NMR IN BIOMEDICINE
NMR Biomed. 2004;**17**:361–381
Published online in Wiley InterScience (www.interscience.wiley.com). DOI:10.1002/nbm.891

Review Article

Issues of spectral quality in clinical ^{1}H-magnetic resonance spectroscopy and a gallery of artifacts

Roland Kreis*

Department of Clinical Research, Unit for MR Spectroscopy and Methodology, University Berne, Switzerland

Received 19 March 2004; Revised 30 April 2004; Accepted 7 May 2004

ABSTRACT: In spite of the facts that magnetic resonance spectroscopy (MRS) is applied as clinical tool in non-specialized institutions and that semi-automatic acquisition and processing tools can be used to produce quantitative information from MRS exams without expert information, issues of spectral quality and quality assessment are neglected in the literature of MR spectroscopy. Even worse, there is no consensus among experts on concepts or detailed criteria of quality assessment for MR spectra. Furthermore, artifacts are not at all conspicuous in MRS and can easily be taken for true, interpretable features. This article aims to increase interest in issues of spectral quality and quality assessment, to start a larger debate on generally accepted criteria that spectra must fulfil to be clinically and scientifically acceptable, and to provide a sample gallery of artifacts, which can be used to raise awareness for potential pitfalls in MRS. Copyright © 2004 John Wiley & Sons, Ltd.

KEYWORDS: clinical MR spectroscopy; artifacts; brain; quality control; quality assessment; review; quantitation; human

QUALITY ASSESSMENT

There is extensive literature on clinical and research use of ^{1}H-magnetic resonance spectroscopy (^{1}H-MRS), but, interestingly, extremely little has been published on issues of quality definition and assurance.[1–5] MRS experts seem to take it for granted that it is obvious what makes a spectrum a good spectrum, and how artifacts can be recognized. However, for clinical users of MRS, judgment of spectral quality is far from evident, and artifact recognition tends to be very difficult. Moreover, with the advent of reimbursability of MRS and of semi-automatic data acquisition, data processing and quantitation, the interest in clinical use of MRS has increased and the threshold for non-experts to use MRS in the clinic has been drastically lowered. Unfortunately, MRS is not like MRI where many artifacts are eye-catching. In MRS, pitfalls are at least as ubiquitous, but much less conspicuous. This is highlighted in Fig. 1, with the extreme example of a patient crawling out of the magnet during a scan. This fact would never go unnoticed in MRI (left), but the spectrum in Fig. 1(b) looks perfectly alright—except for a lower signal-to-noise. However, quantitative analysis would reveal a 50% loss of all metabolites and hence yield the diagnosis of vanishing white matter disease.[6]

Furthermore, there is no agreement among experts on what exactly defines a good spectrum. When asked, how they judge the quality of their spectra, the most common answer from an expert will be: 'it depends'. And indeed quality criteria in MRS do depend on whether we deal with single voxel (SV) or spectroscopic imaging (SI) data, on whether long or short echo time (*TE*) was used, on whether near-normal or grossly pathologic spectra are considered, on whether we deal with spectra of the brain (and even which parts of the brain), or other organs, etc. Hence, this article will provide a list of issues important for the judgment of spectral quality and a series of good and bad spectra to demonstrate the artifacts that one should look out for. Even though I have tried to find objective criteria and queried experts in the field,[†] the following should be considered as the personal view of the author, but it is hoped that it will spark discussion and will finally lead to accepted quality factors that will have to be fulfilled for any spectroscopy data to be interpreted clinically or to be published.

The first part presents popular concepts[7–10] for definition of spectral quality (signal-to-noise ratio, linewidth, estimates of fitting errors), followed by a discussion of

**Correspondence to:* R. Kreis, MR Center 1, University and Inselspital, CH-3010 Berne, Switzerland.
E-mail: roland.kreis@insel.ch

Abbreviations used: Ac, acetate; ADC, analog to digital conversion; Ch, cholines; CRMVB, Cramer–Rao minimum variance bounds; Cr, creatines; CSF, cerebrospinal fluid; FID, free induction decay; FD, frequency domain; FWHM, full-width at half-maximum peak height; Glu, glutamate; Gln, glutamine; Glx, Glu + Gln; Gly, glycine; GM, gray matter; Lac, lactate; MRS, magnetic resonance spectroscopy; mI, myo-inositol; NAA, *N*-acetylaspartate; NAAG, *N*-acetylaspartylglutamate; NA, NAA + NAAG; Phe, phenylalanine; PSF, point spread function; QA, quality asessment; ROI, region of interest; RF, radio frequency; rms, root-mean-square; SNR, signal-to-noise ratio; SV, single voxel; SI, spectroscopic imaging; SD, standard deviation; TD, time domain; *TE*, echo time; *TR*, repetition time; WM, white matter.

[†]Strongly influenced by an e-mail poll among many spectroscopists, some of whom were kind enough to communicate their views on this topic. See acknowledgment.

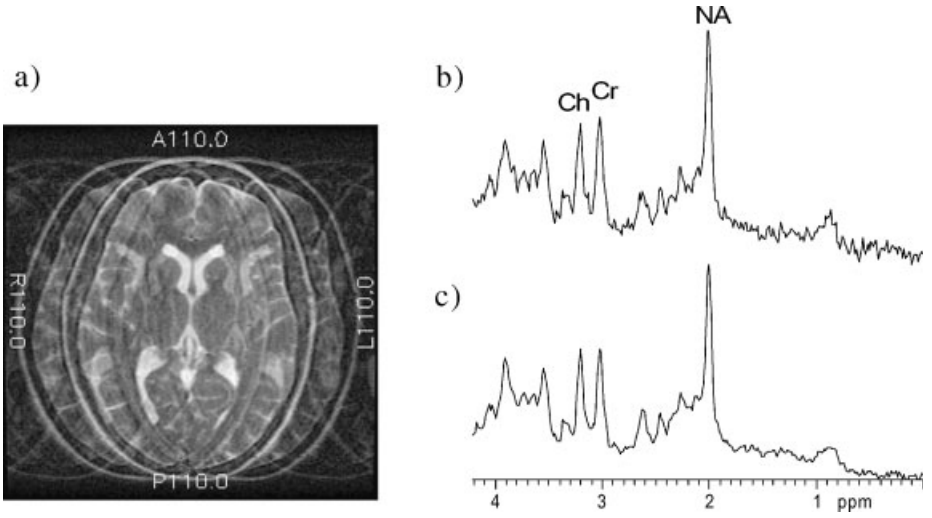

Figure 1. Conspicuity of artifacts in MRI and MRS. If a patient leaves the magnet half-way through a scan, even a layman will refrain from interpreting the resulting image (a). If this happens in a MRS scan, even the expert will not be able to recognize this fact from the resulting spectrum (b), since only signal-to-noise and absolute concentrations will be affected. Spectra (b) (half of the acquired FIDs contain noise only) and (c) (normal acquisition) were scaled to the largest peak, resulting in an apparent signal-to-noise difference, while quantitative analysis would yield a 50% deficit for all metabolites. Unless double-checking mechanisms are put in place and plausibility arguments are used, the resulting diagnosis will be completely wrong. (Scan parameters: 38-year-old healthy woman; MRI, fast spin echo with *TE* 102 ms, *TR* 3 s, 256 × 256, 4 mm slice thickness; MRS, PRESS with *TE* 20 ms, *TR* 3 s, 6.7 cm^3 ROI in periventricular GM, 128 acquisitions)

factors that can potentially lead to bad spectral quality, artifacts and misinterpretation of MR spectra. The final part deals with general concepts on how to use quality criteria in daily clinical practice (rejection criteria, reproducibility, definition of abnormality).

Signal-to-noise ratio

The signal-to-noise ratio (SNR) is often defined in the frequency domain (FD) as the height of the largest metabolite peak divided by the root-mean-square (rms) amplitude (= SD) of the noise in a signal- (and artifact)-free part of the spectrum.[‡] Alternatively, it can be based on signal area in FD. The latter relates to the SNR definition in the time domain (TD), where signal amplitude at time zero is divided by the noise at the end of the FID, where all signals have decayed (or from a signal-free extra scan). This is shown in Fig. 2, where three FIDs and their corresponding spectra are plotted. Each contains a single resonance line at the same frequency, but differing width. The SNRs of the FIDs in TD are all equal (=10), such that in FD SNR defined as signal area divided by noise is also equal, but SNR based on the peak height differs between the spectra. If FD intensity is used, SNR depends inversely on linewidth. Low SNR can be remedied by choosing larger regions of interest (ROIs—SNR proportional to ROI size), increased scan time (SNR proportional to square root of number of acquisitions) or optimized hardware (local receive coil, higher field). The dependence on total scan time is illustrated in Fig. 3. The top left spectrum was recorded with a short *TE* PRESS sequence as the average of 64 scans from a 10 cm^3 ROI in perisylvian gray matter (GM) in a 7-year-old boy. The spectrum was analyzed with LCModel[11] and yielded an SNR of 33. The fitted spectrum and the corresponding difference between experimental and fitted data (residues) are plotted on the right-hand side. Excerpts from the same recordings are displayed below, where only some of the individually stored free induction decays (FIDs) had been used to produce a spectrum. The lowest spectrum was from a single-shot acquisition giving an SNR of 4, confirming the expected square root dependence of SNR [33/4 ≈ sqrt(64)]. The tissue content and its expected error band for some exemplary metabolites, as determined from these spectra by LCModel, are given in Fig. 3. Upper and lower confidence limits were calculated as plus or minus twice the fitting error (comparable to the Cramer–Rao minimum variance bounds—CRMVB,[12] see below). In general, the confidence interval narrows for increased SNR, and CRMVB decrease in proportion with the rms noise level (unless a distinctly different fitting solution is found at different noise levels). At very low SNR, grossly wrong content estimation is possible and the CRMVB may yield error bounds that seem to be too narrow. (As seen below, CRMVB only provide a lower limit for the actual errors and—to be exact—would have to be calculated with true, not estimated, model parameters.) Systematic overestimation is

[‡]In LC-Model,[11] it is not signal vs noise, but vs rms of the residues.

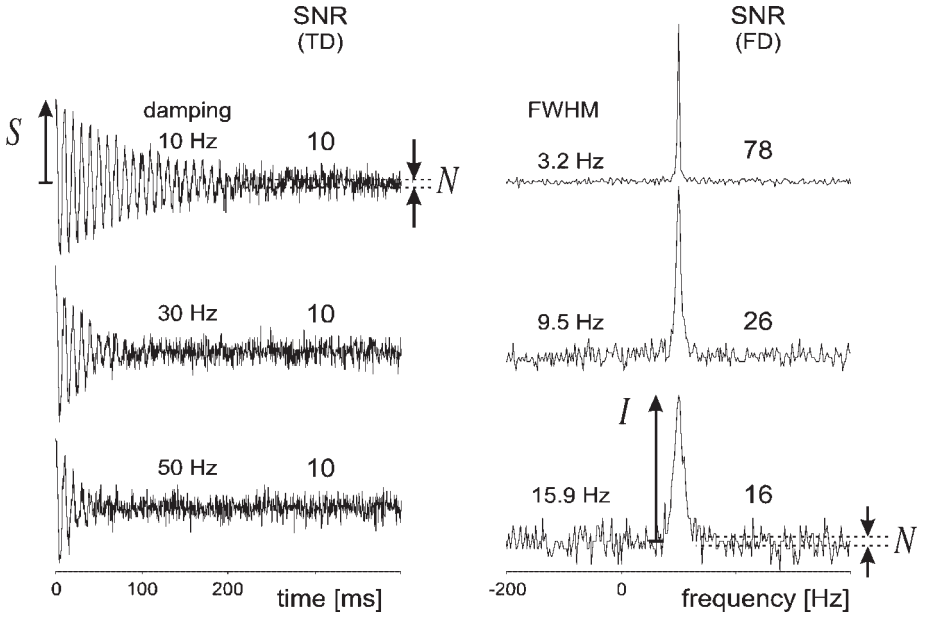

Figure 2. Definition of SNR. SNR can be defined in TD (left) as initial signal amplitude (S) vs standard deviation (SD) of the noise (N). This is independent of lineshape. The equivalent in FD would be peak area vs noise in a signal-free region of the spectrum. However, in the spectrum SNR is more often defined as peak intensity (*I*) vs noise (right side), which depends on line shape and width. The figure contains synthetic data, produced with the jMRUI[81] data processing package. The FD spectra were scaled to match the size of the peak (1024 points, 2500 Hz width, 10% noise level in TD, exponential damping of 10, 30 and 50 Hz, respectively)

likely for low-concentration metabolites[13] and systematically wrong results are also possible for major peaks under certain conditions.[14]

The SNR as determined from FD amplitudes can be a misleading criterion to judge the quality of spectra, since FD-SNR depends on the length of the acquisition window, real-time and post-acquisition filtering, and apodization. Rapidly varying noise in the spectrum, arising from noise towards the end of the TD signal, gives the impression of low SNR, but is largely irrelevant for the precision of peak fitting, particularly if fitting restrictions or prior knowledge on line widths is used. This is illustrated in Fig. 4, where the same spectrum is presented and fitted in three variations. The top spectrum is the originally measured spectrum (0.512 Hz digital resolution, 1.05 s length of FID). The middle spectrum was obtained by adding in additional (measured) noise. As expected, the CRMVB increase for all metabolites. The lowest spectrum appears to have a better SNR than the middle one, because the FID of the middle spectrum was truncated to 0.524 s and then zero-filled to the original length. The fitting errors, however, remain at a similar level or even increase.

SNR is often used to discard bad spectra. However, as SNR is directly reflected in the CRMVB and CRMVB are more directly linked to confidence limits, a rejection criterion based on CRMVB for each metabolite seems more effective. SNR may be convenient as rejection criterion when integrating well isolated peaks (often done in SI with long *TE*), when no peak fitting is performed.

Linewidth/lineshape

Linewidth is usually defined independently of lineshape as the full-width at half-maximum peak height (FWHM) in FD. It determines the resolution available to discern spectral features. This is demonstrated in Fig. 5, which contains two spectra obtained subsequently from the same ROI in the same subject, but with intentionally decreased shim quality. With larger FWHM the singlets of Cr and Ch start to overlap (dashed ellipse), but are still well defined. However, all distinct features of the Glu/Gln region are lost with bad linewidth (dotted ellipse). If linewidths are variables in model fitting, they strongly influence the calculation of CRMVB for peak areas. For example, in the case illustrated in Fig. 2, the CRMVB for peak areas increase from 1.3 to 2.1 to 2.7% with increasing linewidth, i.e. they are inversely proportional neither to the SNRs defined in TD, nor to those from FD.

It turns out that linewidth is critical for model fitting, and bad resolution easily leads to meaningless results in short *TE* spectra. Rejection criteria based on FWHM are very useful for automatic screening of SV[15] and SI[16] data. The effect of reduced resolution (and inherently lower intensity SNR) is illustrated in Fig. 6, which contains more PRESS spectra of the same gray matter ROI, but recorded with intentionally varied shim settings (same case as in Fig. 5). The CRMVB given for some of the metabolites demonstrate that increased linewidths directly lead to a severe increase in fitting inaccuracies, primarily for the minor contributors to the spectrum,

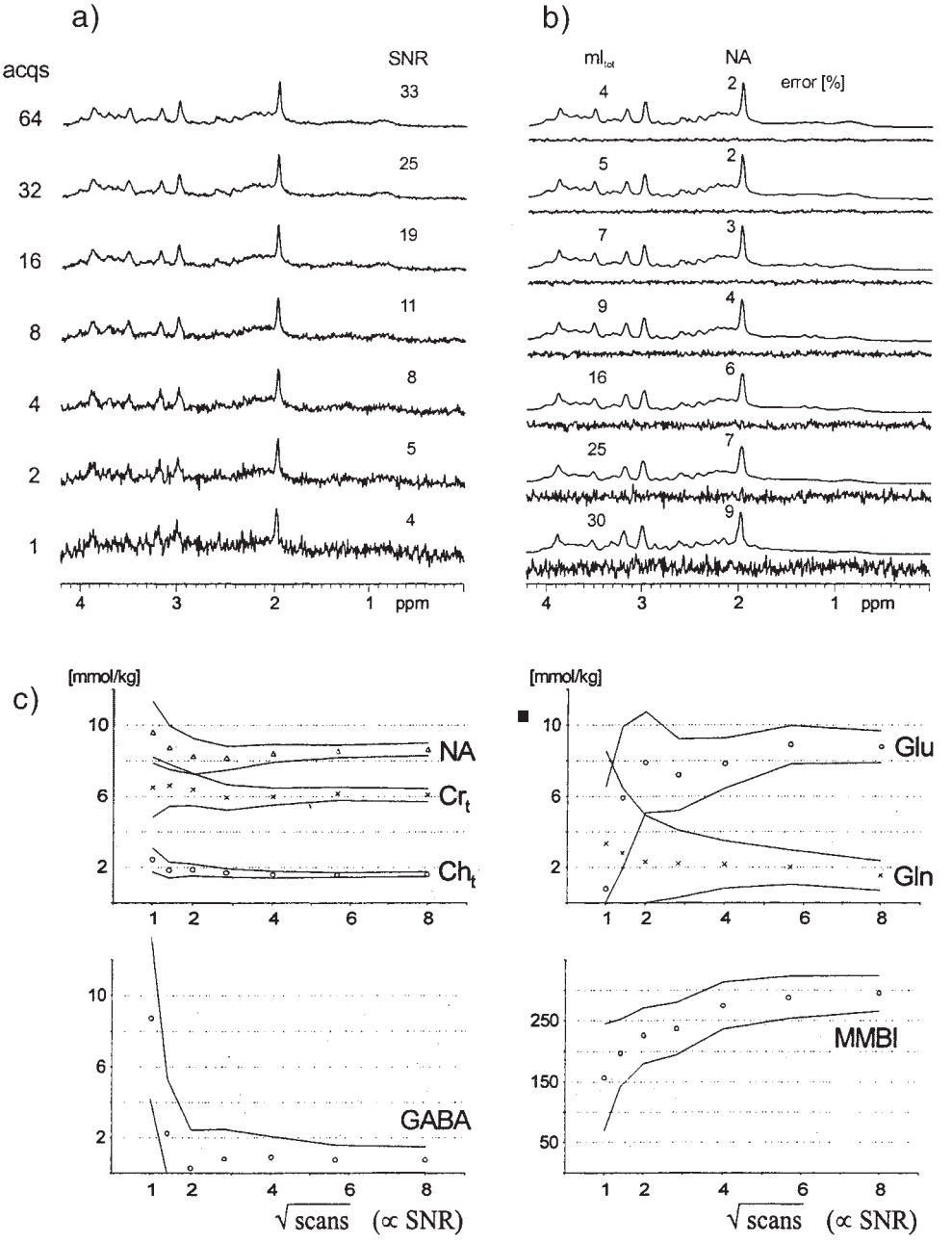

Figure 3. Scan time dependence of SNR and error estimates. Improvement of SNR with longer scan time is illustrated with a spectrum from perisylvian GM in a 7-year-old boy. The top row [experimental spectrum in (a), fit and residuals in (b)] contains the summed data from 64 acquistions. Below, spectra were constructed from parts of the full scan (1, 2, 4, 8, 16, 32 acquistions of the total). SNR scales with the square root of the number of acquistions (i.e. scan time), because signal adds up linearly, while noise only increases with a square root dependence. (b) Estimates of the fitting error for the two metabolites NA and mI_{tot}. The dependence of further estimated concentrations (mean $\pm$ 2 SD) for this case is plotted in (c). The confidence interval clearly decreases with increasing SNR for all metabolites. However, some metabolites (in the current case in particular the macromolecular baseline) can be systematically wrongly estimated at low SNR.[13,14] (NAA, *N*-acetylaspartate; NAAG, *N*-acetylaspartylglutamate; NA, NAA + NAAG; Gly, glycine; mI, myo-inositol; mI_{tot}, mI + Gly; Cr_t, total creatine; Ch_t, total cholines; GABA, γ-aminobutyrate; Glu, glutamate; Gln, glutamine; MMBl, macromolecular baseline. Scan parameters, PRESS with *TE* 20 ms, *TR* 3 s, 9.6 cm^3 ROI in periventricular GM, fitted with LC-Model[11,28])

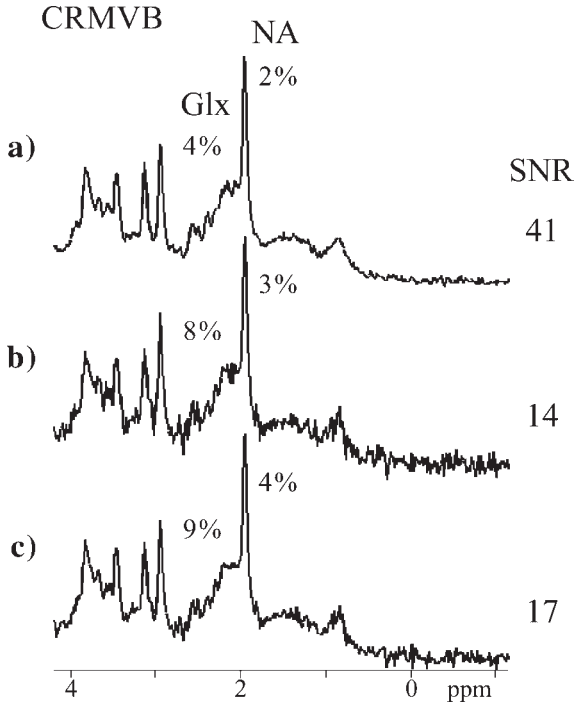

Figure 4. Apparent SNR and fitting errors. Apparent SNR in FD depends on processing. The same spectrum is presented with different apparent SNR: (a) original spectrum; (b) additional noise added; (c) same noise as (b), but second half of FID (0.51–1.02 s) replaced by zeroes. The spectrum in (c) has an apparently better visual appearance and FD-SNR, but fitting errors from LC-Model are somewhat larger, because noise is estimated from the residuals, which tend to increase, because part of the information content was reduced by zeroing in TD. (Scan parameters: occipito-parietal GM, 12 cm^3; 38-year-old man; PRESS, *TE* 20 ms, *TR* 1.6 s, 96 scans)

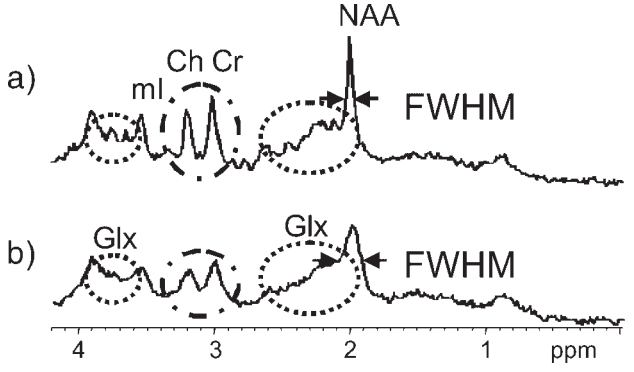

Figure 5. Definition of linewidth. Two spectra from the same ROI in the same subject illustrate the definition of linewidth as FWHM. An increase in linewidth from 0.5 (a) to 0.9 ppm (b) leads to increased overlap of the singlets from Cr and Ch, which is not yet prohibitive for their peak area estimation. However, the complete loss of distinguishing features in the minor contributors (e.g. the patterns of Gln and Glu, dotted ellipse) represents a severe loss of information (scan parameters as in Fig. 4)

which are mostly not well-isolated singlet peaks, but rather broad overlapping patterns.

It was recently found that reduced resolution and decreased SNR in combination with complicated base-

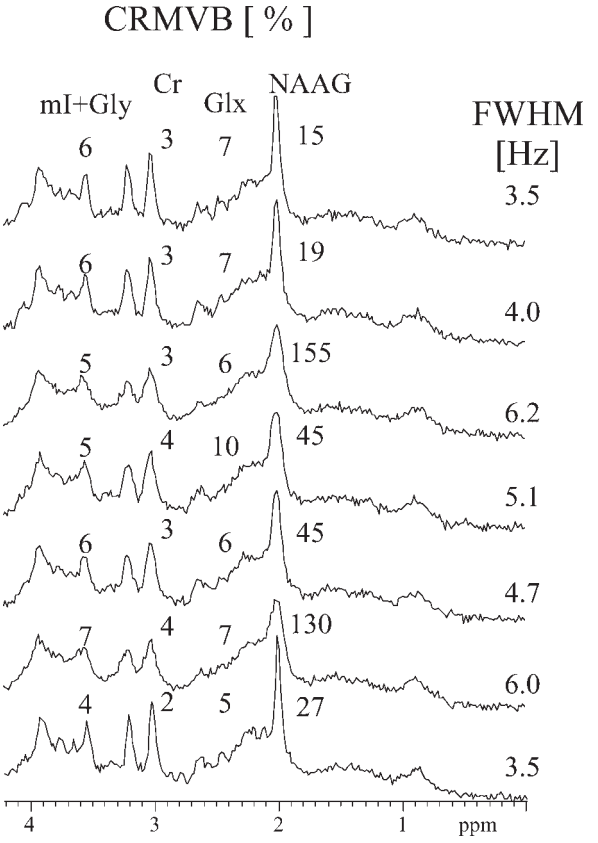

Figure 6. Linewidth and fitting errors. Short-*TE* PRESS spectra that were obtained sequentially with different shim settings but from the same ROI within the same examination are used to show the dependence of fitting error on the lineshape. A 2-fold range of FWHM (estimated from LC-Model) leads to a 2–10-fold increase in estimates of fitting error (scan parameters as in Fig. 4)

lines and possibly inaccurate fitting models can lead to systematic over- or underestimation of low-concentration, but also prominent metabolites,[13,17] which in part will not be reflected in increased CRMVB.[14] Remedies for bad resolution include: better shimming (using higher order shims, if available), smaller ROI size or moving the ROI away from tissue interfaces. A further way to minimize the impact of a bad lineshape is to determine the lineshape with reference scans or with field maps and subsequently correct the distorted FID based on the known shape.[18–20] Similarly, the known lineshape can be used as prior knowledge in model fitting.[21] The success of such schemes to correct or account for shape deviations in terms of preventing increases in CRMVB has not been extensively investigated.

Errors, variability, reproducibility

Quantitative results should always be given with error estimates. The stochastic uncertainty for a single measurement is reflected in the CRMVB.[12] They represent lower bounds for the variance of the estimated parameters, given a certain model, a set of parameters and a

noise level. Hence they include effects of SNR and the inherent limits of fitting with a given model, including the interdependence of fitting parameters and the effect of prior knowledge constraints. Simple aspects of the influence of the inherent limits of a fitting model at particular points in parameter space are visualized in Fig. 7. From this example it becomes clear that CRMVB are clearly superior as parameters to judge the quality of quantified data than mere statements of SNR or FWHM. The figure also illustrates how prior knowledge can substantially improve the fitting accuracy. It is important to note that CRMVB are calculated under the premise that the fitting model and prior knowledge constraints are correct. Any systematic errors are neglected and may lead to over- or underestimated confidence limits.

For the definition of normal ranges, it is necessary to determine reproducibility. Reproducibility must be established locally and cannot usually be inferred from the literature. Ideally, reproducibility would be determined in a patient population,[14,22] not in healthy subjects, but this is usually very difficult to achieve. With regard to absolute quantitation in standard units, the systematic errors are often substantial. Possible sources of systematic error include inaccurate values for T_1 and T_2, inaccuracies in reference measurements and calibrations, operator dependence, limited MR visibility, metabolite compartmentation, wrong baselines and lineshape models. The random errors reflected in the reproducibility can be much smaller[23] than systematic errors and should not be taken as a measure for the latter. The quantitative method with best reproducibility by no means guarantees results closest to the true values. However, best reproducibility guarantees most sensitive detection of pathology.

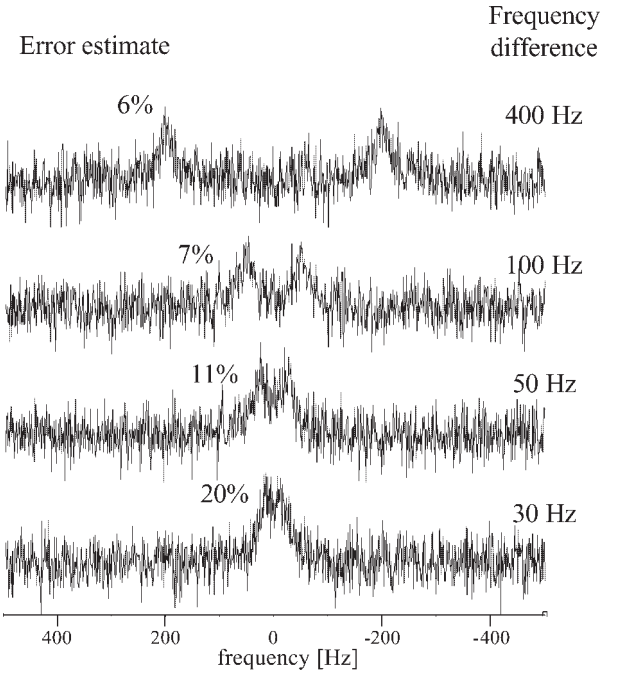

Figure 7. Inherent limits of a fitting model. Cramer–Rao minimum variance bounds (CRMVB) depend on SNR and the inherent limits of the fitting model, realized at the true parameter values. This means that fitting errors depend on the interdependence of the fitting parameters. In the current example the accuracy of peak area estimation depends on the frequency difference between the two peaks in this spectrum. In the well-isolated case (top row, 400 Hz frequency difference) the fitting error is 6%. This increases up to 20% in the bottom row (30 Hz frequency difference). SNR (10 in TD) and all other parameters are equal for all rows. If prior knowledge constraints can be imposed, fitting uncertainties can be severely reduced. In the current case, knowledge of resonance position and phase, relative area ratio (= 1), and difference in width (= 0) reduces the error estimate from 6 to 4% in the top spectrum and from 20 to 7% in the lowest spectrum. [Spectrum details: simulated data and error estimation from jMRUI vs 2.1,[81] 1024 points, 1 kHz spectral width, Lorentz lines with 100 Hz damping (32 Hz linewidth), 0° phase and same area for both lines]

The achievable reproducibility has been published for several methods and can serve as an indication of what is feasible, even though the results scatter considerably. The overall reproducibility is determined by the square root of the additive variances (σ^2). Three sources of variability should be distinguished: (1) variability upon immediate repetition of a scan (σ_{rep}; closely related to CRMVB[24]); (2) intra-individual variability upon re-examination of the same subject in a subsequent scan (σ_{intra}); and (3) inter-individual variability for examination of an equivalent ROI in different subjects (σ_{inter}). σ_{rep} includes effects of the noise in the spectrum, but also, for example, short-term system and subject instabilities. σ_{intra} reflects additional variance due to imperfect repositioning of the subject and the ROI, as well as potential temporal fluctuations in metabolite content and long-term system instabilities. σ_{inter} summarizes contributions from true individuality in metabolite content, but also variations because of differing ROI composition due to different morphology. Overall variability for re-examination of a certain ROI within the same subject in repeated exams is then given by $\sqrt{(\sigma_{rep}^2 + \sigma_{intra}^2)}$, and overall variability for examination of a certain ROI within different subjects by $\sqrt{(\sigma_{rep}^2 + \sigma_{intra}^2 + \sigma_{inter}^2)}$. The following values represent ranges of reproducibility taken from the recent literature for NA, Cr, Cho and mI:

- σ_{rep} 3–22%,[24] σ_{rep} 4–17%,[25§] σ_{rep} 3–7%;[26]
- σ_{intra} 1–4%,[26] $\sqrt{(\sigma_{rep}^2+\sigma_{intra}^2)}$ 9–16%,[25§] $\sqrt{(\sigma_{rep}^2+\sigma_{intra}^2)}$ 4–7%,[27] $\sqrt{(\sigma_{rep}^2 + \sigma_{intra}^2)}$ 4–8%;[17]
- σ_{inter} 4–9%,[27] $\sqrt{(\sigma_{rep}^2 + \sigma_{intra}^2 + \sigma_{inter}^2)}$ 8–15%,[17] $\sqrt{(\sigma_{rep}^2 + \sigma_{intra}^2 + \sigma_{inter}^2)}$ 7–16%.[28]

Additional variation will of course occur, if cerebral regions are included that are not strictly equivalent, e.g. if different WM or GM structures are pooled. It should also be noted that σ_{rep} (related to SNR) and σ_{intra} will depend

§Without mI.

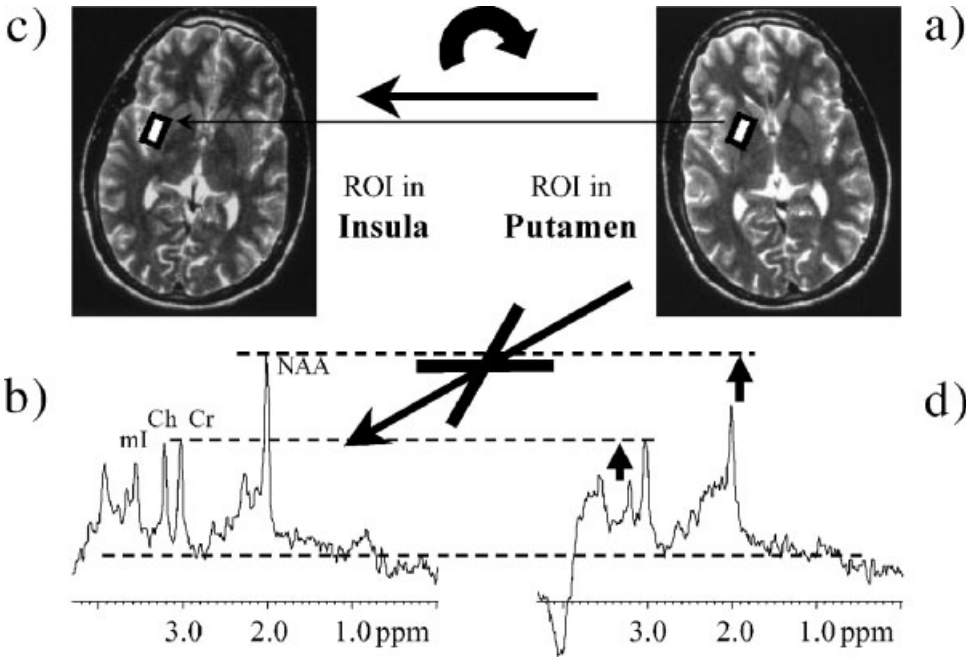

Figure 8. Effect of gross head movement. An ROI was placed in the putamen on a localizer image (a). The resulting spectrum is shown in (b). However, unknown to the operator, the examined subject had turned his head a little towards his left, which was picked up from the repeated localizer scan after the MRS acquisition (c). As ROIs are prescribed in magnet coordinates, the ROI targeted at the putamen ended up in insular GM, such that the spectrum in (b) was in reality acquired from insula, not putamen. This explains the narrow lines, which are atypical for basal ganglia. A spectrum from the putamen was then acquired (d) and shows that a completely wrong diagnosis would have resulted when the spectrum from insular GM was taken as originating from putamen. Verification of proper ROI placement is crucial. (Scan parameters: 26-year-old man; MRI: fast spin echo sequence, echo train length 16, *TR* 3 s, *TE* 100 ms; MRS: 2.2 cm^3 ROI, PRESS, *TE* 20 ms, *TR* 3 s, 128 acquisitions)

on the size of the ROI. This is evident for σ_{rep} because of the SNR argument, but since the success of exact repositioning depends on the size of the ROI, σ_{intra} has also been shown to vary with ROI size.[29]

The above values are for SV studies; similar numbers can be found for SI data.[29–32]

GALLERY OF ARTIFACTS

In the following, several sources of artifacts are discussed and examples of how to recognize and remedy these effects are given. At present, artifact detection has not been automated and requires an experienced eye. The prime focus of this article and in particular this section is single-voxel ^{1}H-MRS of the brain. Most effects are also relevant for SI, but may manifest themselves differently in SI and these specifics are not covered in the following. A selection of artifacts has been put onto the Internet[¶] by the MRS team at the University of Berne and readers are invited to submit artifacts and pitfalls to this address, such that clinical users of MRS can learn from the errors of others and potential misdiagnoses can be prevented.

[¶]www.cx.unibe.ch/dkf1/amsm/MRS_artifacts

ROI location

In single-voxel MRS, there is no direct spatial information inherent in the spectrum. The correct prescription of the ROI, be it manual or graphic, should always be checked to avoid wrong diagnoses because of operator errors or patient movements. Control of data header information, acquisition of an ROI image and the repetition of scout images after the recording of the MR spectrum are advisable. An example of how much difference a fairly small movement of the head can make even in a healthy control without any targeted focal lesions is presented in Fig. 8. As outlined in Fig. 8(a), it was intended to record a spectrum from within the right putamen. The resulting spectrum, Fig. 8(b), was of good quality with nicely narrow lines. However, repetition of the scout images revealed that the subject had turned his head somewhat to his left, Fig. 8(c), after acquisition of the first image, such that the ROI, defined in magnet coordinates, turned out to have been placed in insular GM, rather than the putamen. The proper spectrum of the putamen is plotted in Fig. 8(d) and features substantially wider lines and, compared with the spectrum from the insular region, much lower NA and Ch peaks. If patient movements go unnoticed, the resulting spectra can clearly lead to completely wrong

diagnoses. This can also happen for considerably smaller movements, particularly if small (narrow) ROIs are examined, where ROI composition changes quickly with small movements, or if tissue segmentation (WM/GM content) or compartmentation algorithms are used for absolute quantitation. In the current case, an experienced spectroscopist might have become suspicious, because spectra of the iron-rich putamen do not feature such narrow lines, but in patients any abnormalities are usually attributed to disease, not improper localization.

Movements

In SV MRS, repeated small gross motions or local pulsatile motion (cardiac-related CSF pulsation, respiration) are normally reflected in increased linewidths, overall frequency shifts, possibly reduced peak areas (phase cancellation)[33] and decreased quality of water suppression. Single events of gross bulk motion will obviously result in recording data from a wrongly localized ROI (see above). If the movement occurs during acquisition of the spectrum, this may become evident by a doubling of all peaks (see Fig. 9). A post-acquisition correction is only possible if all acquisitions were stored separately and if signals are present that allow for realignment and/or individual phasing. In SI, motion leads to spatial blurring,[34,35] readily evidenced by substantial metabolite signals found outside the head.

Signal phase

The zero-order phase must be correct for visual inspection of a spectrum, and good phasing helps to find the global χ^2 minimum in fitting. Good phasing is, however, not crucial for quantitative evaluation, if the phase is a variable in the fit. The correct phase can be obtained from a reference scan. This is recommended, particularly for single voxel spectroscopy, where it takes minimal extra time (single acquisition without water suppression) and eliminates operator influence.

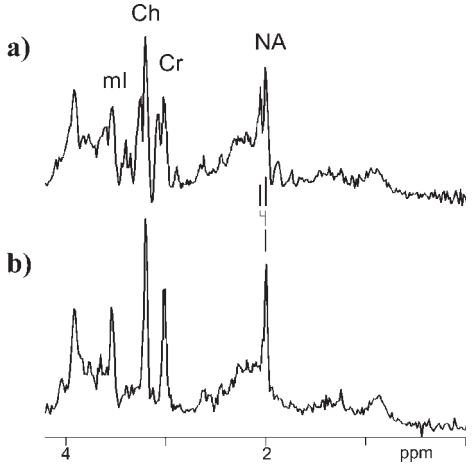

Figure 9. Effect of head movements. All peaks are doubled in a spectrum from a neonate,[21] because the baby had moved its head between two distinct positions during the scan (a). The repeat examination shows single peaks with perfect shim and lineshape, when the baby was soundly asleep. (Scan parameters: neonate of 41 weeks gestational age, ROI in thalamus, PRESS *TE* 20 ms, *TR* 2 s, 128 acquisitions)

In spin-echo-based localization schemes, no first-order phasing should be needed, since signal acquisition should start at the center of the echo. FID-based localization sequences suffer from a delayed start of signal acquisition and therefore first-order phasing is needed for correct display of the spectrum. Because the delay is defined by the pulse sequence, the first-order phase is constant and does not have to be fitted together with the other spectral parameters, but can be calibrated in phantom scans.

There is, however, an artifactual mechanism that can shift a spin echo and first-order phasing may then become relevant. In localization sequences involving the formation of spin echoes, signal from outside the ROI is crushed by gradient spoiler pulses. To obtain the full signal, crusher pulses have to be balanced with respect to the 180° pulse and the timing of RF pulses with respect to the slice selection gradients has to be properly adjusted to obtain full refocusing. Improper adjustments or eddy currents can lead to effectively non-symmetric gradient pulses. If this is combined with static B_0 gradients caused by local inhomogeneities or bad shimming, the maximum echo signal is shifted to times before or after the nominal echo time. The effect of bad gradient crusher trimming is illustrated in Fig. 10. On the left, the shift of the maximum echo signal is clearly visible in the FID in the three lower traces—in these examples provoked by intentionally mis-setting the gradient crusher amplitudes and shim settings. In FD, shifted echoes lead to characteristic lineshapes that feature negative 'feet', most clearly noticeable in the lowest spectrum on the right of Fig. 10. This lineshape is similar to what is obtained with extreme resolution enhancement by Lorentz–Gauss transformation. Quantitation with standard line shape models would find reduced peak areas. In the current example, net peak areas range from 100% in the top trace to 89, 56 and 19%, respectively. Mild misadjustment as in row (b) is most dangerous, because it is hardly noticeable by eye. If crusher gradients are not balanced, manual shimming guided by maximum FD signal intensity as target will automatically lead to a shifted echo and inhomogeneous B_0, if no attention is paid to signal shape or FID. If shifted echoes are observed, this calls for urgent gradient readjustment. If eddy currents are causing the ill-balanced crusher gradient shape, readjustment may be spatially dependent and *TE*-dependent.

ROI shape

It is well known that the localization profile of slice selective RF pulses is not perfect and that voxels

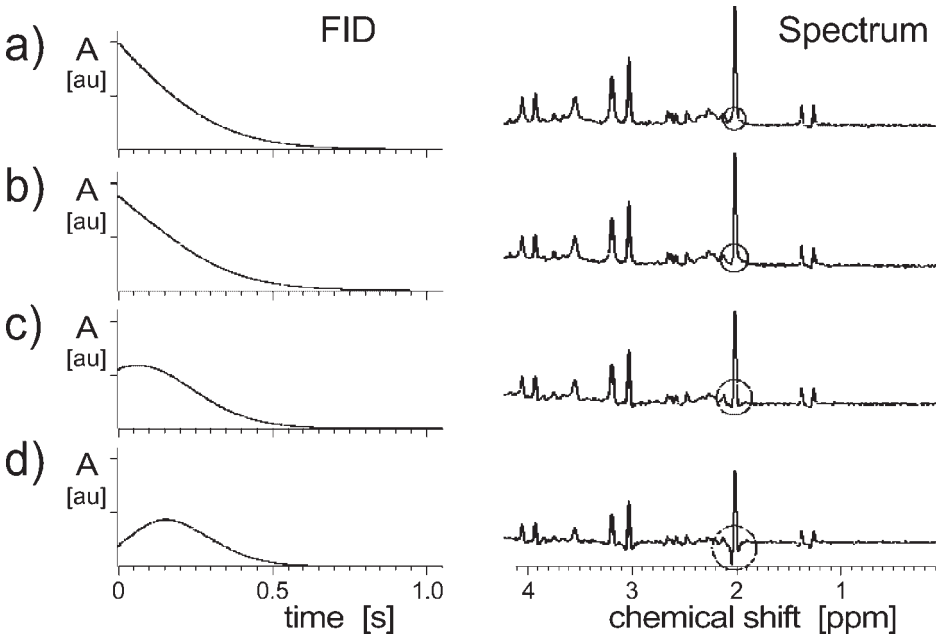

Figure 10. Effect of unbalanced gradient pulses. Unbalanced gradient crusher pulses in combination with B_0 gradients, which make up for the missing crusher gradient amplitudes, lead to a shift of the echo maximum away from the nominal echo time. Shifted echoes lead to characteristic negative feet for each peak, highlighted by the circles. This is illustrated in PRESS spectra (*TE* 20 ms, *TR* 3 s) of a metabolite solution. A shift of the double spin echo was provoked by deliberately misadjusting the relative amplitudes of crusher gradient amplitudes and manual shimming to obtain high peak intensities. (a) FID and spectrum from the well tweaked sequence. For (b)–(d) gradient amplitudes were more and more unbalanced, leading to net peak areas relative to (a) of 89, 56 and 19%, respectively. (Metabolite solution in a QA phantom from General Electric Medical System, Milwaukee, USA, consisting of an aqueous phosphate-buffered solution of NAA, creatine, choline, mI, Glu and Lac, with sodium azide and Gd-DTPA added)

localized by STEAM or PRESS sequences are not perfect cubes, i.e. with equal response within the ROI and zero response from outside. Still, it is very useful to know the actual voxel profile,[36,37] when judging spectra from focal lesions. It not only depends on the kind of RF pulses used (e.g. damped sinc, Gauss or numerically optimized pulses[38]), but also changes drastically for incorrect B_1 settings.[36] For each scan, the appropriate RF voltage to obtain correct flip angles (i.e. correct B_1) has to be determined in the prescan calibration. This is needed because different composition and sizes of human body parts and relative positioning with respect to RF transmit coils lead to individually different electrical loading of the RF coils. If the prescan procedure is not working properly or is inaccurate, or if the patient moves between prescan and scan, spectra will be recorded with wrong RF pulse characteristics. This leads to higher or lower net signal and also to non-ideal spatial selection profiles. This is demonstrated for a PRESS sequence in Fig. 11. The top part contains the pulse profiles as a function of transmitter gain. The automatic prescan procedure adjusted the RF voltage to produce the ROI profile labeled B_1 (top, right). Manual re-adjustment of the RF voltage by +1 dB (+12%), −1 dB (−12%), and −2 dB (−26%) yielded the other three plotted profiles. The selected ROI seems to be excited most uniformly using the setting from auto prescan. Decrease or increase in RF power leads to less uniform excitation within the selected ROI and noticeable excitation of regions outside the prescribed ROI. The corresponding spectra show that excitation with a B_1 amplitude lower than the optimum set by autoprescan produces larger net signals, mainly due to excitation of a larger ROI. The figure also shows that the absolute signal response from this particular PRESS sequence with numerically optimized RF pulses is very sensitive to B_1 calibration. Absolute quantitation procedures relying on signal acquisition at two different ROIs (external standard) or those interpreting the RF voltage needed for a 90° pulse (reciprocity principle, see below) are very susceptible to any error in flip angle calibration.

Chemical shift artifact

The targeted ROI, as presented in Fig. 11, is only selected for on-resonance excitation. Peaks from metabolites at other chemical shifts are excited at slightly shifted voxel locations. The size of this effect depends on the particular RF pulses used and on whether different voxel sizes are realized with RF pulses with variable bandwidth and constant gradient strength or with RF pulses of constant bandwidth, but varying gradient strength. Table 1 gives a

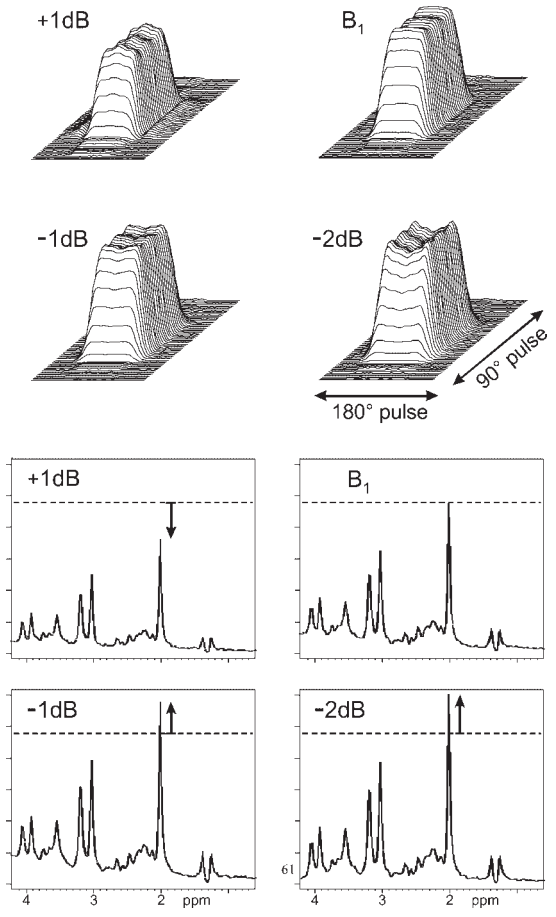

Figure 11. Voxel profile as function of RF pulse amplitude. Real-world voxel profiles from a commercial PRESS sequence, optimized for short echo time, are illustrated by measurements on a calibration phantom (see Fig. 10). The top part contains the pulse profiles as a function of transmitter setting. The automatic prescan procedure adjusted the RF voltage to produce the ROI profile labeled B_1. Deliberate misadjustment of the RF voltage by ± 1, and –2 dB yielded the other three profiles, leading to less uniform excitation within the selected ROI and noticeable excitation of regions outside. The corresponding spectra are plotted below, showing that the best excitation profile does not correspond to maximum signal. Detail of the acquisitions: PRESS with *TE* 35 ms, *TR* 2 s, excitation with numerically optimized RF pulses implemented in standard localization sequences on GE scanners (no outer volume saturation pulses, no reduced flip angle pulses)

realistic example where varying gradient strength is used to change voxel dimensions. The minimum resolution (i.e. at maximum gradient strength) is assumed to be 4 mm and two cases with either 10 or 20 mT/m gradient strength at 1.5 T (corresponding to 20 or 40 mT/m at 3 T) are listed. For relatively small voxels ($\leq$20 mm linear dimension) the slice shift is almost negligible ($\leq$ 1.2 mm) and there is an overlap of 86% for the ROIs of all relevant metabolite peaks in the upfield region (1.3–3.9 ppm)—at least for today's systems with at least 20 mT/m gradient strength. To minimize the voxel shift, the transmit frequency should be set to the middle of the spectral region of interest (e.g. around 2.5 ppm for acquisition of water-suppressed scans and on resonance with water for reference acquisitions). At 3 T, all these shifts double, if the same gradient strength is available, and the common voxel volume drops to 69% for this example! This shift can be even more appreciable if the spectral region is larger, e.g. if Cr is used as internal standard for the quantitation of phenylalanine, only a little more than half of the signals originate from the same cerebral volume (10 mT/m, 1.5 T). Additionally the linear voxel shift for the example of Phe determination, where voxels of up to $50 \times 70 \times 20$ mm are used, becomes substantial ($>$1 cm for Phe vs Cr).

Particular care with respect to the chemical shift artifact should be taken when left–right comparisons are made. As the direction of voxel shift due to the chemical shift artifact depends on the sign of the slice-selection gradient, the ROIs are shifted in the same direction in both cerebral hemispheres. This means that the two ROIs are not symmetrically placed in the brain. This can easily lead to apparent left/right asymmetries in the spectra, which in reality are due to asymmetric ROI location because of the chemical shift artifact. For proper left–right comparison the slice selection gradient for the left/right direction has to be inverted before recording the second spectrum.[23]

A further complicating factor occurs if PRESS or STEAM is combined with additional outer volume suppression pulses that are not based on the same bandwidth and/or gradient direction. In this case the effective ROI size and position depend on the nominal ROI size, the width of the suppression band and the chemical shift of

Table 1. Effect of the chemical shift artifact for three-dimensional slice selection sequences (STEAM, PRESS)

	Frequency difference (ppm)	Shift (mm), 20 mm slice	Shift (mm), 70 mm slice	Common volume (%)	Shift (mm), 20 mm slice	Shift (mm), 70 mm slice	Common volume (%)
Hardware configuration		20 mT/m @ 1.5 T or 40 mT/m @ 3 T			10 mT/m @ 1.5 T or 20 mT/m @ 3 T		
Water/lip (—CH_2—)	3.4	1.3	4.5	82	2.6	9.1	66
Cr (—CH_3)/lip (—CH_2—)	1.75	0.7	2.3	90	1.3	4.7	81
Cr (—CH_2)/lip (—CH_3)	2.65	1.2	4.0	84	2.3	8.1	69
Cr (—CH_3)/Phe	4.32	1.6	5.7	77	3.3	11.5	58
Lip (—CH_3)/Lip (—CH═)	4.45	1.7	5.9	77	3.4	11.8	57

(Example based on assumption that minimum ROI size is 4 mm and that the same RF pulse with varying gradient strength is used for different ROI sizes).

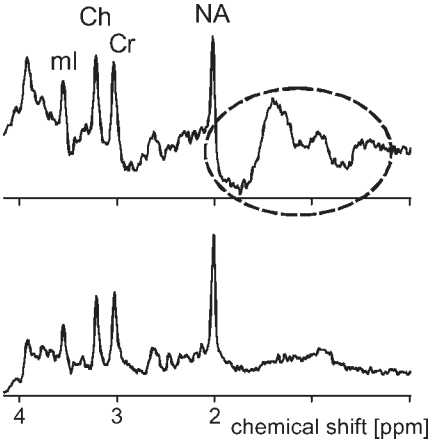

Figure 12. Outer volume lipid contamination. Non-ideal slice selection profiles lead to contamination with outer volume signals. This is particularly relevant if the pulses produce a non-negligible net excitation in areas with lipid deposits that intrinsically give rise to huge signals compared with the metabolite levels. Two fairly old spectra of two male children from a previously published study on brain development, both recorded with a STEAM sequence with identical parameters (*TE* 30 ms, *TR* 1.5 s), and similar ROI position (periventricular WM) demonstrate this effect. The top spectrum shows lipid contamination with unclear signal phase, while the lower spectrum demonstrates, what this spectral region normally looks like (adapted from Kreis *et al.*[40])

the peak.[28] This problem is circumvented if the outer volume suppression pulses define the ROI shape.[39]

Outer volume signal bleed

As seen in Fig. 11, ROI selection pulses are not infinitely selective and will always excite spins outside the targeted ROI as well (particularly within the smooth transition band and an ROI width around the ROI). As the out-of-volume signal will normally only be a small fraction of the signal from the targeted ROI, this will not be relevant in areas with small spectral changes between the inside and the outside of the ROI. The opposite is true if the surrounding tissue provides much stronger signals (e.g. lipids) or if the selected ROI is suspected to feature a spectrum that strongly deviates from the neighboring tissue (focal lesion). Often, signal from well outside the selected volume is characterized by a different phase than the signal from within the ROI. This can be found very often in the older spectroscopy literature, where short *TE* spectra were used, but the RF pulses and gradient crushing schemes on routine scanners were not as optimized as today. Figure 12 illustrates this with spectra from a study on neonatal brain development, conducted in 1990–1992.[40] The encircled region in the top spectrum contains unclear and unphased lipid contributions that are likely to be contaminations from outside the selected ROI, while another spectrum from a different child, but acquired with the same technical parameters, does not show such signal contributions. RF-phase cycling will only reduce this artifactual signal if it is due to a spurious echo or FID formation from an unwanted combination of RF pulses (see outer volume ghosts below). If the outer volume signal is due to inherently non-ideal pulse selection profiles, RF phase cycling is ineffective, but spatial saturation bands placed around the ROI can usually prevent these signals. Another way to identify or eliminate interfering outer volume signal is by shifting the ROI or reducing its size. This is demonstrated in Fig. 13, which contains spectra from superficial temperoparietal cortex. Figure 13(d) and (e) shows the original voxel placement and Fig. 13(a) the resulting spectrum, which is contaminated by signal from outer volume lipids and also disturbed by residual water signal arising from a broad water resonance in close proximity to the skull. A small shift of the ROI prevents most of the interfering lipid signal [Fig. 13(f) and (b)]. The disturbing signal could be eliminated completely when the longest voxel dimension, which was pointing in the direction of the skull, was reduced in size, even though the outer edge of the ROI remained at the same position. The reason for this behavior lies in the fact that the transition bands, from which the spurious signal arises, gets smaller in proportion with the ROI dimensions, since variable gradient strength defines ROI size in this implementation.

Outer volume ghosts

In three-pulse spatial selection schemes (PRESS, STEAM) crusher gradient pulses are inserted to prevent refocusing of unwanted echoes or FIDs (e.g. the single spin echo from a whole slice created by the 90° and one of the two 180° pulses in PRESS). Crusher gradient pulses work fine for homogeneous B_0. If appreciable local gradients are present at tissue interfaces, they can cancel the effect of some of the crusher gradients and lead to refocusing of unwanted echoes within the acquisition window. Refocusing is hardly ever at the expected echo maximum, and this leads to the typical appearance of outer volume ghosts, depicted in Fig. 14. Enhancing the crusher gradient amplitudes, or optimizing the directions in which they are played out,[41] can eliminate this artifact. Enhanced phase cycling can reduce these effects and, if individual scans are saved, can be used to identify these ghost signals.[42]

Pointspread function in spectroscopic imaging (SI)

This article is focused on single voxel MRS and coverage of potential artifacts and quality issues in SI is beyond its scope. Nevertheless it is necessary to point out the most drastic difference with regard to artifacts between single voxel MRS and SI. Classical SI is plagued neither with the

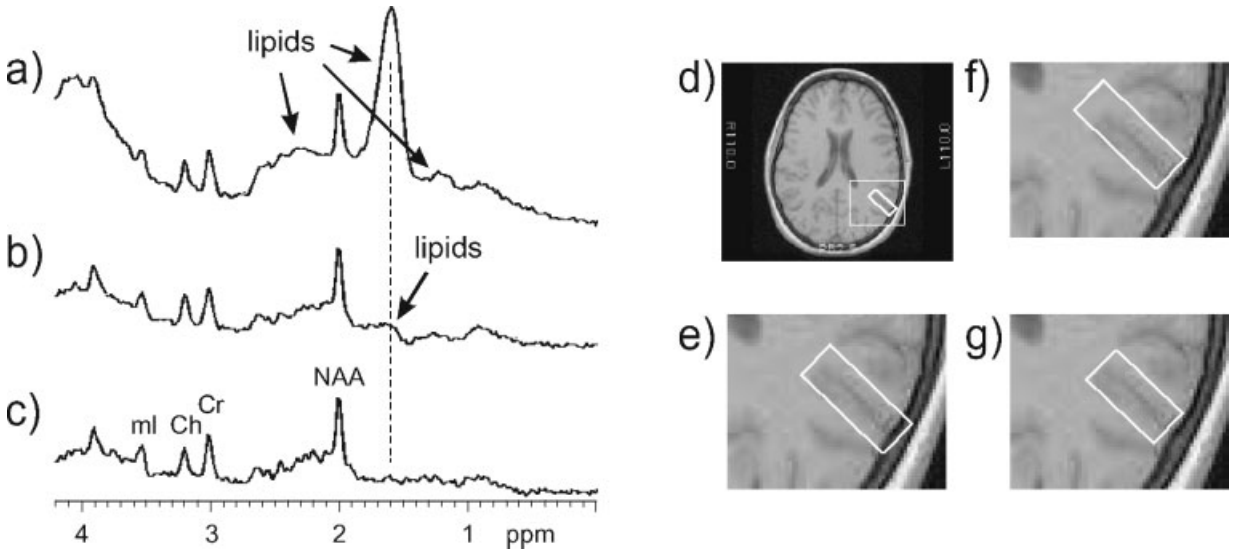

Figure 13. Signal bleed from outside the targeted ROI. Signal from outside the selected ROI can give dominating signal contributions, if the transition zone of the slice selective pulses falls into regions with large lipid content. This is illustrated for PRESS spectra obtained from a 40-year-old woman. The original ROI dimensions of $10 \times 15 \times 27$ mm, used for spectra (a) and (b) were reduced to $10 \times 15 \times 22$ mm for spectrum (c). This diminished the transition zone of the longest dimension of the voxel pointing towards the lipid-containing areas and the lipid contribution vanished. Just moving the ROI away from the skull, (e)–(f), did not completely eliminate the lipid contamination in the spectrum, (a)–(b). (Scan parameters: *TE* 20 ms, *TR* 3 s, 1953 Hz spectral width, 1024 points zero-filled to 2048 points, outer volume suppression pulses disabled)

chemical shift artifact nor with a non-ideal voxel profile, nor with outer volume ghosts (except if SI is combined with ROI preselection by STEAM or PRESS). However, unwanted signal from outside the ROI (i.e. a pixel) is still an inherent problem in SI. In SI—just like in MRI—signal from one pixel is always contaminated with signal from (all) other pixels. The true pixel size does not correspond to the ideal nominal resolution obtained by dividing field of view by number of phase encoding steps. The function that is responsible for the spatial mixing of signal is the 'point spread function' (PSF), which depends on the sampling and post-processing scheme. For techniques of SI and their optimization the reader is referred to the specialized literature.[7,30,35,43–47] The basic feature of signal blurring because of the PSF is illustrated in Fig. 15. Figure 15(a) shows the MRI of an experimental set up with two containers, where a large flask was filled with water only and a small vial within the larger container was filled with an acetate solution. The MR image is overlaid with a grid of spectra obtained from a classic SI scan, recorded with 32×32 phase resolution. The spectra are zoomed to the acetate peak and displayed in magnitude mode. (SI with STEAM prelocalization, *TE* 30 ms, *TR* 1 s, spectra displayed with 5-fold scaling, such that the full-scale acetate peak at the center of the vial has a size of five grid dimensions). The scaling demonstrates that there is also some acetate signal outside of the acetate-containing vial. The signal profile on the right (full line) can be compared with the expected signal for the given grid and vial size for an infinitely sharp PSF (dashed line). Figure 15(c) contains the corresponding metabolite image obtained from acetate peak areas. This contamination effect is fairly small for low-concentration metabolites, but is substantial, if the side band falls into a region with much stronger signals (e.g. lipid deposits[48]). *In vivo* there is often not enough time to attain the resolution needed to reach the spatial dimension of the suspected lesions. Figure 15(b) and (d) illustrates how Gibbs ringing, introduced by zero-filling of low-resolution SI data, exacerbates the problem of signal bleed. The same data as displayed in Fig. 15(a) and (d) were truncated to 8×32 phase steps and then zero-filled to obtain the original 32×32 resolution. The effect is drastic. Not only is the resolution four times worse (reflected in the spatial broadening seen in the signal profile and the signal spill-over around the central vial), but the PSF also leads to contamination of voxels far from the signal origin. In the current case, even voxels at the edge of the field of view show substantial acetate signal. Digital filtering or weighted acquisition schemes can be used to diminish this effect, but this always happens at the further expense of resolution.

Eddy currents

Uncompensated eddy currents lead to asymmetric lineshapes that are particularly striking for spectra with good resolution. Within certain limits, eddy current effects on lineshape can easily be corrected with the phase information from reference scans,[18,49] or by including a general lineshape function in the fitting model.[11,21] The former is illustrated in Fig. 16. On the down side, eddy current correction with inappropriate reference scans (recorded after a patient moved, or containing substantial lipid

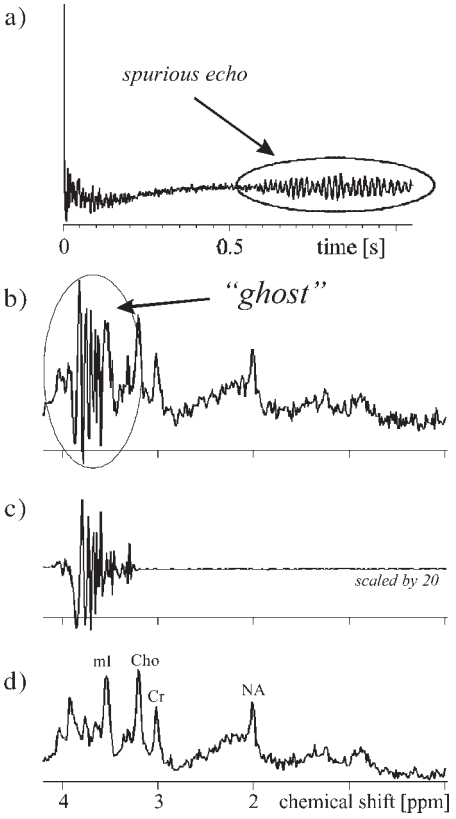

Figure 14. Effect of spurious echoes. Insufficient amplitude of gradient crusher pulses in combination with local B_0 inhomogeneities can lead to the refocusing of unwanted echoes (e.g. 2 pulse echo in a PRESS sequence). (a) The FID from a PRESS acquisition (*TE* 20 ms, *TR* 3 s) localizing developing white matter in a female preterm neonate (34 weeks gestational age). The encircled part of the FID originates from an unwanted echo. (b) The typical appearance of spurious echoes, often called ghosts, in the spectrum. Because extended phase cycling was used (phase rotation[42]) in data acquisition, the origin of the spurious signal could be identified in a separate trace after Fourier transformation along the phase rotation dimension (c). The particular phase evolution proved the spurious signal to arise from a two-pulse echo of the initial 90° and last 180° pulse. In the current case, elimination of the ghosting artifact can easily be accomplished by zeroing the latter half of the FID. The resulting spectrum is plotted in (d)

signal, as in MRS performed outside the brain) introduces substantial sidebands. Particularly at longer *TE*, eddy currents can also lead to signal drop in excess of the T_2 decay. These effects are location-dependent, if eddy currents occur in the gradient coils.

Automated fitting

Several fitting packages are available that all allow for automatic fitting of MR spectra using prior knowledge information from model spectra.[11,21,50–52] Using these programs, it may be tempting to concentrate on the numbers they produce (metabolite content, area ratios, error estimates) and not worry about visual inspection of the spectra and the fit. However, numbers can easily be treacherous. Suspicion concerning artifactual results will only arise if the results do not conform with expectation. It is strongly advisable to check the fitted spectrum as well as the shape of the fitted background signal on every occasion, and in particular to visualize the residuals (original spectrum minus fitted spectrum) for plausibility. Residuals should contain random white noise, otherwise the fitting model is not correct or the fitting procedure did not end in a global χ^2 minimum. Figures 17–19 all illustrate examples where inspection of residuals pinpoints artifactual results.

Figure 17 takes up the example from Fig. 14, where signal from a spurious echo led to a 'ghost' signal. The residuals make it easy to identify this ghost signal, since it did not conform to any of the model spectra. Although it appears that the fitting algorithm was able to separate out the ghost signal from the underlying metabolite spectrum (in particular mI), the resulting metabolite contents from metabolites with spectral components in the distorted region should be considered with caution, and in particular the error bound found for these metabolites may be wrong, because the spurious echo was not part of the fitting model. (In LC-Model, the estimated error bounds of all other metabolites will probably be too large, because noise is determined from the residuals and will therefore be overestimated.) The mean size of the residuals can be used for automated detection of artifacts, if the size of the minimum χ^2 is automatically compared with the SD for random noise.[21,53]

Inadequate fitting model

Unexpected features in the residuals can not only be due to artifacts, but also represent metabolites that were not included in the basis set of the fitting model. Figure 18 illustrates such a case. A short *TE* spectrum from supraventricular WM (PRESS, *TE* 20 ms) was apparently fitted properly except for the spectral region around 1.2 ppm, where obvious non-random signal contributions remain in the residuals. These contributions resemble a triplet with reasonably sized *J*-coupling. In this example, where the investigated subject was known to have been drinking alcoholic beverages before the scan, assignment of the triplet to the methyl part of ethanol is trivial, even though the quartet part of the ethanol model spectrum is not equally visible (reported to feature much wider lines *in vivo* than *in vitro*[54]).

Automatic assignments

Automatic or semiautomatic data processing by the scanner software or independent postprocessing usually leads to automatic peak assignment. This is necessary and useful, but should always be verified. Particularly in grossly abnormal or low SNR spectra, assignments can be wrong. Often, lipid signal is labeled as lactate or alanine. If the NAA signal is very small, automatic

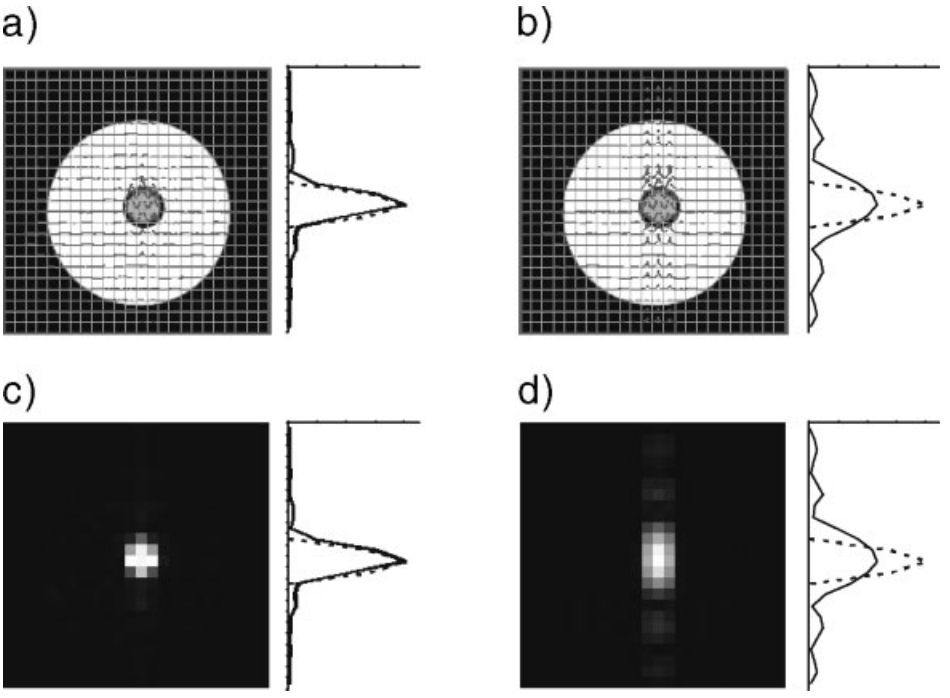

Figure 15. Point spread function in spectroscopic imaging. The 'point spread function' (PSF), which depends on the sampling and post processing scheme, shows how much signal is contributed to a pixel in SI from outside this pixel. The MRI of an experimental setup with two containers is shown in (a) and (b): the large flask is filled with water only and a small vial within the larger container is filled with an acetate solution. The MR image is overlaid with a grid of spectra obtained from a classic SI scan (32×32 phase resolution). The spectra of the acetate peak are overlaid in magnitude mode (SI with STEAM prelocalization, *TE* 30 ms, *TR* 1 s, spectra scaled 5-fold). The signal profile is displayed on the right (full line) and compared with the theoretical acetate distribution (infinitely sharp PSF, dashed line). (c) The acetate image obtained from acetate peak areas. Parts (b) and (d) illustrate how zero-filling of low-resolution SI data exacerbates the problem of signal bleeding: the original 32×32 resolution data was truncated to 8×32 phase steps and then zero-filled to 32×32. Substantial acetate signals are observed way outside the acetate container

assignments can lead to Cr being labeled as Ch or vice versa. The remedy in these cases is clear: 'Know (and check) your assignments!'[‖] If for some reason there is an appreciable frequency shift compared with normally acquired spectra (e.g. failure of center frequency algorithm, or movement between spectrum acquisition and acquisition of reference scans), fitting packages like LC-Model or TDFDfit may not converge to the true minimum and will then provide completely wrong results. This is illustrated in Fig. 19, where a frequency shift was not detected (or outside the correctable range) and all metabolite contents were totally wrong (0 mmol/kg NA, but 20 mmol/kg Gln!). Again, checking the fitting residuals (and baseline) clearly identifies such a failure in fitting. Assignment is particularly tricky if there are spectral peaks that do not correspond to previously known metabolite resonances. Identification can then turn into a research project of its own.[55–57] There are even cases where it turned out that *in vitro* spectra are different from *in vivo* spectra and likely assignment is even more difficult (chemical equilibrium of two isomers, or effect of medium;[58] dipolar coupling in ordered media[59–61]). If resonance peak position and/or linewidths do not agree with expectation from *in vitro* data, single observations should not be trusted and one should consider that artifacts are more likely than new metabolites. As an example, raging debates have been held about the interpretation of spectra with potential, but elusive contributions from silicone metabolites.[62] A further major difficulty arises if there is essentially just one interpretable peak in the spectrum [e.g. Ch in some brain tumor spectra, or in spectra of mamma carcinoma,[63–65] or phenylalanine (Phe) in difference spectra of PKU patients]. In these cases, prior knowledge on frequency and width of this single peak should stringently be enforced (e.g. vs water resonance) and multiple spectra should be recorded (best in repeat examinations or with differing acquisition parameters) in order to avoid interpretation of noise or artifact peaks.

[‖]Quoted from 'Rookie mistakes' assembled by Peter Barker, private communication.

Water suppression

Poor water suppression (WS) is not very problematic, provided that the shape of the residual water signal is well

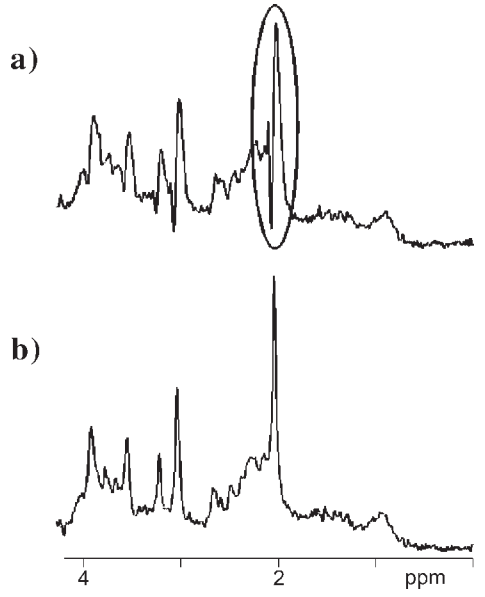

Figure 16. Effect of eddy currents. All lineshapes are distorted in spectrum (a) due to eddy currents in a short-*TE* PRESS spectrum (*TE* 20 ms) of occipital GM in a 14-year-old boy. (b) The same spectrum after restoration of the lineshape using the phase information from a water reference scan

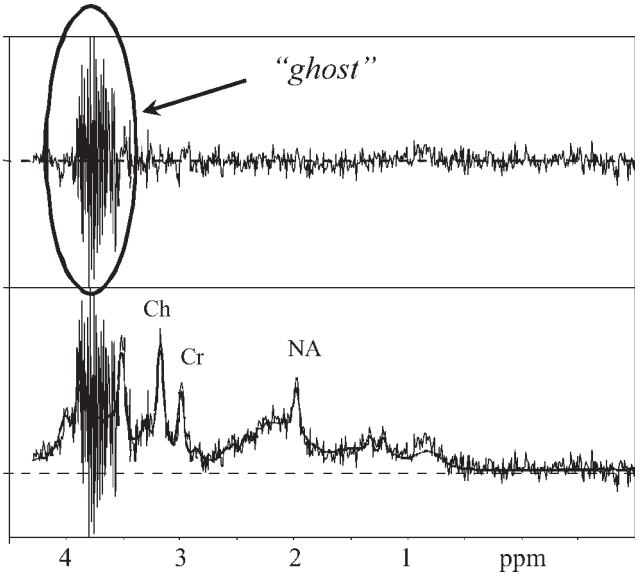

Figure 17. Use of fitting residuals to identify ghosting signals. A ghosting signal originating from a spurious echo can often be identified in the fitting residuals as a beating signal that cannot be due to random noise, nor due to a metabolite signal. The same spectrum as in Fig. 14 (neonate, thalamus) was fitted with LC-Model and the fit is plotted in the lower part of the figure overlying the experimental spectrum. The residuals (experimental–fitted spectrum) in the upper part clearly show the strong wiggles of a shifted echo at around 3.8 ppm

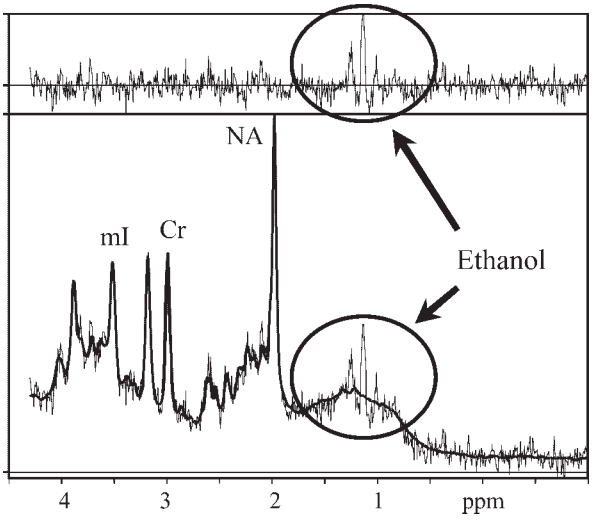

Figure 18. Use of fitting residuals to identify unexpected resonances. A spectrum of parietal WM from a 41-year-old male subject was fitted with LC-Model using a basis set consisting of the usually observed metabolites. The original spectrum and its fit are plotted in the lower part, the residuals after subtracting the fitted spectrum in the upper part of the figure. The latter clearly features signals that are not due to noise and can be used to pinpoint unexpected metabolites in patients. In the current example the triplet is due to ethanol consumed before the scan[54]

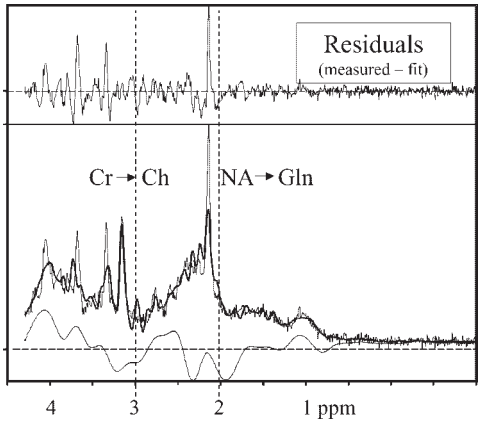

Figure 19. 'Know your assignments'. The importance of knowing and controlling peak assignment is illustrated for a case where a spectrum was shifted in frequency compared with its usual position and the fitting algorithm was not able to correct for the frequency shift. Instead, it converged to a wrong χ^2 minimum. In the given case, the spectral peaks were interpreted as wrong metabolites, yielding absurd quantitative metabolite contents, e.g. the Cr peak was interpreted as Ch and NA was quantified mainly as Gln (0.4 mM Cr instead of 6.2 mM; 20 mM Gln instead of 1 mM). If fitting residuals and baseline are checked for plausibility, such misassignments will not go unnoticed in near-normal spectra. However, in strongly abnormal spectra (e.g. tumor, where NA may be missing, and Cr may be low) correct assignment is more difficult to verify

behaved (smooth and not overwhelming) and can be accounted for in data fitting or preprocessing.[66] If so-called black-box filtering of residual water[67–70] is used in postprocessing, it should be verified that the subtracted signal represents only residual water. Dominating residual water peaks together with mechanical vibrations lead to sidebands at particular system-dependent offsets, that interfere with any method of peak area determination.[71,72] If basic peak integration routines, with or

without manual baseline correction, are used, the foot of the water resonance is often a major problem, in particular at short *TE* for determination of the mI peak area. The bandwidth of WS pulses should be kept constant and WS pulses should be applied at constant frequency offset with respect to the rest of the spectrum, because spectral features above 3.5 ppm may be partly suppressed, lying in the transition band of the WS pulses. In particular, mI determination might be influenced if one is not using identical WS (beware if comparing spectra from different scanners).

Sources of error in procedures of absolute quantitation

The most common procedures for calculating absolute concentrations[23,73] are (1) reference to the water signal as the internal standard, (2) use of a separately recorded reference signal from an external vial, and (3) signal calibration based on the principle of reciprocity.[74] All three methods can lead to large systematic errors and it is advisable to double check the results for consistency and plausibility.

- *Internal water reference*: the water content has to be assumed. This assumption is of course prone to error—particularly in unknown disease and within focal lesions. Compartmentation information from the *TE* decay of water[75] at least excludes error due to unexpectedly high water contributions of CSF or cysts. Gray and white matter differences in water content can be taken into account if segmentation information is acquired. Combining the method using the reciprocity principle with the water referencing technique allows a double check on both methods.
- *Reference signal from external standard*: this approach can easily be flawed by B_1 inhomogeneities between the two ROIs. If B_1 is calibrated for both locations, this uses up more time and depends on the correctness of both calibrations. Double checking is possible, if this technique is combined with the internal water reference and a water content for the ROI is calculated and checked for plausibility.
- *Calibration based on the reciprocity principle*[74,76]: this method is very sensitive to any changes in hardware (receive as well as transmit path, see below). Regular hardware checks are a must. Phantom measurements to verify calibration are recommended to accompany each patient measurement. Amplifier linearity and receive amplification should be verified regularly. As mentioned above, determination of water content using the non-suppressed water signal is a good means of checking for plausibility of the calibration.

Furthermore, the use of T_1- and T_2-insensitive localization sequences (i.e. short *TE*, long *TR*) is recommended in order not to rely on unchanged relaxation characteristics in disease.

Hardware problems

Often, MRS will be the first technique to be affected by failing scanner hardware, as spectral quality and reproducibility depend strongly on optimal and constant hardware (B_0 homogeneity, eddy currents, RF homogeneity and stability, stability of receive channels, amplifier linearity, external noise sources).

- *External source of RF*: an external source of RF leaking into the receive paths may be difficult to recognize in MRS, if the frequency coincides with frequency offsets of known metabolite peaks. Two examples are given in Fig. 20. In the first example (left-side of Fig. 20) the interfering RF happened to appear at the position of acetate and to be 30° out of phase with the spectrum. This can be detected in the minimally-treated spectrum [Fig. 20(c)], but not really in the smoothed spectrum Fig. 20(b). In the second example (right-hand side of Fig. 20), the leaking RF overlaps with NAA. Most often the phase of an external RF signal is not locked to the scanner phase and can have any relation. In this example it was assumed that the external signal happened to be approximately 180° out of phase and would therefore diminish the NAA peak. The example shows that, again, a minimally treated spectrum would clearly show an artifact at the position of NAA, as the external RF is not decaying with a T_2, but would have a very narrow width. If only line-broadened spectra are inspected visually, the artifact can easily be missed and the spectrum would suggest neurodegeneration. If an external RF signal is suspected during signal acquisition (e.g. because a strong signal becomes weaker with signal averaging), it can be easily verified: it would persist when the localization pulses were switched off. If the suspicion arises later during inspection of the spectrum, it would be helpful to have stored individual scans and not just the sum, because adding the magnitude spectra of all single excitations would probably drastically increase the artifactual signal (no phase cancellation).
- *Analog-to-digital conversion (ADC)*: hardware or software errors can occur in the signal acquisition chain. One potential problem occurs if analog-to-digital conversion is incorrect. Problems with a defective bit in analog-to-digital conversion are illustrated in Figs 21 and 22. For illustration purposes it was assumed that one bit in the ADC is faulty, i.e. always zero, irrespective of the detected analog RF input. Other technical defects in ADC are possible, and they cause different effects on the spectra. Figure 21 shows an example for a synthetic single line spectrum, where an intermediate order bit is defective. FID (top) and the corresponding spectrum (bottom) on the left are from a properly functioning system, while the graphs on the right illustrate the effect of a faulty bit on the TD and FD signals. In the usually visualized FD spectrum, the

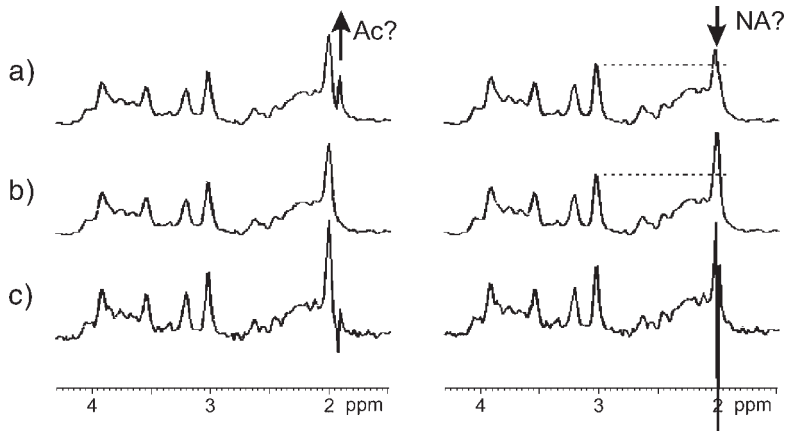

Figure 20. Spurious RF signals. The effect of an RF leak is demonstrated in a PRESS spectrum of a 9-year-old boy (perisylvian GM). The spectrum was modified by adding a constant external RF signal to its FID. (b) The original spectrum (2 Hz exponential apodization on the left, 2 Hz Gauss apodization on the right). In spectrum (a) on the left-hand side, the external RF signal could easily be mistaken for a peak of acetate at 1.93 ppm. The unapodized spectrum in (c) shows that the added signal is out of phase with the rest of the spectrum and is narrower. Spectrum (a) on the right-hand side could erroneously be read as proof for an NA deficit, since the external RF signal happens to coincide with the NA peak and have negative phase. The asymmetric lineshape for the NA peak in the smoothed spectrum might lead an expert to inspect the unapodized spectrum (c), which would clearly reveal an artifactual RF signal as the cause for the apparently low NA area

peak appears smaller, with unusual lineshape and additional sideband peaks, which in real data would not be conspicuous because of signal overlap and limited SNR. Figure 22 contains another potential effect originating from the same scanner defect, but this time affecting a high-order bit (i.e. only affecting large numbers in the FID, at and right after the signal maximum at time 0). Figure 22(a) contains the first part of an FID from an *in vivo* brain spectrum showing the cause of the artifact in TD. Figure 22(b) shows that visual inspection of the zoomed spectral region of interest does not give an obvious hint that the FID is faulty, except that the experienced eye may detect a somewhat unusual baseline or signal phase. If signal-free regions of the spectrum are inspected, the artifact is evident in the baseline roll.

- *RF linearity*: RF amplifiers should have linear output characteristics. If not, modulated pulses are distorted, leading to unexpected spatial excitation profiles.[77] Similarly, water suppression pulses, which are usually played out at much lower amplitudes, can also be affected severely, leading to saturation outside the expected bandwidth. Figure 23 illustrates another consequence of nonlinearity in RF output. The transmit gain, expressed as a voltage and adjusted by auto-prescan or the operator, normally translates linearly into the output voltage after power amplification (dashed diagonal line). Hence, the transmit gain is taken as a relevant measure of the excitation current needed to apply the reciprocity principle (see above).[74] In this real-world example, power amplification was defective (measured amplification plotted by diamond symbols), leading to a jump in output power upon a small increase in transmit gain. In the illustrated case, where normally 4.5 V input voltage would be needed to obtain the 90° pulse power, only 3.2 V are needed for this power because of the nonlinear amplification characteristics. If the reciprocity principle is used for

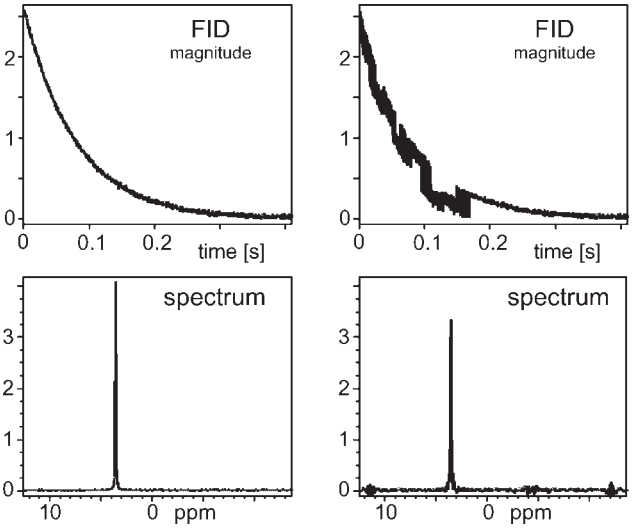

Figure 21. Data acquisition error due to a faulty analog-to-digital converter. Theoretical data were simulated to illustrate a faulty bit in ADC that is zero irrespective of the analog signal to be converted. The synthetic spectrum was created with 1024 points, a spectral width of 2500 Hz and a single peak 250 Hz off resonance. As evident from the faulty FID on the right in comparison to the properly digitized FID on the left, a medium order bit that affected almost the full length of the FID was assumed to be defective. In FD this leads to an unusual lineshape and sideband peaks

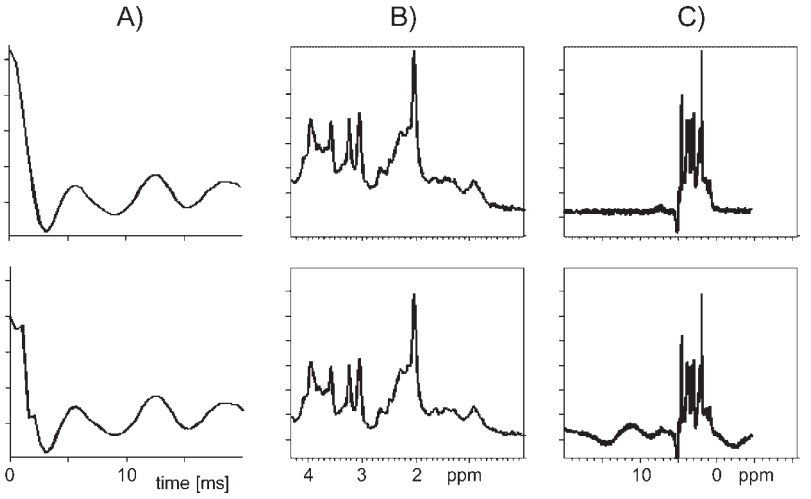

Figure 22. Data acquisition error due to a faulty analog to digital converter. The same defect as in Fig. 21 can lead to quite different consequences if a higher order bit is involved that only affects the first few data points of the FID, this time illustrated on an experimental spectrum (short-*TE* PRESS, *TE* 20 ms, *TR* 3 s, 1953 Hz spectral width, 2048 data points) from an ROI in frontal GM of a young healthy woman. The faulty bit was emulated by signal processing. The top row contains the undistorted data. (A) The distortion of the FID (lower trace). (B) and (C) Differently zoomed regions of the corresponding spectra. The close up of the region of usual interest does not show the artifact clearly, but quantitation would of course be affected by the baseline role, which is clearly visible in (C)

quantitation, this would be interpreted as a less-loaded RF coil (need for little RF input) that will produce a high output signal for a given spin response. Metabolite concentrations would be underestimated by 30%. The hardware defect turned out to be caused by a defective digital input signal in a preamplification stage.

- *Quality Assurance*: to detect any defects in the excitation and receive paths, regular quality control[3,17,78,79] with phantom and volunteer scans is highly recommended. Applying the reciprocity principle to quantitate solution spectra, which are known to be constant, is very useful in this context. Figure 24 shows a series of such spectra, recorded over 7 months. The solution spectra were fitted with LC-Model and absolute metabolite content was determined with two methods: first, using the water signal as internal reference, and second, using the reciprocity principle assuming constant hardware. Clearly, the hardware was not constant. Replacement of the RF coil and also problems with amplifier linearity affected the calculated Cr concentration severely, while a deterioration of image quality in clinical MRI went unnoticed. The combination of the two methods allows the measurements to be used as calibration scheme for clinical exams performed with the corresponding hardware setups.

GENERAL CONSIDERATIONS

Criteria for quality checks

To conclude, potential guidelines that allow to judge the quality of individual data sets are listed in quality checks below:

Quality checks in clinical MRS.

- Verify sufficient SNR.
- Verify sufficient spectral resolution.
- Respect minimum error bounds from data fitting.

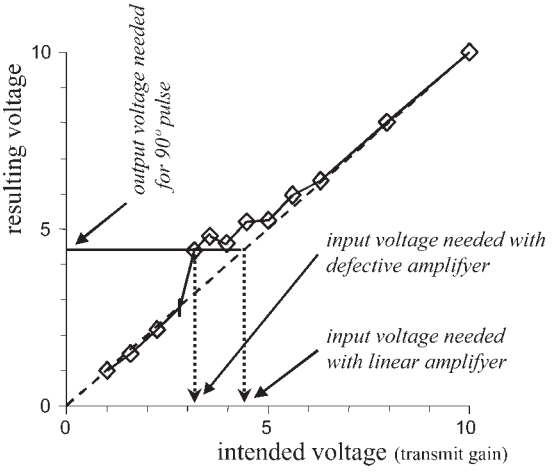

Figure 23. Effect of nonlinear power amplification. The transmit gain, expressed as a voltage and adjusted by auto-prescan or the operator, normally translates linearly into the output voltage after power amplification (dashed diagonal line). Hence, the transmit gain is taken as relevant measure for the excitation current needed to apply the reciprocity principle. In this example power amplification was defective (measured amplification plotted by diamond symbols) leading to a jump in output power upon a small increase in transmit gain. In the illustrated case, where normally 4.5 V would be needed as input, only 3.2 V are needed for this power because of nonlinear amplification. If the reciprocity principle is used for quantitation, metabolite concentrations would be underestimated by 30%

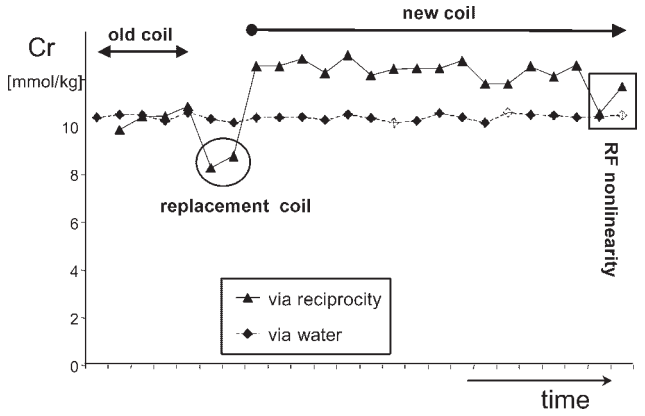

Figure 24. Regular MRS examinations on a QA phantom to detect hardware failure. The Cr content of a phantom (cf. Fig. 10) as determined by two methods is plotted against time (full range of 7 months). Reference method 1 (diamonds) used the water signal as internal reference. As expected for a constant solution, these values show very little fluctuation. Values obtained through application of the reciprocity principle (triangles) feature larger fluctuation and are indicative of hardware failure or changes in hardware. If the reciprocity principle is used to obtain quantitative values from clinical examinations, these phantom measurements are indispensable for calibration after changes in hardware and as seismographs to detect incipient hardware failure

- Verify ROI location:
  - scout images pre- and post-MRS;
  - header information in spectra;
  - ROI images.
- Check minimally processed TD signal to detect:
  - shifted echoes;
  - spurious echoes;
  - faulty bits in or wrong scaling for ADC.
- Check minimally processed spectrum to detect:
  - peaks with abnormal phase;
  - peaks that are narrower or with strange lineshape;
  - peak doubling;
  - failure of WS;
  - abnormal eddy currents;
  - need for first order phasing.
- Check fitting residuals (and baseline) for:
  - ghosts;
  - unidentified metabolites;
  - outer volume signal;
  - wrong assignments;
  - SD of residuals vs SD of noise.
- Check assignments.
- Verify unassigned resonances vs artifacts:
  - exact chemical shift;
  - phase and shape.
- Store individual FIDs to be able to check and possibly correct individual data traces in case of system instabilities or patient motion.
- Acquire reference scans (water) for:
  - phasing of spectrum;
  - eddy correction of spectrum;
  - determination of compartment sizes.[49]
- Calculate water content from reference scans as plausibility check.
- Check plausibility of metabolite contents (overall increase/deficit, water content)
- Regular QA on phantoms to:
  - detect hardware problems [B_0 homogeneity, eddy currents, RF amplifier characteristics, gradient amplifiers, various noise sources, software up- (and down-) grades];
  - update calibration measurements for absolute quantitation.

Criteria for rejection of data

Error estimates from CRMVB, confidence limits,[53] minimum least squares sum (χ^2), and/or general reproducibility at the local site should be considered before defining abnormal values (see below). Confidence images[30,53] and/or rejection masks[16] should be constructed for SI data. Based on above statements, the literature and an opinion poll among colleagues, criteria for rejection of spectra or individual metabolite values can be formulated:
Reject data if:

- FWHM of metabolites > 0.07–0.1 ppm;
- CRMVB $> 50\%$;
- unexplained features in residuals—reject, if artifact or expand model, if unexpected metabolite;
- peaks doubled or patient moved (post-acquisition MRI);
- lineshape strongly asymmetric after eddy correction;
- outer volume ghosts or other artifacts present (at least exclude metabolites overlaid with artifact).

Criteria to define abnormality

A single metabolite level in a single spectrum from a single subject can be considered abnormal if it lies outside the normal range defined by the mean ± 2 SD of the control cohort (SD from variance over measurements in a control group). Additionally, SNR and FWHM should also fall within normal limits. The control values must originate from truly comparable exams (ROI size, acquisition parameters, location, subject age). Given the reproducibility found in the literature, deviations from the norm must usually be at least 15%, possibly much more, depending on which metabolite is considered, the local reproducibility values and ROI size. Smaller changes may be confirmed in repeated studies, but should still be judged against incidental individuality (σ_{inter}).[27] In SI studies, detection of abnormalities can be based on similar principles, but SI also offers additional options. One possibility is to compare intraindividually with the contralateral side, others are to compare correlations with

tissue content,[80] combinations of metabolite levels from different brain regions or whole brain measures (Andrew Maudsley, private communication).

Acknowledgements

Supported by the Swiss National Foundation (31-059082). Particular thanks go to all spectroscopists near and far, who helped me with this manuscript or shared with me their views on quality issues in MRS; in particular to Drs Peter Barker, Chris Boesch, Patrick Bolan, Else Danielsen, Arend Heerschap, Franklyn Howe, Ralph Hurd, Michael Ith, Andrew Maudsley, Mary McLean, Petra Pouwels, Napapon Sailasuta, Stephen Provencher, Hans Slotboom and Peter Vermathen.

REFERENCES

1. Kubelka J, Burian M, Hajek M. Quality control and routine stability check for single voxel 1H *in vivo* quantitative magnetic resonance spectroscopy. *J. Magn. Reson. Anal.* 1997; **3**: 93–98.
2. Podo F, Henriksen O, Bovée WMMJ, Leach MO, Leibfritz D, de Certaines JD. Absolute metabolite quantification by *in vivo* NMR spectroscopy: I. Introduction, objectives and activities of a concerted action in biomedical research. *Magn. Reson. Imag.* 1998; **16**: 1085–1092.
3. Hajek M, Burian M, Dezortova M. Application of LCModel for quality control and quantitative *in vivo* 1H MR spectroscopy by short echo time STEAM sequence. *MAGMA* 2000; **10**: 6–17.
4. Griffiths JR, Tate AR, Howe FA, Stubbs M. Magnetic resonance spectroscopy of cancer–practicalities of multi-centre trials and early results in non-Hodgkin's lymphoma. *Eur. J. Cancer* 2002; **38**: 2085–2093.
5. de Certaines JD, Cathelineau G. Safety aspects and quality assessment in MRI and MRS: a challenge for health care systems in Europe. *J. Magn. Reson. Imag.* 2001; **13**: 632–638.
6. van der Knaap MS, Barth PG, Gabreels FJM, Franzoni E, Begeer JH, Stroink H, Rotteveel JJ, Valk J. A new leukoencephalopathy with vanishing white matter. *Neurology* 1997; **48**: 845–855.
7. De Graaf RA. *In vivo NMR Spectroscopy: Principles and Techniques*. Wiley: Chichester, 1999.
8. in't Zandt HJA, van der Graaf M, Heerschap A. Common processing of *in vivo* MR spectra. *NMR Biomed.* 2001; **14**: 224–232.
9. Vanhamme L, Sundin T, Hecke PV, Huffel SV. MR spectroscopy quantitation: a review of time-domain methods. *NMR Biomed.* 2001; **14**: 233–246.
10. Mierisova S, Ala-Korpela M. MR spectroscopy quantitation: a review of frequency domain methods. *NMR Biomed.* 2001; **14**: 247–259.
11. Provencher SW. Estimation of metabolite concentration from localized *in vivo* proton NMR spectra. *Magn. Reson. Med.* 1993; **30**: 672–679.
12. Cavassila S, Deval S, Huegen C, van Ormondt D, Graveron-Demilly D. Cramer–Rao bounds: an evaluation tool for quantitation. *NMR Biomed.* 2001; **14**: 278–283.
13. Tkac I, Ugurbil K, Gruetter R. On the quantification of low concentration metabolites by ^{1}H NMR spectroscopy in the human brain at 7 Tesla. In *10th Meeting of the International Society of Magnetic Resonance in Medicine*, Honolulu, 2002; 528.
14. Kreis R, Boesch C. Bad spectra can be better than good spectra. In *11th Meeting of the International Society of Magnetic Resonance in Medicine*, Toronto, 2003; 264.
15. Kreis R, Hofmann L, Kuhlmann B, Boesch C, Bossi E, Hüppi PS. Brain metabolite composition during early human brain development as measured by quantitative *in vivo* ^{1}H magnetic resonance spectroscopy. *Magn. Reson. Med.* 2002; **48**: 949–958.
16. Ebel A, Soher BJ, Maudsley AA. Assessment of 3D proton MR echo-planar spectroscopic imaging using automated spectral analysis. *Magn. Reson. Med.* 2001; **46**: 1072–1078.
17. Schirmer T, Auer DP. On the reliability of quantitative clinical magnetic resonance spectroscopy of the human brain. *NMR Biomed.* 2000; **13**: 28–36.
18. de Graaf AA, van Dijk JE, Bovee WMMJ. QUALITY: quantification improvement by converting lineshapes to the Lorentzian type. *Magn. Reson. Med.* 1990; **13**: 343–357.
19. Webb P, Spielman D, Macovski A. Inhomogeneity correction for *in vivo* spectroscopy by high-resolution water referencing. *Magn. Reson. Med.* 1992; **23**: 1–11.
20. Bartha R, Drost DJ, Menon RS, Williamson PC. Spectroscopic lineshape correction by QUECC: combined QUALITY deconvolution and eddy current correction. *Magn. Reson. Med.* 2000; **44**: 641–645.
21. Slotboom J, Boesch C, Kreis R. Versatile frequency domain fitting using time domain models and prior knowledge. *Magn. Reson. Med.* 1998; **39**: 899–911.
22. Mullins PG, Rowland L, Bustillo J, Bedrick EJ, Lauriello J, Brooks WM. Reproducibility of 1H-MRS measurements in schizophrenic patients. *Magn. Reson. Med.* 2003; **50**: 704–707.
23. Kreis R. Quantitative localized ^{1}H-MR spectroscopy for clinical use. *Prog. NMR Spectrosc.* 1997; **31**: 155–195.
24. Bartha R, Drost DJ, Menon RS, Williamson PC. Comparison of the quantification precision of human short echo time (1)H spectroscopy at 1.5 and 4.0 Tesla. *Magn. Reson. Med.* 2000; **44**: 185–192.
25. Marshall I, Wardlaw J, Cannon J, Slattery J, Sellar RJ. Reproducibility of metabolite peak areas in ^{1}H MRS of brain. *Magn. Reson. Imag.* 1996; **14**: 281–292.
26. Brooks WM, Friedman SD, Stidley CA. Reproducibility of ^{1}H-MRS *in vivo*. *Magn. Reson. Med.* 1999; **41**: 193–197.
27. Kreis R, Fusch C, Maloca P, Felblinger J, Boesch C. Supposed pathology may be individuality: interindividual and regional differences of brain metabolite concentrations determined by ^{1}H MRS. In *2nd Meeting of the Society of Magnetic Resonance*, San Francisco, 1994; 45.
28. Hofmann L, Slotboom J, Jung B, Maloca P, Boesch C, Kreis R. Quantitative 1H-magnetic resonance spectroscopy of human brain: influence of composition and parameterization of the basis set in linear combination model fitting. *Magn. Reson. Med.* 2002; **48**: 440–453.
29. Li BS, Babb JS, Soher BJ, Maudsley AA, Gonen O. Reproducibility of 3D proton spectroscopy in the human brain. *Magn. Reson. Med.* 2002; **47**: 439–446.
30. McLean MA, Woermann FG, Barker GJ, Duncan JS. Quantitative analysis of short echo time (1)H-MRSI of cerebral gray and white matter. *Magn. Reson. Med.* 2000; **44**: 401–411.
31. Wiedermann D, Schuff N, Matson GB, Soher BJ, Du AT, Maudsley AA, Weiner MW. Short echo time multislice proton magnetic resonance spectroscopic imaging in human brain: metabolite distributions and reliability. *Magn. Reson. Imag.* 2001; **19**: 1073–1080.
32. Chard DT, McLean MA, Parker GJ, MacManus DG, Miller DH. Reproducibility of *in vivo* metabolite quantification with proton magnetic resonance spectroscopic imaging. *J. Magn. Reson. Imag.* 2002; **15**: 219–225.
33. Felblinger J, Kreis R, Boesch C. Effects of physiologic motion of the brain upon quantitative ^{1}H-MRS: analysis and correction by retro-gating. *NMR Biomed.* 1998; **11**: 107–114.
34. Haupt CI, Kiefer AP, Maudsley AA. In-plane motion correction for MR spectroscopic imaging. *Magn. Reson. Med.* 1998; **39**: 749–753.
35. Kim DH, Adalsteinsson E, Spielman DM. Spiral readout gradients for the reduction of motion artifacts in chemical shift imaging. *Magn. Reson. Med.* 2004; **51**: 458–463.
36. Ryner LN, Ke Y, Thomas MA. Flip angle effects in STEAM and PRESS-optimized versus sinc RF pulses. *J. Magn. Reson.* 1998; **131**: 118–125.
37. Keevil SF, Newbold MC. The performance of volume selection sequences for *in vivo* NMR spectroscopy: implications for quantitative MRS. *Magn. Reson. Imag.* 2001; **19**: 1217–1226.
38. Pauly J, Le Roux P, Nishimura D, Macovski A. Parameter relations for the Shinnar-Le Roux selective excitation pulse design algorithm. *IEEE Trans. Med. Imag.* 1991; **10**: 53–65.

39. Tran TK, Vigneron DB, Sailasuta N, Tropp J, Le Roux P, Kurhanewicz J, Nelson S, Hurd R. Very selective suppression pulses for clinical MRSI studies of brain and prostate cancer. *Magn. Reson. Med.* 2000; **43**: 23–33.
40. 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**: 424–437.
41. Moonen CTW, van Zijl PCM. Highly effective water suppression for *in vivo* proton NMR spectroscopy (DRYSTEAM). *J. Magn. Reson.* 1990; **88**: 28–41.
42. Hennig J. The application of phase rotation for localized *in vivo* proton spectroscopy with short echo times. *J. Magn. Reson.* 1992; **96**: 40–49.
43. Nelson SJ, Vigneron DB, Star-Lack J, Kurhanewicz J. High spatial resolution and speed in MRSI. *NMR Biomed.* 1997; **10**: 411–422.
44. Hugg JW, Maudsley AA, Weiner MW, Matson GB. Comparison of k-space sampling schemes for multidimensional MR spectroscopic imaging. *Magn. Reson. Med.* 1996; **36**: 469–473.
45. Pohmann R, von Kienlin M, Haase A. Theoretical evaluation and comparison of fast chemical shift imaging methods. *J. Magn. Reson.* 1997; **129**: 145–160.
46. Ebel A, Maudsley AA. Improved spectral quality for 3D MR spectroscopic imaging using a high spatial resolution acquisition strategy. *Magn. Reson. Imag.* 2003; **21**: 113–120.
47. Soher BJ, Vermathen P, Schuff N, Wiedermann D, Meyerhoff DJ, Weiner MW, Maudsley AA. Short TE *in vivo* (1)H MR spectroscopic imaging at 1.5 T: acquisition and automated spectral analysis. *Magn. Reson. Imag.* 2000; **18**: 1159–1165.
48. Haupt CI, Schuff N, Weiner MW, Maudsley AA. Removal of lipid artifacts in 1H spectroscopic imaging by data extrapolation. *Magn. Reson. Med.* 1996; **35**: 678–687.
49. Ordidge RJ, Cresshull ID. The correction of transient B0 field shifts following the application of pulsed gradients by phase correction in the time domain. *J. Magn. Reson.* 1986; **69**: 151–155.
50. Elster C, Link A, Schubert F, Seifert F, Walzel M, Rinneberg H. Quantitative MRS: comparison of time domain and time domain frequency domain methods using a novel test procedure. *Magn. Reson. Imag.* 2000; **18**: 597–606.
51. Soher BJ, Young K, Govindaraju V, Maudsley AA. Automated spectral analysis III: application to *in vivo* proton MR spectroscopy and spectroscopic imaging. *Magn. Reson. Med.* 1998; **40**: 822–831.
52. Mierisova S, van den Boogaart A, Tkac I, Van Hecke P, Vanhamme L, Liptaj T. New approach for quantitation of short echo time *in vivo* ^{1}H MR spectra of brain using AMARES. *NMR Biomed.* 1998; **11**: 32–39.
53. Young K, Khetselius D, Soher BJ, Maudsley AA. Confidence images for MR spectroscopic imaging. *Magn. Reson. Med.* 2000; **44**: 537–545.
54. Kreis R, Hofmann L, Boesch C. Visibility of the methylene signals of ethanol in ^{1}H-MR spectra of the human brain and implications for model fitting. In *8th Meeting of the International Society of Magnetic Resonance in Medicine*, Denver, 2000; 426.
55. Michaelis T, Helms G, Merboldt K-D, Haenicke W, Bruhn H, Frahm J. Identification of *scyllo*-inositol in proton NMR spectra of human brain *in vivo*. *NMR Biomed.* 1993; **6**: 105–109.
56. Cady EB, Lorek A, Penrice J, Reynolds EO, Iles RA, Burns SP, Coutts GA, Cowan FM. Detection of propan-1,2-diol in neonatal brain by *in vivo* proton magnetic resonance spectroscopy. *Magn. Reson. Med.* 1994; **32**: 764–767.
57. Ith M, Bigler P, Scheurer E, Kreis R, Hofmann L, Dirnhofer R, Boesch C. Observation and identification of metabolites emerging during postmortem decomposition of brain tissue by means of *in situ* ^{1}H-magnetic resonance spectroscopy. *Magn. Reson. Med.* 2002; **48**: 915–920.
58. Kreis R, Ross BD. Cerebral metabolic disturbances in patients with subacute and chronic diabetes mellitus: detection with proton MR spectroscopy. *Radiology* 1992; **184**: 123–130.
59. Kreis R, Boesch C. Spatially localized, one- and two-dimensional NMR spectroscopy and *in vivo* application to human muscle. *J. Magn. Reson. Ser. B* 1996; **113**: 103–118.
60. Kreis R, Koster M, Kamber M, Hoppeler H, Boesch C. Peak assignment in localized ^{1}H MR spectra based on oral creatine supplementation. *Magn. Reson. Med.* 1997; **37**: 159–163.
61. Boesch C, Kreis R. Dipolar coupling and ordering effects observed in MR spectra of skeletal muscle. *NMR Biomed.* 2001; **14**: 140–148.
62. Hull WE. A critical review of MR studies concerning silicone breast implants. *Magn. Reson. Med.* 1999; **42**: 984–995; discussion 995–996.
63. Yeung DK, Cheung HS, Tse GM. Human breast lesions: characterization with contrast-enhanced *in vivo* proton MR spectroscopy—initial results. *Radiology* 2001; **220**: 40–46.
64. Katz-Brull R, Lavin PT, Lenkinski RE. Clinical utility of proton magnetic resonance spectroscopy in characterizing breast lesions. *J. Natl Cancer Inst.* 2002; **94**: 1197–1203.
65. Bolan PJ, DelaBarre L, Baker EH, Merkle H, Everson LI, Yee D, Garwood M. Eliminating spurious lipid sidebands in 1H MRS of breast lesions. *Magn. Reson. Med.* 2002; **48**: 215–222.
66. Coron A, Vanhamme L, Antoine JP, Van Hecke P, van Huffel S. The filtering approach to solvent peak suppression in MRS: a critical review. *J. Magn. Reson.* 2001; **152**: 26–40.
67. Marion D, Ikura M, Bax A. Improved solvent suppression in one- and two-dimensional NMR spectra by convolution of time-domain data. *J. Magn. Reson.* 1989; **84**: 425–430.
68. Sundin T, Vanhamme L, Van Hecke P, Dologlou I, van Huffel S. Accurate quantification of (1)H spectra: From finite impulse response filter design for solvent suppression to parameter estimation. *J. Magn. Reson.* 1999; **139**: 189–204.
69. Laudadio T, Mastronardi N, Vanhamme L, Van Hecke P, van Huffel S. Improved Lanczos algorithms for blackbox MRS data quantitation. *J. Magn. Reson.* 2002; **157**: 292–297.
70. Cabanes E, Confort-Gouny S, Le Fur Y, Simond G, Cozzone PJ. Optimization of residual water signal removal by HLSVD on simulated short echo time proton MR spectra of the human brain. *J. Magn. Reson.* 2001; **150**: 116–125.
71. Kreis R, Boesch C. Localized ^{1}H-MRS without water saturation: techniques and initial results for human brain and muscle. In *6th Meeting of the International Society of Magnetic Resonance in Medicine*, Sydney, 1998; 24.
72. Hurd RE, Gurr D, Sailasuta N. Proton spectroscopy without water suppression: the oversampled *J*-resolved experiment. *Magn. Reson. Med.* 1998; **40**: 343–347.
73. Tofts P, Waldman AD. Spectroscopy: ^{1}H metabolite concentrations. In *Quantitative MRI of the Brain: Measuring Changes Caused by Disease*, Tofts P (ed.). Wiley: New York, 2003; 299–340.
74. Hoult DI. The principle of reciprocity in signal strength calculations—A mathematical guide. *Concepts Magn. Reson.* 2000; **12**: 173–187.
75. Ernst T, Kreis R, Ross BD. Absolute quantitation of water and metabolites in the human brain. I. Compartments and water. *J. Magn. Reson. Ser. B* 1993; **102**: 1–8.
76. Michaelis T, Merboldt KD, Bruhn H, Hänicke W, Frahm J. Absolute concentrations of metabolites in the adult human brain *in vivo*: quantification of localized proton MR spectra. *Radiology* 1993; **187**: 219–227.
77. Chan F, Pauly J, Macovski A. Effects of RF amplifier distortion on selective excitation and their correction by prewarping. *Magn. Reson. Med.* 1992; **23**: 224–238.
78. Simmons A, Smail M, Moore E, Williams SCR. Serial precision of metabolite peak area ratios and water referenced metabolite peak areas in proton MR spectroscopy of the human brain. *Magn. Reson. Imag.* 1998; **16**: 319–330.
79. Keevil SF, Barbiroli B, Brooks JCW, Cady EB, Canese R, Carlier P, Collins DJ, Gilligan P, Gobbi G, Hennig J, Kugel H, Leach MO, Metzler D, Mlynarik V, Moser E, Newbold MC, Payne GS, Ring P, Roberts JN, Rowland IJ, Thiel T, Tkac I, Topp S, Wittsack HJ, Wylezinska M, Zaniol P, Henriksen O, Podo F. 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. Imag.* 1998; **16**: 1093–1106.
80. Hetherington HP, Pan JW, Mason GF, Adams D, Vaughn MJ, Twieg DB, Pohost GM. Quantitative ^{1}H spectroscopic imaging of human brain at 4.1 T using image segmentation. *Magn. Reson. Med.* 1996; **36**: 21–29.
81. Naressi A, Couturier C, Devos JM, Janssen M, Mangeat C, de Beer R, Graveron-Demilly D. Java-based graphical user interface for the MRUI quantitation package. *MAGMA* 2001; **12**: 141–152.

