Bootstrap in MRSI: a non-parametric way to assess quantification standard error

H. Ratiney[1], S. Chung[2,3], R. G. Henry[2,3], R. Srinivasan[3], S. J. Nelson[2,3], and D. Pelletier[1]

[1]Neurology, University of California San Francisco, San Francisco, CA, United States, [2]Graduate Group in Bioengineering, University of California San Francisco & Berkeley, University of California San Francisco, San Francisco, CA, United States, [3]Radiology, University of California San Francisco, San Francisco, CA, United States

Introduction:

MR spectroscopic Imaging (MRSI) by providing a multi-voxel data set offers the possibility to apply a resampling technique such as bootstrap [1] to estimate different standard errors (SE). Although used in DTI and fMRI, the bootstrap has been poorly explored in MR spectroscopy. Cramér Rao Lower Bounds [2] (CRLB) is the usual method used to assess quantification results quality by providing, an estimation of the lowest standard deviation theoretically attainable for a given parameter. We propose to estimate a bootstrap amplitude (or concentration) standard error from a large amount of MR spectroscopic imaging in vivo data acquired on Multiple sclerosis patients and to compare the results with the classical, analytically calculated CRLB. The calculations were performed on the quantification results of two short echo time quantitative methods: LCModel and Quest. Results showed that the calculated bootstrap standard error behaves similarly (strong correlation) compared to the method CRLBs provided by Quest for the metabolite containing a singlet (NAA, Cre, Cho) but showed a lower correlation with the CRLB provided by LCModel.

Method :

231 Multiple Sclerosis patient data were used for the study. 3D PRESS short echo time CSI data set were acquired on these patients on a 3T GE scanner using 8 channel-coil in reception. T1 weighted SPGR were also acquired and segmented using the FAST algorithm leading to the assessment of the proportion of white matter *p(WM,i),* proportion of gray matter *p(GM,i)* within each voxel i. A total of spectroscopic voxels were quantified by LCModel and Quest. The metabolite basis set used for the two methods contained: NAA+NAAG, tCr, tCho, mI, Glu, Gln, GABA, Tau and Glc.

The sample considered as the original data set was composed of the points $\mathbf{a}=\{a_i{}_i=a(p(WM,i),p(GM,i)), i=1\ldots N\}$ where a is the amplitude/concentration estimated either by QUEST or LCModel. for the voxel i, N being the total number of voxels used in the analysis. For each patient:

- B=500 bootstrap samples were calculated**.** A bootstrap data set is $\mathbf{a^*}=\{a_1^*,a_2^*,\ldots a_N^*\}$ where each a_i^* equals any one of the N members of **a** with equal probability (drawing with replacement). Then the component of a* were sorted considering the *p(WM,i)* in ascending order. We thus obtained a*as a function of the WM content of the voxel. A sliding window averaging filter was then applied on a*(WM) to obtain an estimated mean value $\overline{a}_i$ for each a_i assuming that the metabolite amplitude of voxel with similar tissue content should have similar amplitude. Only the voxel verifying *(WM)+p(GM)*>50% were included in the analysis. Standard errors are then estimated on the mean values obtained at each WM content value.
- To be able to compute a fair correlation between the drawn bootstrap standard errors and the CRLBs of the two quantitative methods, we also applied a sliding window filter to the CRB_{quest}(WM) and the $CRB_{lcmodel}$(WM) and we calculated the correlation between the relative standard errors : smoothed CRB_{quest}(WM,i)/a(WM,i) versus $SE_{bootstrap}$(WM,i)/ $\overline{a}_i$ (WM) as well as the correlation smoothed $CRB_{lcmodelt}$(WM,i)/a(WM,i) versus $SE_{bootstrap}$(WM,i)/ $\overline{a}_i$ (WM) for each patient.
- We computed the distance between the mean values of CRB_{quest}(WM,i)/a(WM,i) and $SE_{bootstrap}$(WM,i)/ $\overline{a}_i$ (WM) as well as the distance between the mean values of $CRB_{lcmodel}$(WM,i)/a(WM,i) and $SE_{bootstrap}$(WM,i)/ $\overline{a}_i$ (WM)

Results:

We found that the calculated bootstrap standard error depended on the size of the sliding window (SW size) that is the number of voxel used to calculate the mean value in the proposed bootstrap method. The correlation between the relative bootstrap standard error and the smoothed- relative $CRLB_{quest}$ as well as the correlation between the relative bootstrap standard error and the smoothed - relative $CRLB_{lcmodel}$ for NAA for two sliding window size and 231 patients are shown figure 1. The CRLBs provided by QUEST correlate with the non parametric bootstrap method standard error. The correlation between the CRLB from LCModel and the bootstrap increases if we reduce the size of the sliding window. The distance between the mean values of the relative bootstrap standard error and the relative CRLB of Quest and Lcmodel for NAA are shown figure (2). The CRLB given by LCModel are clearly higher than the standard error provided by the bootstrap or Quest.

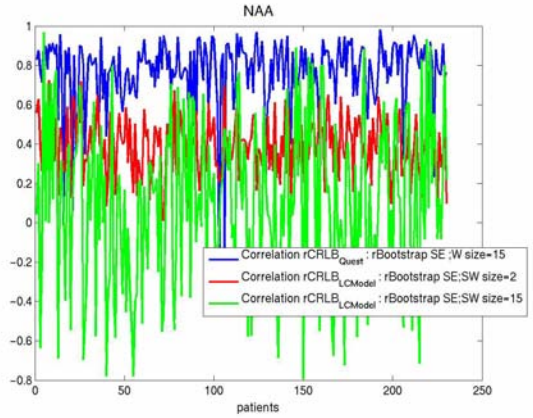

Figure 1: Correlation Results between CRLB and bootstrap SE

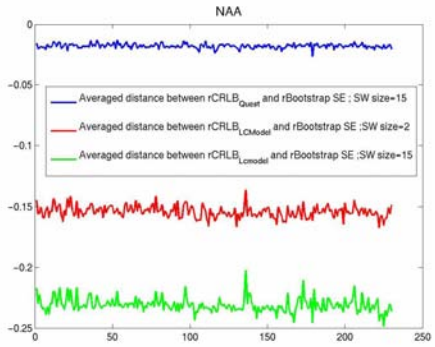

Figure 2 : mean distance value between CRLB and bootstrap SE

Conlusion:

This work proposed one application among others of bootstrap for MRSI data. The proposed method incorporates in the calculation of the amplitude parameter standard deviation the knowledge of the other voxels quantification results. Moreover, it has the advantage compared to the CRLB of not suffering from an approximate model function as it is non-parametric and does not suppose an underlying model. In case of short echo time where the behavior of the macromolecules signal, especially in pathology, is not clearly defined such method have a great interest. Refinement of the method could investigate regional standard error differences by applying the bootstrap separately on the different slices. From this preliminary work it seems that the standard error given by LCModel are clearly penalized compared to the CRLB given by QUEST or the one given by the bootstrap error, allowing in practice a safer rejection criterium. On the other hand, the CRLB provided by Quest clearly correlate with the standard error found by bootstrap.

Acknowledgement: Study supported in part by a National Multiple Sclerosis Society grant (RG-3517-A).

Reference: [1] Efron B. Annals of Staitistic [2] Cavassila S., Deval S., Huegen C., van Ormondt D., and Graveron-Demilly D., NMR in Biomedicine, vol. 14, pp. 278--283, 2001.[3] S.W. Provencher: Estimation of metabolite concentrations from localized in vivo proton NMR spectra. Magn Reson Med 30, 672 (1993).
1993[4] H. Ratiney, et al. NMR in Biomedicine, 2005 Feb; 18(1):1-13