

Quantification of magnetic resonance spectroscopy data using a combined reference: Application in typically developing infants

Ryan J. Larsen¹  | Borjan Gagoski^{2,3,4} | Sarah U. Morton^{4,5} | Yangming Ou^{2,3,4} |
 Rutvi Vyas² | Jonathan Litt^{4,5,6} | P. Ellen Grant^{2,3,4,5} | Bradley P. Sutton^{1,7} 

¹Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA

²Fetal-Neonatal Neuroimaging and Developmental Science Center, Boston Children's Hospital, Boston, Massachusetts, USA

³Department of Radiology, Boston Children's Hospital, Boston, Massachusetts, USA

⁴Harvard Medical School, Boston, Massachusetts, USA

⁵Division of Newborn Medicine, Boston Children's Hospital, Boston, Massachusetts, USA

⁶Department of Neonatology, Beth Israel Deaconess Medical Center, Boston, Massachusetts, USA

⁷Department of Bioengineering, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA

Correspondence

Ryan J. Larsen, Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, IL, 61801, USA.

Email: larsen@illinois.edu.

Funding information

Abbott Nutrition

Quantification of proton magnetic resonance spectroscopy (¹H-MRS) data is commonly performed by referencing the ratio of the signal from one metabolite, or metabolite group, to that of another, or to the water signal. Both approaches have drawbacks: ratios of two metabolites can be difficult to interpret because study effects may be driven by either metabolite, and water-referenced data must be corrected for partial volume and relaxation effects in the water signal. Here, we introduce combined reference (CRef) analysis, which compensates for both limitations. In this approach, metabolites are referenced to the combined signal of several reference metabolites or metabolite groups. The approach does not require the corrections necessary for water scaling and produces results that are less sensitive to the variation of any single reference signal, thereby aiding the interpretation of results. We demonstrate CRef analysis using 202 ¹H-MRS acquisitions from the brains of 140 infants, scanned at approximately 1 and 3 months of age. We show that the combined signal of seven reference metabolites or metabolite groups is highly correlated with the water signal, corrected for partial volume and relaxation effects associated with cerebral spinal fluid. We also show that the combined reference signal is equally or more uniform across subjects than the reference signals from single metabolites or metabolite groups. We use CRef analysis to quantify metabolite concentration changes during the first several months of life in typically developing infants.

KEY WORDS

combined reference, GABA, infant development, magnetic resonance spectroscopy, quantification, water scaling

1 | INTRODUCTION

Proton magnetic resonance spectroscopy (¹H-MRS) provides concentration measurements of metabolites involved in a variety of biomolecular processes, including signaling, energy production and membrane turnover.¹ In infant populations, ¹H-MRS is used clinically to assess the severity

Abbreviations used: Cho, combined glycerophosphocholine and phosphocholine; CRef, combined reference; CSF, cerebral spinal fluid; FWHM, full width at half maximum; GABA, γ -Aminobutyric acid; Glu, glutamate; Glx, combined glutamate and glutamine; GSH, glutathione; Ins, myo-inositol; Lip, lipids; M1, Month 1; M3, Month 3; MEGA, MEGhcher-GArwood; MOI, metabolite group of interest; NAA, N-acetylaspartate; NAAG, N-acetylspartylglutamate; PCr, phosphocreatine; PMA, postmenstrual age; Tau, taurine; tCr, combined creatine and phosphocreatine; tNAA, combined NAA and NAAG; WSc, water-scaling method.

Ryan J. Larsen and Borjan Gagoski contributed equally to this work

of hypoxic ischemic encephalopathy, screen for inherited metabolic disorders, assess cerebral involvement in hepatic encephalopathy, and assess the severity of injury in other events such as drowning.² Metabolite concentrations are also dynamic during the rapid brain development that occurs in the months following birth and are therefore helpful in characterizing typical or atypical early brain development.^{3,4} Metabolites, or metabolite groups, commonly studied using ¹H-MRS include *N*-acetylaspartate (NAA), combined glycerophosphocholine and phosphocholine (Cho), combined creatine and phosphocreatine (tCr), myo-inositol (Ins), and combined glutamate and glutamine (Glx). Concentrations of the neurotransmitter γ -Aminobutyric acid (GABA) can also be measured using specialized methods, such as the MEGHcher-GArwood (MEGA) sequence.⁵ GABA is of particular interest as it plays a critical role in the early development and experience-related modulation of neuronal connectivity.^{6,7}

The use of ¹H-MRS data requires that the raw metabolite signal be referenced to a second signal, thereby removing spurious variance from factors, such as coil loading, which scale all signals.^{8–10} A common reference is the water signal, which can be used to calculate molarity or molal concentration values.¹¹ However, the use of water scaling is complicated by the fact that it often includes signal from the cerebral spinal fluid (CSF). Typically, the metabolite signal is assumed to originate only from tissue,^{12–14} and the water signal is adjusted to not include the contribution of the CSF signal, thereby providing estimates of metabolite concentration in tissue. This removal requires knowledge of the proportion of the water signal that originates from the CSF. This information can be gained from supplemental measurements that distinguish CSF from brain tissue water based on differences in their relaxation properties.^{15–18} Knowledge of relaxation constants gained from these experiments can also be used to correct for relaxation effects in the water reference signal. Corrections for partial volume and relaxation effects decrease spurious variance, thereby increasing statistical power.^{19,20} Such corrections are particularly important during the first months of life, when relaxation constants change dramatically with development.¹⁸ However, the additional time required for these measurements may be prohibitive in some studies.

Some studies avoid the use of the water signal by using another metabolite signal as the reference signal, such as tCr. Unlike the water signal, a metabolite signal has little contribution from CSF, causing the metabolite ratios to be insensitive to the CSF content of the scan area. Nevertheless, the reference metabolite concentrations can depend on the relative amounts of gray and white matter within the scan area.¹⁸ However, these effects can be mitigated by consistent within-study voxel placement, or statistically corrected by performing regressions across subjects to remove average tissue dependence.²¹

Ideally, a reference metabolite group, consisting of one or more similar metabolites, will exhibit a uniform concentration across subjects or will at least be independent of study effects. However, when only simple ratios are used, this independence is difficult to assess, and such an assessment is rarely performed. Therefore, for many studies it is not known which signal in a ratio is driving a given effect, causing study results to be ambiguous. This ambiguity limits the usefulness of studies and can lead to false interpretations of their results. This deficiency can be mitigated by analyzing multiple combinations of simple ratios²²; however, such analysis is complicated by the fact that many ratios can be generated from a few metabolite signal groups (e.g. 15 combinations can be generated from six groups). What is required is a lower dimensional framework for distilling information from the relative intensities of multiple metabolite signals.

Here, we propose the combined reference (CRef) analysis, in which metabolites are referenced to a weighted sum of signals from multiple metabolite groups. The advantage of this approach is that it allows for relationships among multiple metabolite groups to be probed using fewer ratios, and is not dependent on the uniformity of any single reference signal. Previous work has used a combined signal for referencing²³; however, there have been no reports comparing the uniformity of combined reference across subjects with that of the single references, such as tCr. Here we evaluate single versus combined metabolite references by correlating both with the “gold standard” reference signal, which is the water signal that has been corrected for partial volume and relaxation effects.

We first present the theoretical framework for quantification using CRef analysis, followed by validation using data collected from 140 infants aged 10–156 days (0.3–5.2 months) scanned one or two times. We demonstrate the robustness of this analysis by showing that combined reference ratios are numerically consistent, whether or not water scaling is used. To validate CRef analysis, we compare it with molal concentration values calculated from water-scaled data that have been corrected for relaxation and partial volume effects, using a standard approach that we refer to as the corrected water-scaling method (WSc).^{17,18} We demonstrate that the combined reference signal exhibits excellent consistency with the corrected water signal. We also show that the combined metabolite reference has similar or better uniformity than single reference metabolites. These results imply that the combined reference may be useful for estimating molal concentration when the water signal corrections are not available. We demonstrate that this estimation can be performed by rescaling the combined reference ratio to units of molal concentration using expectation values of reference metabolite concentrations, and we evaluate the accuracy of this estimate.

2 | THEORY

A simple multiple-reference metabolite ratio, R_{CRef_simp} , is formed by dividing the signal, S_{MOI} , from the metabolite group of interest (MOI), by the average of signals, S_{Ref} , from the reference metabolites,

$$R_{CRef_simp} = \frac{S_{MOI}}{\left(\sum S_{Ref} / N\right)} = N \frac{S_{MOI}}{\sum S_{Ref}}, \quad (1)$$

where N is the number of reference metabolites and the summation is performed over all signals from the reference metabolites. A metabolite signal can originate from a single metabolite or from a metabolite group; we take the metabolite signal to be the result of an integration of either the metabolite spectrum or a fit to the metabolite spectrum, and we assume that it has been corrected for the number of protons per molecule, such that S is proportional to the number of molecules. As will be demonstrated in the Results section, an issue with R_{CRef_simp} is that metabolite signals of greater variance will contribute more variance to R_{CRef_simp} , causing it to be more sensitive to those signals. To equalize the variance sensitivity of the reference signal among all reference metabolites we propose to divide each reference signal, S_{Ref} , by the standard deviation value, S_{Ref_SD} , of that signal, calculated across all subjects. This normalization is similar to the practice of Z-scoring variables before performing principle component analysis²⁴; in both cases the purpose is to equalize the contributed variance among the variables. The resulting combined reference ratio is

$$R_{CRef} = N \frac{S_{MOI}/S_{M_SD}}{\sum(S_{Ref}/S_{Ref_SD})}, \quad (2)$$

where S_{M_SD} is a representative standard deviation value yet to be specified. A challenge of using R_{CRef} is that calculation of S_{Ref_SD} requires measurements from multiple subjects. This could be a limiting factor for small studies or single measurements. In principle this can be solved by using standard deviation values measured from comparable studies; however, the absolute scale of S_{Ref_SD} varies among sites and measurement conditions because it is sensitive to a variety of factors such as the scanner hardware and the use of water scaling. Our solution to the scaling problem is to choose S_{M_SD} as the harmonic mean of S_{Ref_SD} values,

$$\frac{1}{S_{M_SD}} = \frac{1}{N} \sum \frac{1}{S_{Ref_SD}}. \quad (3)$$

With this definition, the reference signals are multiplied by a series of weighting factors,

$$w_{Ref} = \frac{S_{M_SD}}{S_{Ref_SD}} = \frac{\sum S_{Ref_SD}}{S_{Ref_SD}}, \quad (4)$$

which sum to unity, such that the combined ratio is given by

$$R_{CRef} = \frac{S_{MOI}}{\sum(w_{Ref} S_{Ref})}. \quad (5)$$

Each weighting coefficient w_{Ref} is the inverse of the fractional contribution of a metabolite's standard deviation to the sum of the standard deviations of the reference metabolites. With this approach, R_{CRef} can be calculated without knowledge of the absolute scale of S_{Ref_SD} . Values of the weighting factors, w_{Ref} , which we hereafter denote simply as $w=w_{Ref}$ for the sake of brevity, can be obtained from normative datasets, thereby enabling calculations of R_{CRef} in small studies or individual measurements.

Because CRef analysis involves ratios of metabolites, it can be calculated without the use of a water-reference signal. However, in ¹H-MRS the water signal from a reference scan is useful for performing coil combination and eddy current correction. In such cases, it may be expedient to calculate R_{CRef} by substituting the raw metabolite signals, S , in Equation 2 with water-scaled metabolite signals, A . To ensure consistency among studies, it is important to determine whether values of R_{CRef} are sensitive to this choice. In the Results section of this paper, we demonstrate that values of R_{CRef} are numerically consistent between raw and water-scaled data, provided that the postmenstrual age (PMA)-dependent variance is removed from the signals before the calculation of w factors.

Although R_{CRef} depends little on water scaling, it is dependent on other choices required for implementation, including the choice of reference metabolites. By default, we recommend the use of the largest possible number of independent metabolites, or metabolite groups, which can be measured with reasonable confidence. For clarity of interpretation we also recommend using reference metabolites, or metabolite groups, which are independent of the MOI; for example, not using Glutamate + Glutamine as a reference when Glutamate is the MOI. This would also rule out using the same metabolite as both the MOI and a reference.

When a metabolite is referenced to multiple metabolites in a nonbiased way, it is reasonable to expect that for a large number of reference metabolites, values of R_{CRef} may be statistically consistent with molality. In this case, R_{CRef} may be used as a surrogate of measurements of molality. To test this hypothesis, we calculate the statistical consistency between the references signals of R_{CRef} and those of molal concentration values, C_{Wsc} . For R_{CRef} , the reference signal is $\sum(wS_{Ref})$ and for C_{Wsc} the reference is the water signal, corrected for relaxation and partial volume effects. The good statistical consistency we demonstrate between these measures motivates an optional rescaling of R_{CRef} to an approximation of

the molal concentration, which we denote as C_{CRef} . This is similar to the practice of expressing ratios to tCr in units of concentration by multiplying the ratio by the expected concentration of tCr.²⁵ By extension, C_{CRef} is given by

$$C_{CRef} = \left(\sum w C_{WSc_Ref_E} \right) R_{CRef}, \quad (6)$$

where the terms $C_{WSc_Ref_E}$ are PMA-dependent expectation values of C_{WSc} , taken from measurements of C_{WSc} performed on the reference metabolites. In the Results section of the paper, we assess the accuracy of C_{CRef} for estimating C_{WSc} .

3 | METHODS

3.1 | Recruitment

This study was approved by the Institutional Review Board at Boston Children's Hospital. Families of infants were recruited at Beth Israel Deaconess Medical Center and Brigham and Women's Hospital, and written informed consent of the parent or legal guardian was obtained. This protocol corresponds to clinicaltrials.gov Identifier: NCT02058225 ("Developing Advanced MRI Methods for Detecting the Impact of Nutrients on Infant Brain Development"; <https://clinicaltrials.gov/ct2/show/NCT02058225>). ¹H-MRS data were obtained from 140 singleton infants born to mothers with no known medical conditions or complications during pregnancy. Of the 140 infants in the study, 75 were male and 65 were female. Families were asked to participate in two scan sessions, one approximately 1 month after birth (M1), and a second approximately 3 months after birth (M3). Data were obtained from 202 scan sessions; 110 at M1, and 92 at M3. The average infant age for the M1 scan was 27 days (range: 10–52 days, standard deviation: 9 days). The average age for the M3 scan was 106 days (range: 77–156 days, standard deviation: 17 days). Development is reported as a function of PMA, given by the sum of gestational ages at birth (mean: 39.2 weeks, range: 35–42 weeks, standard deviation: 1.3 weeks) and chronological age at the time of the scan.

3.2 | ¹H-MRS data acquisition

Scans were either performed at Boston Children Hospital's location in Boston, MA, or Waltham, MA. All scans were performed on a Siemens TIM Trio 3-T system equipped with gradients with a maximum amplitude of 40 mT/m and maximum slew rate of 200 mT/m/ms (software: syngo MR B17, implementation VB17B). Experiments were performed using a standard 32-channel head coil array. Infants were fed before the scan and swaddled in a blanket. After naturally falling asleep, neonates were settled in a special infant positioning vacuum bag, or padded with cushions, to minimize motion. Infants who woke up were taken out of the scanner.

Our scan protocol consisted of a multiecho MPRAGE acquisition enhanced with volumetric EPI navigators for prospective motion correction²⁶ (voxel size = 1 mm³, FOV 160 mm², GRAPPA—GeneRalized Autocalibrating Partial Parallel Acquisition—acceleration factor R = 2; TR = 2350 ms, TI = 1450 ms, flip angle of 7°, for an overall scan time of 4 min 51 s). This scan was used to position a single voxel of 30 x 40 x 20 mm³ superior to the corpus callosum. The voxel was positioned as superior as possible such that the voxel did not touch the skull to avoid lipid contamination (Figure 1). Metabolite spectra were obtained with a MEGA sequence⁵ that included a Point-RESolved Spectroscopy (PRESS) excitation and a Gaussian editing pulse of bandwidth 44 Hz that alternated between an "edit-on" condition at 1.9 ppm and an "edit-off" condition at 7.5 ppm (TE/TR = 68/2000 ms, water suppression, no outer volume suppression, 1024 (nominal)/2048 (raw) complex data points in the FID, 2000 (nominal)/4000 (raw) Hz sampling, frequency offset: -1.7 ppm). Initially, this sequence was performed with 144 averages for all subjects. However, to minimize the effect of changes in center frequency due to drift and head motion, we revised the protocol to include one to three scans of 48 averages, with a manual readjustment of center frequency before each scan. These scans were immediately followed by an identical acquisition with no water suppression (TR = 30 s, one average, offset frequency of -1.7 ppm). Of 202 scan sessions, 13 did not include a nonwater-suppressed scan and were excluded from all further analysis.

To further characterize the water signal we also performed single-voxel PRESS acquisitions with no water suppression, and TE = 30, 70, 150, 250, 600, 900, 1100, 1300 and 1500 ms (TR = 30 s, no outer volume suppression, 512 (nominal)/1024 (raw) complex data points in the FID, 1250 (nominal)/2500 (raw) Hz sampling, frequency offset: 0 ppm, one average). Because of infant waking and removal, some datasets lacked a portion of these acquisitions.

Ideally, all of the nonwater-suppressed acquisitions would have been performed at an offset frequency of 0 ppm. However, the water reference scan used to scale the metabolites was performed with an offset frequency of -1.7 ppm, as previously indicated. The effect of this error on the accuracy of the water signal corrections is reviewed in the Discussion section.

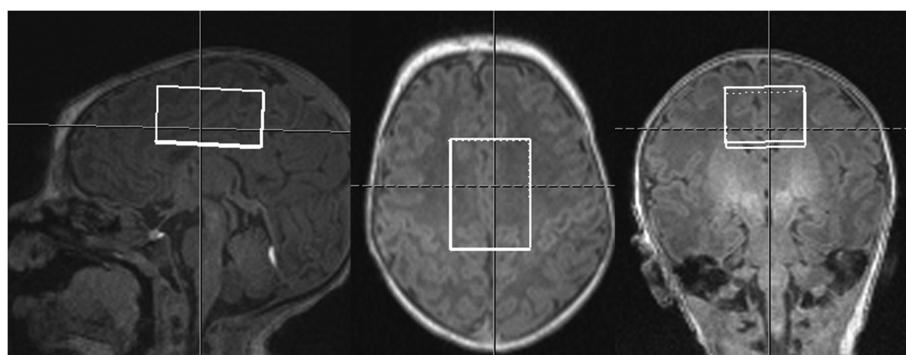


FIGURE 1 Typical placement of MEGA-PRESS voxel superior to the corpus callosum

3.3 | Spectral analysis

Acquired data were postprocessed using code adapted from the FID-A processing toolbox (github.com/CIC-methods/FID-A).²⁷ The code was applied to the free induction decay data that were manually exported from the scanner after each scanning session. These data contained 2048 complex time samples from each of the receive coil elements of the selected MEGA-PRESS voxel. We used the water signal from the nonwater-suppressed acquisition to phase the data and to determine the coil combination weights. We then performed a weighted coil array recombination, using the same weights for both the water-suppressed and the nonwater-suppressed data. We removed averages from the water-suppressed data contaminated by motion; these were identified by calculating an “unlikeness” metric for each average and flagging those that deviated by more than a threshold number of standard deviations away from the mean value.²⁸ This threshold was usually set at 3.2 (194 of 202 scan sessions); however, in eight sessions it was manually adjusted in the range of 2.2–4.5, based on visual inspection of the data, to ensure that spectra of poor quality were removed. We then performed frequency-drift corrections for both the “edit-on” and “edit-off” conditions.²⁹ In some infants, head movement during the scan produced abrupt changes in frequency. We discarded “edit-on”/“edit-off” pairs of averages that were acquired after the frequency from the “edit-off” acquisition had drifted by more than 2 Hz from the start of each scan. Basic statistics on the numbers of averages scanned, discarded and retained are reported in Table 1.

We summed all nondiscarded pairs of averages, and then performed a manual adjustment of frequency and phase to minimize residual Cho and tCr peaks in the difference spectrum. Difference spectra were examined visually, and two were eliminated due to poor spectral quality.

LCModel (version 6.3-0 L) was used to perform an eddy current correction and to provide quantitation of metabolite signal. Spectral quality was acceptable, as shown by the representative spectra in Figure 2. Analyzed metabolite signals include GABA, combined N-acetylaspartic acid and N-acetylaspartyglutamic acid (tNAA), NAA, tCr, combined glycerophosphocholine and phosphocholine (Cho), Glx, glutamate (Glu), Ins, glutathione (GSH) and taurine (Tau). These 10 designations pertain to individual metabolites, or metabolite groups; for simplicity we refer to all of them as metabolite groups, even if the group consists of only one metabolite. The GABA signal was quantified from the difference spectra, and all other metabolites were quantified from the “edit-off” spectra. The LCModel analysis was run twice for each dataset, with and without water scaling. Information about the LCModel basis sets used can be found in the supporting information.

We identified spectra of lower quality using absolute Cramér-Rao Lower Bounds (CRLBs), calculated from the water-scaled data.³⁰ For each metabolite group, we identified high outliers as points that were greater than the median by more than 1.5 times the interquartile range. Six spectra, for which more than four of the 10 metabolite groups were identified as high outliers, were deemed to be of low quality and discarded. Nonwater-scaled spectra from the same scan sessions were also discarded. A summary of the discarded spectra is shown in Table 2. Values of signal-to-noise ratios (SNRs) and full width half max (FWHM) calculated from LCModel analysis on nondiscarded spectra were found to be within an acceptable range (Table 1).

For GABA, values of the water-scaled signal, A, tend to increase with FWHM (Figure S1). We statistically corrected for this effect by performing a linear regression of A as a function of FWHM. Expected values calculated from the fit were subtracted from the measured values. The resulting deviation values were added to mean values of A to generate revised estimates of A.

3.4 | Quantification with the combined reference method, CRef

Values of $R_{C\text{Ref}}$ were calculated in two ways: using raw metabolite signals, S, as shown in Equation 2 and by replacing the raw signals with water-scaled metabolite signals, A. We denote these two measures as $R_{C\text{Ref},S}$ and $R_{C\text{Ref},A}$, respectively. Values of $R_{C\text{Ref},A}$ were calculated using seven reference metabolites: GABA, tNAA, Cho, tCr, Ins, Glx and GSH. In this case, the use of water scaling allowed GABA to share the same scale as the

TABLE 1 (a)–(d): Descriptive statistics (Min, Max, Mean and Median) of the numbers of “edit-on”/“edit-off” pairs of averages that were (a) acquired, (b) discarded based on the “unlikeness” metric, (c) discarded based on drift, and (d) used for spectral quantification in LCModel. Statistics in columns (a)–(d) are calculated from 202 scan sessions. (e)–(h): Spectral quality statistics from LCModel: signal-to-noise ratios (SNRs) from (e) difference spectra and (f) “edit-off” spectra; and full width half maximum (FWHM) values from (g) difference spectra and (h) “edit-off” spectra. Statistics in columns (e)–(h) are calculated from the nondiscarded spectra: 181 difference spectra and 183 “edit-off” spectra

					Spectral quality statistics from LCModel			
Numbers of pairs of averages					SNR		FWHM (ppm)	
(a) Acquired	(b) Discarded due to “unlikeness” metric	(c) Discarded due to drift	(d) Used for spectral quantification		(e) difference	(f) “edit-off”	(g) difference	(h) “edit-off”
Min	48	0	0	21	14	12	0.016	0.023
Max	144	38	104	144	32	34	0.101	0.078
Mean	134.7	8.5	8	118.1	23.9	23.1	0.037	0.036
Median	144	5	0	132.5	24	23	0.031	0.031

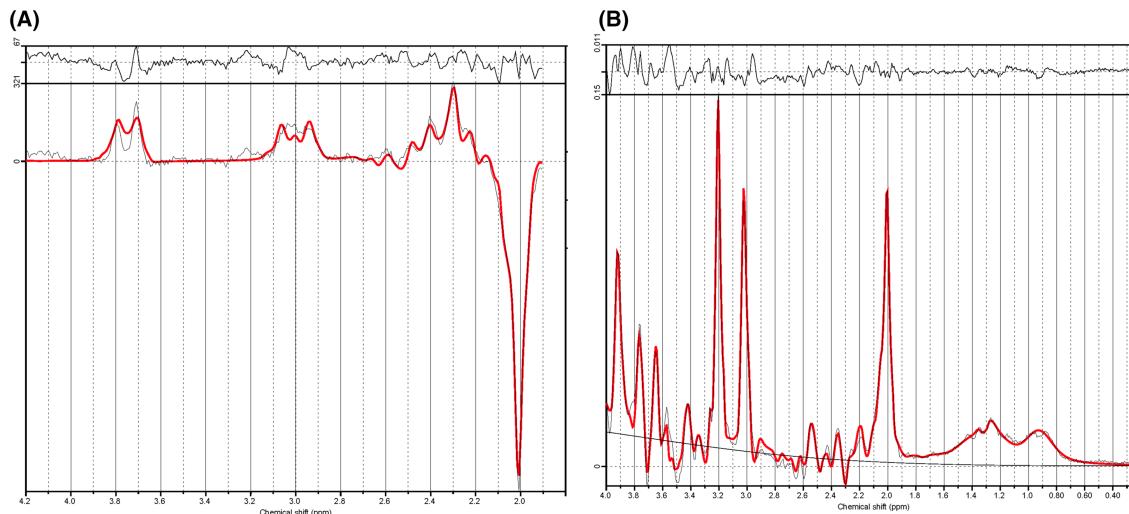


FIGURE 2 Representative (A) difference and (B) “edit-off” spectra for an infant aged 22 days

TABLE 2 Summary of discarded MRS scan sessions

Total scan sessions:	202
Discarded scan sessions	
Missing nonwater-suppressed acquisition	13
Visual inspection; poor difference spectra (GABA only)	2
High Cramér-Rao Lower Bounds across 4 of 10 metabolites	6
Remaining scan sessions	181 for GABA
	183 for other metabolites

metabolites in the “edit-on” spectrum in LCModel. However, when water scaling is not used, the difference spectrum and the “edit-on” spectrum do not share the same scale in LCModel because they are calculated from different basis sets. For this reason, values of R_{CRef_S} were not calculated for GABA, and metabolites were referenced to the six “edit-off” reference metabolites. For direct comparisons of R_{CRef_S} with R_{CRef_A} , the latter was calculated using the same six metabolites as used for R_{CRef_S} .

Standard deviation values for the normalizations were calculated after removing variance due to PMA dependence. The PMA correction of S_{Ref} and A_{Ref} signals was performed by first pooling data from M1 and M3 and then calculating expectation values, using second-order polynomials.

$$\begin{aligned} S_{Ref_E} &= S_0 + S_1 \hat{x} + S_2 \hat{x}^2 \\ A_{Ref_E} &= A_0 + A_1 \hat{x} + A_2 \hat{x}^2 \end{aligned} \quad (7)$$

where the variables with numerical subscripts denote fitting constants, and \hat{x} is the normalized PMA, such that $\hat{x} = (x - \bar{x})/SD_x$, where $\bar{x} = 338.4$ days is the mean and $SD_x = 44.9$ days is the standard deviation of PMA values. Second-order polynomials were chosen to account for the nonlinear dependence of some metabolites on age. The standard deviations, S_{Ref_SD} and A_{Ref_SD} , for each metabolite were calculated from the deviation values ($S_{Ref} - S_{Ref_E}$) and ($A_{Ref} - A_{Ref_E}$).

To characterize the PMA dependence of R_{CRef_S} , R_{CRef_A} , C_{WSc} and C_{CRef} , we calculated expectation values $R_{CRef_S_E}$, $R_{CRef_A_E}$, C_{WSc_E} and C_{CRef_E} as functions of PMA using second-order polynomials, like those shown in Equation (7).

3.5 | Quantification with corrected water scaling

We validate CRef analysis by comparing our results with molal concentration values measured using the corrected WSc.¹³ In this method, the molal concentration, C_{WSc} , in units of moles of metabolite per kg of ¹H-MRS-visible water, is proportional to the ratio of metabolite signal, S , to the corrected water signal, S_{Wc} ,

$$C_{WSc} = \frac{S}{S_{Wc}} [W], \quad (8)$$

where $[W]$ is the molality of neat water, 55.6 mol/kg. As in the Theory section of the paper, we neglect metabolite relaxation and take the signal to be the result of an integration of the fit to the metabolite spectrum, which has been corrected for the number of protons per molecule.

The corrected water signal, S_{Wc} , is related to the observed signal, S_W , according to

$$S_{Wc} = \frac{S_W f_{Br}}{R_W}, \quad (9)$$

where f_{Br} is the fraction of water signal that originates from the metabolite-containing tissue, and the term R_W corrects for relaxation of the water signal. Combining Equation 8 and Equation 9 we obtain

$$C_{WSc} = \frac{R_W}{f_{Br}} A [W], \quad (10)$$

where A is the water-scaled metabolite signal. All of the metabolite signal is assumed to originate from brain tissue, which includes both gray and white matter.¹³ In this case, f_{Br} will be the mole fraction of brain tissue water, including both gray and white matter, relative to total water signal, which includes contributions from gray matter, white matter and CSF. R_W accounts for the relaxation of the total water signal. Because our reference water signal was obtained at a high TR of 30 s, it does not require a correction for T_1 relaxation. The relaxation factor R_W can be estimated with a variety of methods, including methods that distinguish relaxation in gray matter, white matter and CSF.¹³ One standard approach estimates R_W from measurements of the water signal attenuation as a function of TE.¹⁵ By attributing short-TE attenuation to brain tissue (combined gray and white matter), and long-TE relaxation to CSF, this method also provides estimates of f_{Br} . Details of how these corrections were calculated can be found in the supporting information.

Our validation of CRef analysis is performed by correlating the combined reference signal with the water signal, S_W . A signal proportional to S_W was calculated from the ratios S/A , where the scale of S is set by the basis set of LCModel. In principle, this estimate could be obtained from any metabolite; however, to minimize bias we calculated S/A from the six reference metabolites of the “edit-off” spectrum. The within-scan session variance of S/A across metabolites was less than 1% for all subjects, indicating that this ratio provides an estimate of S_W that depends little on the metabolite used. Although S/A is only proportional to S_W , it is still valid for characterizing the statistical consistency of S_W with CRef measures using Pearson correlations, which are independent of absolute scale. When used for this purpose, S_W for each subject is taken to be the average of S/A across metabolite groups.

We calculated corrected water-scaled molal concentrations, C_{WSc} , using Equation 10, with values of f_{Br} and R_W calculated from data obtained at various values of TE, as described in the supporting information. Of the 202 total scan sessions, 170 datasets had a complete set of single-voxel acquisitions at all nine TE values. However, many of the scan sessions were contaminated by motion, as shown in Figure S3. Using the criteria described in the supporting information, we discarded 94 scan sessions, leaving valid T_2 data from 76 scan sessions. Five of these were from scan

sessions from which the metabolite data had been discarded, leaving 71 scan sessions for comparisons between R_{CRef} and molality. Organization and analysis of data was performed using custom scripts in MATLAB (version R2019b; MathWorks, Natick, MA, USA).

4 | RESULTS

We evaluate R_{CRef} by first showing that similar values are obtained whether raw signal intensities or water-scaled signals are used. We then demonstrate the importance of normalizing reference metabolite signals by the standard deviations. This is done by showing that in the absence of this correction the reference signal is biased toward metabolites of greatest variance. Finally, we compare R_{CRef} with the “gold standard” of ^1H -MRS metabolite quantification: molality values obtained by correcting water-scaled spectra for partial volume effects.

4.1 | Calculations of R_{CRef} from raw signal versus water-scaled values

Measurements of raw signal, S , and water-scaled signal, A , exhibit PMA dependence, as shown in Figure S4. Values of w_S calculated from raw metabolite signals range from 0.09 for tNAA to 0.32 for Cho, as shown by column (a) of Table 3. Values of w_A are similar to w_S , differing by 0.01 to 0.03, as shown by a comparison of columns (a) and (b) in Table 3. Because of the similarity of w_S and w_A , Pearson correlation coefficients between R_{CRef_S} and R_{CRef_A} are greater than 0.994 for all metabolites, as shown in column (c) of Table 3, indicating high statistical consistency of the two measures. Moreover, values of R_{CRef_S} and R_{CRef_A} are also numerically similar; the average across subjects of the absolute % difference is 1.6% or less, as shown in column (d) of Table 3. The consistency between R_{CRef_S} and R_{CRef_A} is reduced when standard deviation values are calculated without first removing PMA dependence (see the supporting information for further discussion, including Tables S1 and S2, and Figure S5). The inclusion of GABA as a reference metabolite for the water-scaled data has little effect on R_{CRef_A} values, as shown in Table S3.

A key feature of CRef analysis is the normalization of the reference metabolite signals by their standard deviations. We demonstrate the importance of this normalization by showing that in its absence the variance of the reference signal is more sensitive to metabolites of high variance, such as Ins and tNAA, and less sensitive to those of low variance, such as Cho and GABA. These differences are large, ranging up to factor of 6.6, as shown in Table S4.

Values of R_{CRef_A} from different metabolites exhibit distinct patterns of change with PMA, as shown in Figure 3A. The dashed lines in these plots show average PMA dependence calculated from the fitting coefficients shown in Table S5.

4.2 | Comparison of CRef with corrected water-scaled concentrations

Our validation of combined referencing is based on a comparison with C_{WSc} . Values of C_{WSc} exhibit PMA dependence that is qualitatively similar to R_{CRef_A} , as shown in Figure 3B and Table S6.

Measurements of molality require corrections for brain water fraction, f_{Br} , and total water relaxation, R_W , where $C_{WSc} = (R_W/f_{Br})A[W]$ (see Equation 10). These are calculated from supplemental relaxation experiments, as described in the supporting information. Values of T_2 in brain water tend to decrease with PMA, as shown in Figure 4A, whereas values of T_2 in CSF exhibit little PMA dependence (Figure 4B). Because the water reference signal was acquired at a high TR, the correction factor, R_W , of the total water signal exhibits no T_1 -related attenuation and can

TABLE 3 Weighting factors (a) w_S and (b) w_A , calculated using six reference metabolites. (c) Pearson correlation coefficients of R_{CRef_S} versus R_{CRef_A} . (d) Mean estimation differences $|R_{CRef_A} - R_{CRef_S}| / R_{CRef_S}$ (%)

	(a) w_S	(b) w_A	(c) Pearson R: R_{CRef_S} vs. R_{CRef_A}	(d) Mean of $ R_{CRef_A} - R_{CRef_S} / R_{CRef_S}$ (%)
tNAA	0.09	0.10	0.9991	1.3
NAA	-	-	0.9994	1.2
tCr	0.12	0.14	0.9948	1.6
Cho	0.32	0.33	0.9983	0.9
Ins	0.07	0.06	0.9996	1.1
Glx	0.11	0.08	0.9993	1.1
Glu	-	-	0.9997	1.1
GSH	0.29	0.28	0.9964	0.7
Tau	-	-	0.9994	0.4

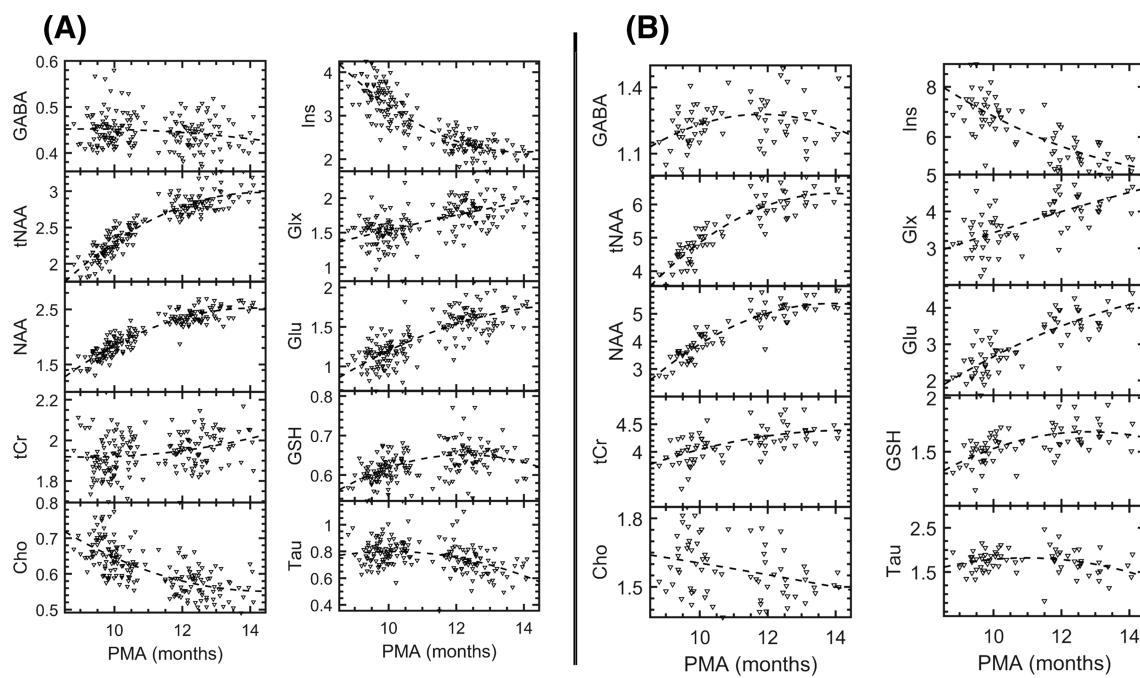


FIGURE 3 (A) Combined ratios, $R_{C\text{Ref}}$, as functions of postmenstrual age (PMA) (months). Average PMA dependence is calculated from polynomial fits to the data, shown with the dashed lines. The fitting parameters are shown in Table S5. (B) Metabolite molal concentrations, $C_{W\text{Sc}}$, as functions of PMA (months). Average PMA dependence is calculated from polynomial fits to the data, shown with the dashed lines. The fitting parameters for $C_{W\text{Sc}}$ are shown in Table S6. Molal concentration units are mmol/L

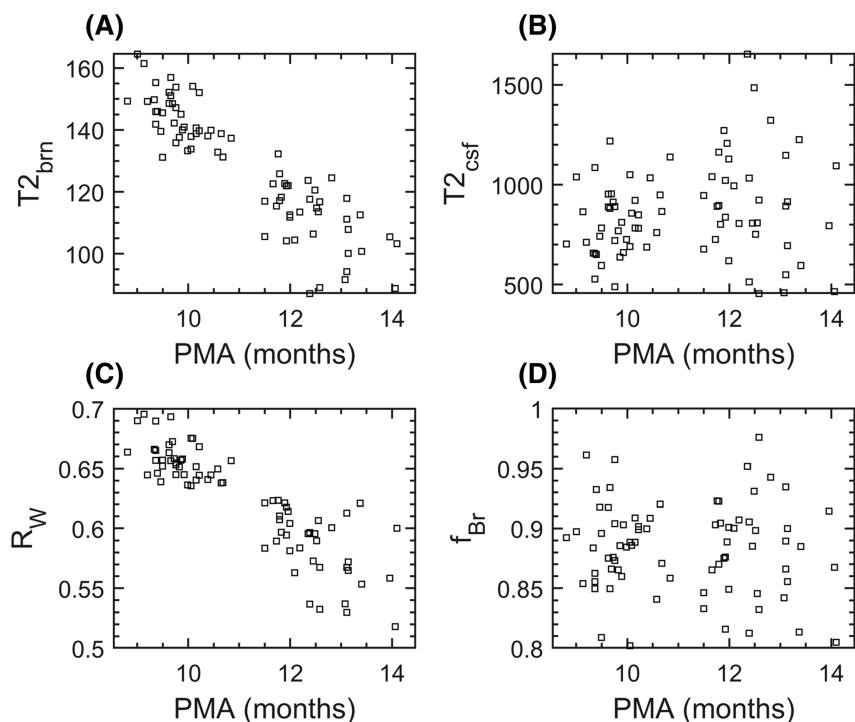


FIGURE 4 Postmenstrual age (PMA) dependence of (A) T_2 relaxation in brain tissue, including gray and white matter, (B) T_2 relaxation of CSF, (C) correction for T_2 relaxation of total water signal, and (D) the fraction of brain water relative to total water. Data are from 76 measurements from 68 subjects

therefore be calculated from the T_2 relaxation times. The factor R_W tends to decrease with age, as shown in Figure 4C, reflecting the decreasing T_2 relaxation time of brain water with age. Values of f_{Br} exhibit little PMA dependence, as shown in Figure 4D.

We investigate the statistical consistency of R_{CRef} with C_{Wc} by comparing their reference signals. For raw metabolite signals, the reference of C_{Wc} is $S_{Wc} = (S_W f_{Br}) / R_W$ and the reference of R_{CRef} is $\sum(wS_{Ref})$. For water-scaled signals, the reference of C_{Wc} is f_{Br} / R_W and the reference of R_{CRef} is $\sum(wA_{Ref})$. By comparing both pairs of references, we separately assess the usefulness of the combined reference for replacing both the corrected water signal, S_{Wc} , and the f_{Br} / R_W portion of this reference.

The factor $\sum wS_{Ref}$ from the six reference metabolites used to calculate $R_{CRef,S}$ is highly correlated with S_{Wc} , with a Pearson correlation coefficient, $R = 0.96$ (see column (a) of Table 4 and Figure 5A). The shared variance vanishes when the ratio $\sum wS_{Ref} / S_{Wc}$ is taken. Multiplication of this factor by the molality of neat water, $[W]$, yields a measure with units of molality, which we call the reference concentration. It is a weighted average of the molality of the reference metabolites, neglecting metabolite relaxation. The uniformity of the reference concentration is characterized by the coefficient of variation, CV_{RC} , of the reference concentration, which is 0.062, as shown in column (b) of Table 4. Values of R and CV_{RC} change little when one of the reference metabolites is excluded for the sake of avoiding self-referencing, as shown by columns (a) and (b) of Table 4.

Our data can also be used to assess the uniformity of a single reference. In this case $w = 1$, and the reference concentration is simply the concentration of the reference metabolite group, $[W]S_{Ref} / S_{Wc}$. Correlation coefficients, R , of S_{Ref} versus S_{Wc} , are generally lower and more widely distributed than the R values for the combined references (see column (c) of Table 4). Similarly, CV_{RC} values from single references are higher and more widely distributed than those from combined references (see column (d) of Table 4). When PMA-dependent variance is removed from $\sum wS_{Ref}$ and S_{Wc} , values of R and CV_{RC} are similar (Figure 5B), and the combined references continue to exhibit greater uniformity, as shown in Table S7.

To evaluate the ability of the combined reference to compensate for partial volume effects and relaxation of the water signal, we repeated the analysis described above, but compared $\sum(wA_{Ref})$ with f_{Br} / R_W . These results show that the Pearson coefficient R values are in the range of 0.84 to 0.87 without controlling for PMA (Figure 5C), and 0.65 to 0.71 when controlling for PMA (Figure 5D), as shown in Table S8. As with the raw signal correlations, values of R from single reference metabolites were generally lower than those from combined references.

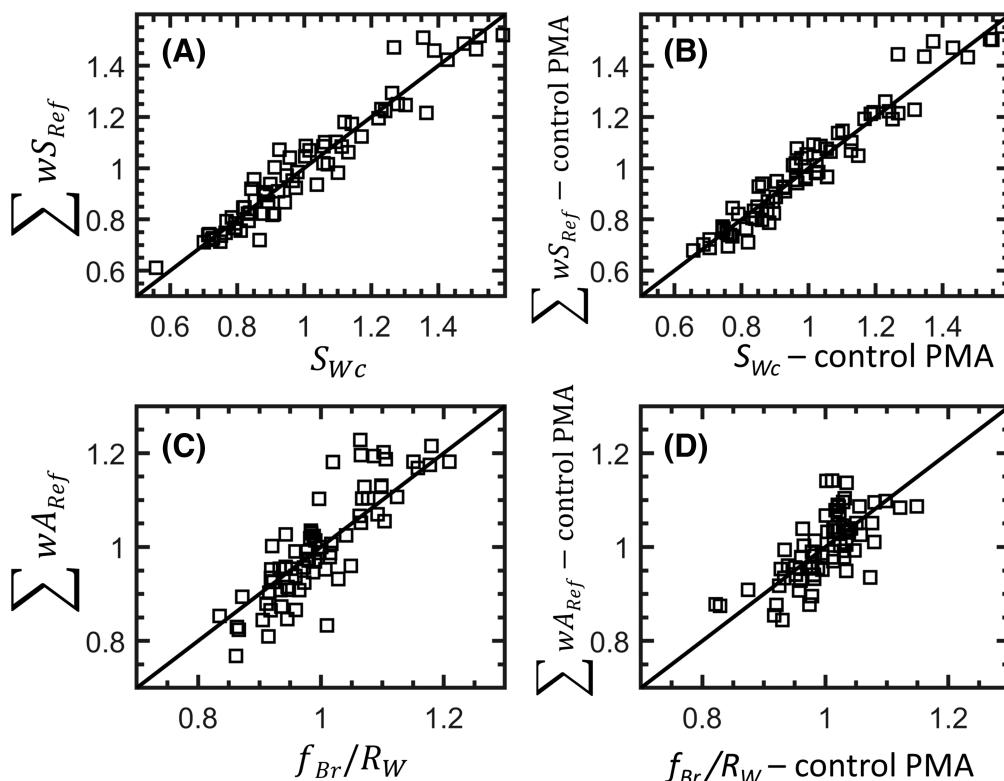


FIGURE 5 Comparison of the combined reference from six metabolites, $\sum wS_{Ref}$, with the corrected water signal, S_{Wc} , (A) without controlling for postmenstrual age (PMA) and (B), where PMA-dependent variance has been removed. To examine the ability of the combined reference to account for correction factors to the water signal we also correlate the combined reference of the water-scaled signal from seven metabolites, $\sum wA_{Ref}$, with the water corrections, $\frac{f_{Br}}{R_W}$, (C) without controlling for PMA, and (D) where PMA-dependent variance has been removed. Each measure is normalized by its mean value and the solid lines indicate unity. Correlations are calculated from 71 measurements on 64 subjects

The statistical consistency of the combined reference signal with the corrected water signal suggests the possibility of rescaling C_{CRef} to units of molal concentration using PMA-dependent expectation values of C_{WSc} , as shown in Equation 6. The resulting concentration estimates, C_{CRef} , exhibit an average PMA dependence that closely matches C_{WSc} , as shown in Table S9 and Figure S6. We evaluate the numerical accuracy of C_{CRef} in predicting C_{WSc} at the individual level by calculating Pearson correlation coefficients, R_C , of C_{CRef} versus C_{WSc} . These vary from 0.65 for tCr to 0.97 for Glx, as shown in column (a) of Table 5. To remove PMA dependence, we correlate deviation values C_{CRef_D} and C_{WSc_D} , calculated by subtracting PMA-dependent expectation values from C_{CRef} and C_{WSc} . The Pearson correlation coefficients R_D of C_{CRef_D} versus C_{WSc_D} are systematically lower than R_C , ranging from 0.44 for tCr to 0.93 for Tau, as shown in column (b) of Table 5. The value of R_D for a given metabolite is related to the PMA-independent variability in the molality of that metabolite, which is characterized by the coefficient of variation, $CV_{C_WSc_D}$, which is the standard deviation of C_{WSc_D} divided by the mean values of C_{WSc} . Metabolites of high R_D tend to exhibit higher values of $CV_{C_WSc_D}$; examples include Ins, Glx and Tau, as shown in Figure 6A. These metabolites also tend to exhibit higher values of average % CRLBs, as shown in Figure 6B. For these metabolites, C_{CRef_D} more accurately predicts C_{WSc_D} , which is distributed over a relatively wide range, $CV_{C_WSc_D}$, likely due to the higher measurement uncertainty indicated by higher average % CRLBs. For metabolites of low $CV_{C_WSc_D}$ and low average % CRLBs, such as tCr, C_{CRef_D} provides a less accurate prediction of C_{WSc_D} , or lower values of R_D . For these metabolites, especially tCr, variation within a relatively small range is not accurately estimated using a combined reference of metabolites of lesser uniformity.

The error associated with using C_{CRef_D} to approximate C_{WSc_D} is given by the standard deviation, σ_{CRef_WSc} , of $C_{CRef_D} - C_{WSc_D}$. Ratios of σ_{CRef_WSc} to the standard deviation, $\sigma_{C_WSc_D}$, of C_{WSc_D} values indicate the extent to which the combined reference can decrease the uncertainty of the molal concentration beyond what would be achieved with only the use of population averages, with no individual measurement being performed. Ratios $\sigma_{CRef_WSc}/\sigma_{C_WSc_D}$ are in the range of 0.36 for Tau to 1.04 for tCr, as shown in column (c) of Table 5.

TABLE 4 Comparison of reference metabolite signals, $\sum wS_{Ref}$, with water signals, S_{Wc} , that have been corrected for partial volume and relaxation effects. In columns (a) and (b), each row pertains to the combined reference for a metabolite group of interest (MOI), where the MOI is excluded from serving as its own reference. In columns (c) and (d), each row refers to a single reference metabolite. Columns (a) and (c) report the Pearson correlation coefficient of the metabolite signal(s) with S_{Wc} . Columns (b) and (d) report the coefficient of variation (CV), or standard deviation divided by the mean, of the reference concentrations. The data are drawn from 71 scan sessions from 64 infants.

Metabolite	Metabolite as MOI		Single metabolite or metabolite group	
	(a) R of $\sum wS_{Ref}$ vs. S_{Wc}	(b) CV_{RC} of $\frac{\sum wS_{Ref}}{S_W}$	(c) R of S_{Ref} vs. S_{Wc}	(d) CV_{RC} of $\frac{S_{Ref}}{S_{Wc}}$
All reference metabolites	0.96	0.062	-	-
tNAA	0.97	0.056	0.78	0.159
tCr	0.96	0.064	0.96	0.067
Cho	0.95	0.07	0.95	0.076
Ins	0.94	0.077	0.84	0.174
Glx	0.97	0.058	0.79	0.17
GSH	0.97	0.058	0.91	0.102

TABLE 5 Comparison of C_{CRef} with C_{WSc} using seven reference metabolites (GABA, tNAA, tCr, Cho, Ins, Glx and GSH). Here, R_C is the Pearson correlation coefficient of C_{CRef} versus C_{WSc} , R_D is the Pearson correlation coefficient of deviation values C_{CRef_D} versus C_{WSc_D} , σ_{CRef_WSc} is the standard deviation of $C_{CRef_D} - C_{WSc_D}$, and $\sigma_{C_WSc_D}$ is the standard deviation of C_{WSc_D} values. Ratios $\sigma_{CRef_WSc}/\sigma_{C_WSc_D}$ in column (c) indicate the extent to which the combined reference can narrow the uncertainty of the molal concentration beyond what would be achieved with only the use of population statistics. All statistics are performed on 71 measurements from 64 infants

	(a) R_C of C_{CRef} vs. C_{WSc}	(b) R_D of C_{CRef_D} vs. C_{WSc_D}	(c) $\frac{\sigma_{CRef_WSc}}{\sigma_{C_WSc_D}}$
GABA	0.75	0.72	0.74
tNAA	0.94	0.61	0.81
NAA	0.96	0.73	0.69
tCr	0.65	0.44	1.04
Cho	0.72	0.68	0.78
Ins	0.96	0.88	0.46
Glx	0.95	0.9	0.42
Glu	0.97	0.9	0.41
GSH	0.88	0.81	0.6
Tau	0.94	0.93	0.36

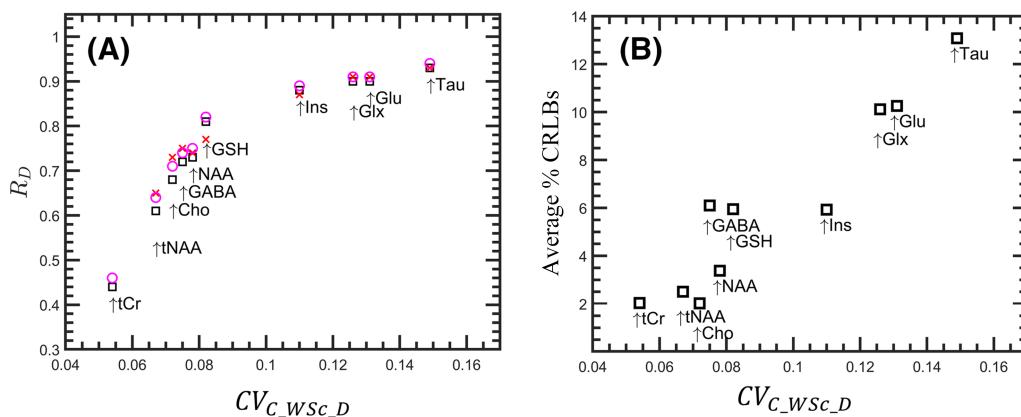


FIGURE 6 (A) The Pearson correlation coefficient, R_D , from correlations of deviation values C_{WSc_D} versus C_{CRef_D} is an increasing function of the postmenstrual age (PMA)-controlled coefficient of variance, $CV_{C_WSc_D}$, of molal concentration. Values of R_D are calculated from C_{CRef_D} values referenced by seven metabolites (black squares), creatine only (red x's), and five metabolites (magenta circles). (B) Metabolites of greater $CV_{C_WSc_D}$ tend to exhibit higher average values of % Cramér-Rao Lower Bounds (CRLBs)

Given that tCr exhibits the lowest values of $CV_{C_WSc_D}$, and the lowest average % CRLBs (Figure 6), it is likely to be the most accurate single reference for estimating C_{CRef} . Indeed, values of R_C and R_D are slightly higher, and values of $\sigma_{CRef_Wsc}/\sigma_{C_WSc_D}$ are slightly lower, for a tCr reference than for the use of seven references, as shown in Figure 6A and Table S10. However, these seven reference metabolites include two metabolites, Glx and Ins, of high variability. When these are removed, the accuracy of the resulting five-reference C_{CRef} values are similar to those calculated using tCr as a single reference, as shown in Figure 6A and Table S10.

5 | DISCUSSION

We have demonstrated the use of CRef analysis to generalize the use of metabolite ratios to multiple reference metabolites. An important feature of CRef analysis is the normalization of reference metabolite signals by their relative inverse standard deviations. The purpose of the normalization is to avoid biasing the combined reference towards metabolites with the greatest variance. Without this correction, the reference metabolite signal can be more than six times more sensitive to metabolites of high variance than to those with low variance, thereby complicating the interpretation of results.

We have shown that the weighting coefficients, w , derived from standard deviation values show little dependence on whether water scaling is used, provided that PMA-dependent variance is removed before calculation of standard deviation values. The numerical consistency of the resulting R_{CRef} values is due to normalization of S_{Ref_SD} values by their harmonic mean.

To the extent that the combined reference signal is not biased towards a reference metabolite, it is possible that R_{CRef} may be used as a surrogate for molal concentration. We have demonstrated that the combined metabolite reference signal, $\sum(wS_{Ref})$, is highly correlated with the water signal, S_{Wc} , corrected for partial volume and relaxation (Pearson $R = 0.96$, when no PMA dependence is removed). This reflects good uniformity of the reference concentration, $[W] \sum(wS_{Ref})/S_{Wc}$, which is a weighted average of the molal concentrations of the reference metabolites. Generally, the reference concentrations from multiple metabolites groups are more uniform than those of single metabolites groups.

The use of R_{CRef} as a surrogate for molal concentration has the advantage of circumventing corrections of the water signal due to the partial volume and relaxation. Because variance from these effects is present in the water-scaled values, A , the reference signal $\sum(wA_{Ref})$ is strongly correlated with f_{Br}/R_W (Pearson $R = 0.86$, when no PMA dependence is removed). Moreover, a greater correlation strength might have been achieved without the mismatch of offset frequency values that caused these signals to be obtained from slightly different locations. Additional unshared variance among the measurements may be due to differences in the proton density among infants, variance in the sum of the metabolite concentrations, and measurement errors. Nevertheless, the strong correlation of the reference signal $\sum(wA_{Ref})$ with f_{Br}/R_W implies that $\sum(wA_{Ref})$ can provide a surrogate for the factor f_{Br}/R_W . The potential to circumvent the supplemental measurement required to obtain f_{Br}/R_W is important to our dataset, given the tendency of these measurements to be compromised by head motion.

The consistency of the combined metabolite reference with the corrected water signal suggests the possibility of rescaling R_{CRef} to units of molal concentration, C_{CRef} , using expected values of the reference metabolites. We have shown that the numerical consistency of C_{CRef} with C_{Wsc} is greatest for metabolites of the highest variance. This result suggests that when estimating metabolites of high variance or measurement uncertainty, the distinction between C_{CRef} and C_{Wsc} is diminished. In this case the choice of reference becomes less critical, and the limiting factor for accurate estimates of concentration is likely to be the uncertainty associated with spectral analysis.

Total creatine, tCr, as a single reference is generally well suited for accurate estimates of molal concentration based on metabolite referencing. This arises from the fact that in our study tCr is the single metabolite group with the most uniform reference concentration, exhibiting a uniformity that is similar to, or greater than, that of the combined references. The implication is that when single references are used, tCr is likely to be the preferred reference, provided that its concentration is not dependent on study effects. Further work is needed to assess whether this result applies to other protocols and populations; for example, in human adults, the use of tCr as a reference is complicated by large concentration differences between gray and white matter.³¹

A challenge of CRef analysis is choosing which metabolites to use as references. This choice may depend on the purpose of the study and the population being tested. For accurate estimates of concentration in typically developing infants, $C_{C\text{Ref}}$ can be calculated using a small number of the more uniform metabolites, or even tCr as a single reference. However, for many studies the highest priority is to identify which metabolites, including possibly tCr, are dependent on study conditions. For these studies, $R_{C\text{Ref}}$ values can be calculated using a relatively inclusive set of reference metabolites, thereby diluting study-dependent effects from any of the reference metabolites. Because of this dilution effect, the combined reference may provide less ambiguity of interpretation than the use of a single reference.

In a combined reference ratio, the asymmetry between a single reference in the numerator and multiple references in the denominator is a key feature for isolating study-dependent effects among metabolites. When a metabolite with study-dependent effects is used as a reference, such effects will be diluted by the presence of other reference metabolites. However, when the same metabolite is employed as the MOI its effect is not diluted by summation with other metabolites, causing study-dependent effects to be more salient. Employing each reference metabolite as the MOI allows the relative contributions of the metabolites to be characterized. This feature may be especially useful for detecting disease or damage states in which one or two metabolites are unusually high or low, as can be seen in neonatal hypoxic ischemic brain injury and metabolic disorders such as Canavan Disease or Transporter Deficiency.²

When study-dependent effects are present in multiple metabolites, CRef analysis offers a framework for reducing relationships among metabolites to a lower dimensionality than would be possible with the use of multiple simple ratios. Future work will demonstrate the use of CRef analysis for aiding the efficiency of descriptive statistics among multiple metabolites.

The use of $R_{C\text{Ref}}$ is most straightforward when applied within a study. However, comparisons of $R_{C\text{Ref}}$ or $C_{C\text{Ref}}$ across studies present several challenges, including variation in the tissues sampled, and the sequences and protocol used. Ratios of metabolites such as $R_{C\text{Ref}}$ are likely to exhibit partial volume effects that arise from metabolite concentration differences between gray and white matter.¹⁸ Partial volume effects and sequence parameters may also affect the relative variances, thereby affecting optimal w factors. Values of w factors reported here pertain to our sequence and associated parameters, voxel placement, and processing steps. The use of w factors reported here for unmatched protocols may not fully correct sensitivity bias among metabolites. To aid with interpretation of data, it is important that values of $R_{C\text{Ref}}$ that are compared across scans or subjects are calculated using the same w factors.

To the extent that protocols and voxel placement are consistent among sites and studies, $R_{C\text{Ref}}$ and $C_{C\text{Ref}}$ may be useful for comparisons of single subject data with normative data. To aid with these comparisons, we have developed a freely available standalone application that can be downloaded and installed on Windows or Macintosh operating systems.³² This program provides a graphical user interface where users can provide metabolite signal values, and specify any combination of reference metabolite groups. The program calculates the w factors from our normative data using the same reference metabolites and plots the user-provided data along with data from our database and average growth curves. The program also reports percentiles of $R_{C\text{Ref}}$ and $C_{C\text{Ref}}$ values relative to PMA-matched infants in our dataset. To aid with the exploration of our database, the program also allows the plotting of any scan from our database in place of the user-provided data. For the subset of infants with available data, values of $C_{W\text{Sc}}$ can be plotted for comparison with $C_{C\text{Ref}}$ at the individual level. We also provide the MATLAB source code for the standalone application and a MATLAB script that performs many of the same functions.

For the estimates from user-provided data to be accurate, it is important to use our scanning sequence and acquisition parameters, voxel placement, and spectral processing and analysis. However, the effect of alternative spectral processing and analyses pipelines on the values of $R_{C\text{Ref}}$ and $C_{C\text{Ref}}$ can be investigated using raw and processed spectra that we are making available.

To the best of our knowledge, these data are the largest sample of normative subjects in this age range, particularly for GABA. We therefore provide a summary of the % changes of molal concentration, $C_{C\text{Ref}}$, of all metabolites during the first 3 months of life (Table 6). Our measurements of GABA concentration obtained using the WSc method show a trend of increasing GABA concentration with PMA (Pearson correlation coefficient $R = 0.21$, $p = 0.07$, as shown in Table 6). The increase in GABA we observe is consistent with observations of increases in GABA concentration in rat brains beginning at postnatal Day 7,³³ which approximately corresponds to the neural developmental stage of a term infant at birth in humans.³⁴ Moreover, the GABA is lower in the temporal lobe of asymptomatic preterm infants compared with PMA-matched term infants,⁷ as well as in term infants with congenital heart disease,³⁵ suggesting that GABA concentration may serve as a marker of maturity or subclinical injury. This may be consistent with pathology data from humans showing that the development of the GABAergic system continues well into infancy.³⁶

Overall trends in PMA dependence of absolute metabolite concentrations are broadly consistent with previous reports.^{3,18,37,38} The substantial increases in NAA likely reflect neuronal maturation,³ which includes the development of dendrites and synapses,^{39,40} and oligodendrocyte proliferation related to myelination.⁴¹ The modest increase in tCr signal during this period may be due to the increasing energy demands

	(a) C_{WSc} vs. PMA R	(b) C_{WSc} vs. PMA p	(c) % change- C_{WSc}
GABA	0.21	0.07	8
tNAA	0.88	1.7×10^{-24}	44
NAA	0.89	3.4×10^{-25}	57
tCr	0.59	4.8×10^{-8}	10
Cho	-0.32	0.006	-5
Ins	-0.77	3.5×10^{-15}	-25
Glx	0.67	1.7×10^{-10}	27
Glu	0.83	5.8×10^{-19}	58
GSH	0.55	8.6×10^{-7}	17
Tau	-0.13	0.27	4

TABLE 6 Changes in absolute metabolite concentration with age calculated from WSc data. (a) The Pearson correlation coefficients, R, and (b) p-values from correlations between metabolite concentration and postmenstrual age (PMA). (c) The % change in metabolite concentration over a 90-day period beginning at a PMA of 280 days (40 weeks). These values are calculated using the average growth curves for WSc data, which are shown in Figure 4B

associated with myelination.⁴⁰ Choline-containing compounds include subcomponents of cell membranes such as phosphatidylcholine; the decrease in choline signal during the perinatal period may be associated with accelerated myelination.^{18,39,40} The decrease in Ins is consistent with other reports and may be precipitated by birth.¹⁸ To the best of our knowledge, the increase in GSH has not been observed in other studies on infants of a similar age; a previous study reported that GSH decreased over 32–43 weeks PMA.⁴² Tau is an osmolyte, and is known to be elevated in the developing brain and to decrease with maturation,^{3,43} which is consistent with our results. Elevated Tau levels during maturation relative to adulthood suggest that it is involved with brain development, possibly by promoting cell migration and organization.⁴⁴

Our results demonstrate the potential usefulness of quantification using a combined reference metabolite signal. CRef analysis does not require the water signal or its associated corrections. Moreover, by forming ratios with multiple metabolites, this approach transcends the pairwise comparisons of metabolites that are typically used, thereby reducing the number of required variables. This allows a more efficient determination of which metabolites are study dependent, and therefore most useful as potential references. Further work could explore the use of additional ways to simultaneously examine relationships among multiple metabolite ratios, including statistical methods.⁴⁵

Data Sharing

Study data, including the original and processed spectra from all scan sessions, are available at the Illinois Data Bank, https://doi.org/10.13012/B2IDB-3548139_V1. The dataset also includes executable files for installing the app described in the Discussion section on Windows and Macintosh operating systems, and a file for installing the app as a MATLAB app. A video demonstration of the app is available at https://mediaspace.illinois.edu/media/t/1_sl2kjnnk. The MATLAB source code of the app, and the MATLAB script that performs similar functions as the app, are available at <https://github.com/rylarsen/InfantCRefLibrary>. MATLAB scripts for generating all figures and tables in the paper from the individual spectra are available at <https://github.com/rylarsen/InfantCRefAnalysis>.

ACKNOWLEDGEMENTS

This work is funded by Abbott Nutrition through the Center for Nutrition, Learning, and Memory at the University of Illinois at Urbana-Champaign. We thank Marie Drottar, Thea Francel, Alana Matos, Clarissa Carruthers and Catherine Vu for assistance with study management and subject recruitment.

ORCID

Ryan J. Larsen  <https://orcid.org/0000-0003-0661-4321>

Bradley P. Sutton  <https://orcid.org/0000-0002-8443-0408>

REFERENCES

1. Rae CD. A guide to the metabolic pathways and function of metabolites observed in human brain H-1 magnetic resonance spectra. *Neurochem Res*. 2014;39(1):1-36.
2. Oz G, Alger JR, Barker PB, et al. Clinical proton MR spectroscopy in central nervous system disorders. *Radiology*. 2014;270(3):658-679.
3. Bluml S, Wisniewski JL, Nelson MD, et al. Metabolic maturation of the human brain from birth through adolescence: insights from in vivo magnetic resonance spectroscopy. *Cereb Cortex*. 2013;23(12):2944-2955.
4. Kimura H, Fujii Y, Itoh S, et al. Metabolic alterations in the neonate and infant brain during development - evaluation with proton MR spectroscopy. *Radiology*. 1995;194(2):483-489.

5. Mescher M, Merkle H, Kirsch J, Garwood M, Gruetter R. Simultaneous in vivo spectral editing and water suppression. *NMR Biomed.* 1998;11(6):266-272.
6. Ben-Ari Y. Excitatory actions of GABA during development: The nature of the nurture. *Nat Rev Neurosci.* 2002;3(9):728-739.
7. Kwon SH, Scheinost D, Lacadie C, et al. GABA, resting-state connectivity and the developing brain. *Neonatology.* 2014;106(2):149-155.
8. Helms G. The principles of quantification applied to in vivo proton MR spectroscopy. *Eur J Radiol.* 2008;67(2):218-229.
9. Kreis R. Quantitative localized H-1 MR spectroscopy for clinical use. *Prog Nucl Mag Res Sp.* 1997;31:155-195.
10. Christiansen P, Henriksen O, Stubgaard M, Gideon P, Larsson HBW. Invivo quantification of brain-metabolites by H-1-MRS using water as an internal standard. *Magn Reson Imaging.* 1993;11(1):107-118.
11. Knight-Scott J, Haley AP, Rossmiller SR, et al. Molality as a unit of measure for expressing H-1 MRS brain metabolite concentrations in vivo. *Magn Reson Imaging.* 2003;21(7):787-797.
12. Hetherington HP, Pan JW, Mason GF, et al. Quantitative H-1 spectroscopic imaging of human brain at 4.1 T using image segmentation. *Magn Reson Med.* 1996;36(1):21-29.
13. Gasparovic C, Song T, Devier D, et al. Use of tissue water as a concentration reference for proton spectroscopic imaging. *Magn Reson Med.* 2006;55(6):1219-1226.
14. Keevil SF, Barbiroli B, Brooks JCW, et al. Absolute metabolite quantification by in vivo NMR spectroscopy: II. A multicentre trial of protocols for in vivo localised proton studies of human brain. *Magn Reson Imaging.* 1998;16(9):1093-1106.
15. Ernst T, Kreis R, Ross BD. Absolute quantitation of water and metabolites in the human brain. 1. Compartments and water. *J Magn Reson Ser B.* 1993;102(1):1-8.
16. Knight-Scott J, Dunham SA, Shanbhag DD. Increasing the speed of relaxometry-based compartmental analysis experiments in STEAM spectroscopy. *J Magn Reson.* 2005;173(1):169-174.
17. Kreis R, Ernst T, Ross BD. Absolute quantitation of water and metabolites in the human brain. 2. Metabolite concentrations. *J Magn Reson Ser B.* 1993;102(1):9-19.
18. Kreis R, Ernst T, Ross BD. Development of the human brain - in-vivo quantification of metabolite and water-content with proton magnetic-resonance spectroscopy. *Magn Reson Med.* 1993;30(4):424-437.
19. Larsen RJ, Newman M, Nikolaidis A. Reduction of variance in measurements of average metabolite concentration in anatomically-defined brain regions. *J Magn Reson.* 2016;272:73-81.
20. Gussew A, Erdtel M, Hiepe P, Rzanny R, Reichenbach JR. Absolute quantitation of brain metabolites with respect to heterogeneous tissue compositions in H-1-MR spectroscopic volumes. *Magn Reson Mater Phys.* 2012;25(5):321-333.
21. Paul EJ, Larsen RJ, Nikolaidis A, et al. Dissociable brain biomarkers of fluid intelligence. *Neuroimage.* 2016;137:201-211.
22. Ferguson KJ, MacLullich AMJ, Marshall I, et al. Magnetic resonance spectroscopy and cognitive function in healthy elderly men. *Brain.* 2002;125:2743-2749.
23. Girard N, Gouny SC, Viola A, et al. Assessment of normal fetal brain maturation in utero by proton magnetic resonance spectroscopy. *Magn Reson Med.* 2006;56(4):768-775.
24. Jolliffe IT, Cadima J. Principal component analysis: a review and recent developments. *Philos Trans A Math Phys Eng Sci.* 2016;374(2065):1-16, 20150202.
25. Frahm J, Bruhn H, Gyngell ML, Merboldt KD, Hanicke W, Sauter R. Localized proton NMR spectroscopy in different regions of the human brain in vivo. Relaxation times and concentrations of cerebral metabolites. *Magn Reson Med.* 1989;11(1):47-63.
26. Tisdall MD, Hess AT, Reuter M, Meintjes EM, Fischl B, van der Kouwe AJ. Volumetric navigators for prospective motion correction and selective reacquisition in neuroanatomical MRI. *Magn Reson Med.* 2012;68(2):389-399.
27. Simpson R, Devenyi GA, Jeppard P, Hennessy TJ, Near J. Advanced processing and simulation of MRS data using the FID appliance (FID-A)An open source, MATLAB-based toolkit. *Magn Reson Med.* 2017;77(1):23-33.
28. Near J, Andersson J, Maron E, et al. Unedited in vivo detection and quantification of gamma-aminobutyric acid in the occipital cortex using short-TE MRS at 3 T. *NMR Biomed.* 2013;26(11):1353-1362.
29. Near J, Edden R, Evans CJ, Paquin R, Harris A, Jeppard P. Frequency and phase drift correction of magnetic resonance spectroscopy data by spectral registration in the time domain. *Magn Reson Med.* 2015;73(1):44-50.
30. Kreis R. The trouble with quality filtering based on relative Cramer-Rao Lower Bounds. *Magn Reson Med.* 2016;75(1):15-18.
31. Wang Y, Li SJ. Differentiation of metabolic concentrations between gray matter and white matter of human brain by in vivo H-1 magnetic resonance spectroscopy. *Magn Reson Med.* 1998;39(1):28-33.
32. Larsen RJ, Gagoski B, Morton SU, et al. Dataset for "Quantification of Magnetic Resonance Spectroscopy data using a combined reference: Application in typically developing infants. University of Illinois at Urbana-Champaign. 2021. https://doi.org/10.13012/B2IDB-3548139_V1
33. Hedner T, Iversen K, Lundborg P. Central GABA mechanisms during postnatal-development in the rat - neurochemical characteristics. *J Neural Transm.* 1984;59(2):105-118.
34. Clancy B, Finlay BL, Darlington RB, Arland KJS. Extrapolating brain development from experimental species to humans. *Neurotoxicology.* 2007;28(5):931-937.
35. Harbison AL, Votava-Smith JK, Del Castillo S, et al. Clinical factors associated with cerebral metabolism in term neonates with congenital heart disease. *J Pediatr.* 2017;183:67-73.
36. Xu G, Broadbelt KG, Haynes RL, et al. Late development of the GABAergic system in the human cerebral cortex and white matter. *J Neuropathol Exp Neurol.* 2011;70(10):841-858.
37. Tomiyasu M, Aida N, Endo M, et al. Neonatal brain metabolite concentrations: An in vivo magnetic resonance spectroscopy study with a clinical MR system at 3 Tesla. *Plos One.* 2013;8(11):1-7.
38. Miscevic F, Foong J, Schmitt B, Blaser S, Brudno M, Schulze A. An MRspec database query and visualization engine with applications as a clinical diagnostic and research tool. *Mol Genet Metab.* 2016;119(4):300-306.
39. Kok RD, van den Berg PP, van den Bergh AJ, Nijland R, Heerschap A, Heerschap A. Maturation of the human fetal brain as observed by H-1 MR spectroscopy. *Magn Reson Med.* 2002;48(4):611-616.

40. Brighina E, Bresolin N, Pardi G, Rango M. Human fetal brain chemistry as detected by proton magnetic resonance spectroscopy. *Pediatr Neurol*. 2009; 40(5):327-342.
41. Bhakoo KK, Pearce D. In vitro expression of N-acetyl aspartate by oligodendrocytes: Implications for proton magnetic resonance spectroscopy signal in vivo. *J Neurochem*. 2000;74(1):254-262.
42. Kreis R, Hofmann L, Kuhlmann B, Boesch C, Bossi E, Huppi PS. Brain metabolite composition during early human brain development as measured by quantitative in vivo H-1 magnetic resonance spectroscopy. *Magn Reson Med*. 2002;48(6):949-958.
43. Sturman JA. Taurine in development. *Physiol Rev*. 1993;73(1):119-147.
44. Pasantes-Morales H, Hernandez-Benitez R. Taurine and brain development: Trophic or cytoprotective actions? *Neurochem Res*. 2010;35(12):1939-1943.
45. Erickson KI, Weinstein AM, Sutton BP, et al. Beyond vascularization: aerobic fitness is associated with N-acetylaspartate and working memory. *Brain Behav*. 2012;2(1):32-41.

SUPPORTING INFORMATION

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

How to cite this article: Larsen RJ, Gagoski B, Morton SU, et al. Quantification of magnetic resonance spectroscopy data using a combined reference: Application in typically developing infants. *NMR in Biomedicine*. 2021;34:e4520. <https://doi.org/10.1002/nbm.4520>