Resampling strategies to estimate mean concentrations from low SNR in vivo MR spectra

G. Helms[1], I. Savic[2], and P. Dechent[1]

[1]MR-Research in Neurology and Psychiatry, University of Göttingen, Faculty of Medicine, Göttingen, Germany, [2]MR-Centrum, Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden

Introduction:

Summing of in vivo MR spectra reduces noise and spurious artefacts, thus improving the definition of metabolite resonances and baseline. Contrary to the average of all spectra, the variation between different subsets yields also the standard deviation (SD). The influence of subset size was systematically evaluated and compared with basic bootstrap resampling (1).

Methods:

Ten spectra were acquired with STEAM (TE/TM/TR/AQ=30/14/6000/64) on a 1.5 T GE Signa in thalamus (mean volume = 3.6 ml, SNR = 3.8) of healthy controls. Each acquisition was aligned in phase and frequency by means of the residual water signal before averaging (2). Molar concentrations were estimated using LCModel 5.2 (3) with a scaling (TRAMP) accounting for calibration, coil loading and inhomogeneity, and CSF partial volume (4). In order to estimate the average concentrations from a subset of p subjects, the single spectra were multiplied by TRAMP prior to summation. The sum was evaluated with TRAMP=1 and the total volume. The SD may be calculated from SD(p) of all p out of n=10 possible subsets:

$$\mathbf{SD}(p) = \mathbf{SD}\ \mathbf{sqrt}\ [(n-p)/pn] = \mathbf{SEM}\ \mathbf{sqrt}\ [(n-p)/p]. \quad [1]$$

This so-called finite population correction accounts for both the number and the redundancy of the subgroups when compared to re-sampling with replacement (bootstrap). 1000 random bootstrap resamples of 10 draws with replacement yielded

$$\mathbf{SD(bs)} = \mathbf{SD} / \mathbf{sqrt}\ (n). \quad [2]$$

Results:

The mean SNR of summed spectra increased with p^0.5 (Fig. 1), thus improving the reliability of LCModel estimates. Figure 2 shows the decrease of Cramer-Rao bounds for the sum of glutamate+glutamine (Glx) to values between 14% and 19%. As the bootstrap resamples were similar in SNR and SD to subsets of p = 5, larger subsets are more efficient in increasing SNR and reliability.

The percent rank display in Figs. 3 and 4 visualizes deviations of the subset distributions of Glx and tNAA (N-acetyl-aspartate+NAA-glutamate) from a normal distribution (crossing the mean at 50% with a slope of SD). Glx demonstrates nicely the decrease in SD(p). The original data of tNAA showed severe skewness and kurtosis. Nevertheless, the distributions were close to normality for medium p (colored). For large p (e.g. the "jacknife" with $p = n - 1 = 9$), the large correction factors strongly enhanced minor deviations. The mean tNAA increased with subset size. This bias was due to the noise dependent baseline. The corrected SD values (Eq. [1], Fig.5) were slightly larger than the one obtained from the original data.

Discussion:

Subset averaging is especially beneficial for strongly-coupled metabolites (Glx) and noisy spectra. Subsets of medium size (around n/2) are recommended because SNR and Cramer-Rao-bounds improve most for small p and because they are normally distributed. The large number of subsets also permits a precise definition of confidence intervals, which can be converted to the original SD by Eq. [1]. If the number of subsets becomes excessive (for n > 10) a random sample of subsets can be evaluated. The subsets derived SD values result in more conservative statistical testing. Especially for large p, the SD is overestimated since the fitting error decreases more slowly than Eq. [1]. For the classical bootstrap approach, the SNR (and thus probably the estimates) will depend on the group size, n. This can be avoided with the subset method. It also permits to match the SNR by choice of p in order to reduce SNR-dependent baseline effects if the SNR differs widely between two groups.

Conclusion:

In addition to improved mean concentrations averaging across subsets also yields the SD or confidence intervals. Subsets sizes of about half the sample group sizes are recommended.

References:

(1)Efron: *Soc Indust Appl Math*, Philadelphia 1982.
(2)Helms, Piringer: *MRM* 2001; 46:395-400
(3)Provencher: *MRM* 1993; 30:672-679.
(4)Helms: *NMR Biomed* 2000; 13:398-406.

Fig. 1: SNR increase

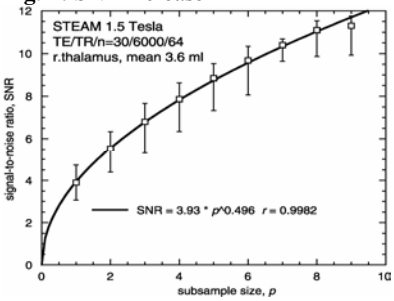

Fig. 2: Cramer-Rao-bounds of Glx

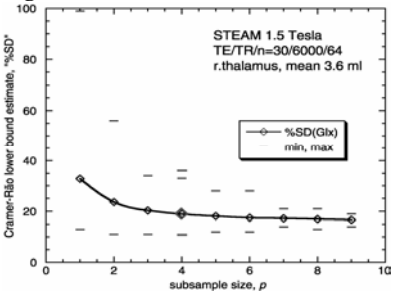

Fig. 3: Distributions of mean Glx

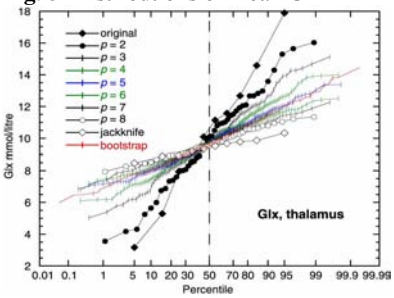

Fig. 4: Distributions of mean tNAA

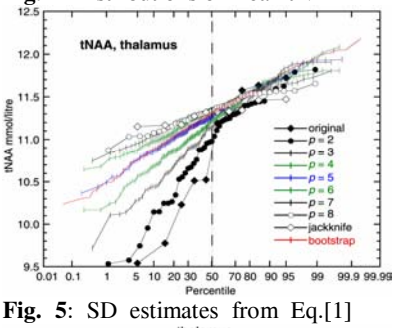

Fig. 5: SD estimates from Eq.[1]