10.1371/journal.pbio.3000602

Oz, G., Alger, J. R., Barker, P. B., Bartha, R., Bizzi, A., Boesch, C., Bolan, P. J., Brindle, K. M., Cudalbu, C., Dinçer, A., Dydak, U., Emir, U. E., Frahm, J., González, R. G., Gruber, S., Gruetter, R., Gupta, R. K., Heerschap, A., Henning, A., … MRS Consensus Group. (2014). Clinical proton MR spectroscopy in central nervous system disorders. *Radiology*, *270*(3), 658–679. https://doi.org/10.1148/radiol.13130531

Pasanta, D., He, J. L., Ford, T., Oeltzschner, G., Lythgoe, D. J., & Puts, N. A. (2023). Functional MRS studies of GABA and glutamate/Glx—A systematic review and meta-analysis. *Neuroscience & Biobehavioral Reviews*, *144*, 104940. https://doi.org/10.1016/j.neubiorev.2022.104940

Provencher, S. W. (1993). Estimation of metabolite concentrations from localized in vivo proton NMR spectra. *Magnetic Resonance in Medicine*, *30*(6), 672–679. https://doi.org/10.1002/mrm.1910300604

R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Retrieved from. https://www.R-project.org/

Schirmer, T., & Auer, D. P. (2000). On the reliability of quantitative clinical magnetic resonance spectroscopy of the human brain. *NMR in Biomedicine*, *13*(1), 28–36. https://doi.org/10.1002/(Sici)1099-1492(200002)13:1<28::Aid-Nbm606>3.0.Co;2-L

Vaisvilaite, L., Hushagen, V., Gronli, J., & Specht, K. (2022). Time-of-day effects in resting-state functional magnetic resonance imaging: Changes in effective connectivity and blood oxygenation level dependent signal. *Brain Connectivity*, *12*(6), 515–523. https://doi.org/10.1089/brain.2021.0129

van Veenendaal, T. M., Backes, W. H., van Bussel, F. C. G., Edden, R. A. E., Puts, N. A. J., Aldenkamp, A. P., & Jansen, J. F. A. (2018). Glutamate quantification by PRESS or MEGA-PRESS: Validation, repeatability, and concordance. *Magnetic Resonance Imaging*, *48*, 107–114. https://doi.org/10.1016/j.mri.2017.12.029

Wang, M. Y., Korbmacher, M., Eikeland, R., & Specht, K. (2022). Deep brain imaging of three participants across 1 year: The Bergen breakfast scanning club project. *Frontiers in Human Neuroscience*, *16*, 1021503. https://doi.org/10.3389/fnhum.2022.1021503

Wang, M. Y., Korbmacher, M., Eikeland, R., & Specht, K. (2023). The Bergen breakfast scanning Club dataset: A deep brain imaging dataset. *bioRxiv*. https://doi.org/10.1101/2023.05.30.542072

Wilson, M., Andronesi, O., Barker, P. B., Bartha, R., Bizzi, A., Bolan, P. J., Brindle, K. M., Choi, I.-Y., Cudalbu, C., Dydak, U., Emir, U. E., Gonzalez, R. G., Gruber, S., Gruetter, R., Gupta, R. K., Heerschap, A., Henning, A., Hetherington, H. P., Huppi, P. S., … Howe, F. A. (2019). Methodological consensus on clinical proton MRS of the brain: Review and recommendations. *Magnetic Resonance in Medicine*, *82*(2), 527–550. https://doi.org/10.1002/mrm.27742

Zollner, H. J., Povazan, M., Hui, S. C. N., Tapper, S., Edden, R. A. E., & Oeltzschner, G. (2021). Comparison of different linear-combination modeling algorithms for short-TE proton spectra. *NMR in Biomedicine*, *34*(4), e4482. https://doi.org/10.1002/nbm.4482

SUPPORTING INFORMATION

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

How to cite this article: Wang, M.-Y., Korbmacher, M., Eikeland, R., Craven, A. R., & Specht, K. (2024). The intra-individual reliability of ^{1}H-MRS measurement in the anterior cingulate cortex across 1 year. *Human Brain Mapping*, *45*(1), e26531. https://doi.org/10.1002/hbm.26531