Probabilistic averaging: an instinctive method of averaging according to data confidence applied to Cardiac 31P MRS

L. E. Cochlin[1], and K. Clarke[1]
[1]Cardiac Metabolism Research Group, University of Oxford, Department of Physiology, Anatomy and Genetics, Oxford, United Kingdom

Objective

Cardiac ^{31}P MRS is a potentially highly useful tool for probing cardiac energetics and informing clinical management of heart disease, but whose precision suffers from low signal to noise ratio and overlapping peaks. In addition, however, valuable information describing the variance of a spectrum's fitting remains unused or ineffectively used when determining changes in cardiac PCr/ATP in response to treatment or disease. The widely used t-test (paired and unpaired) assumes that the variance remains unchanged between experiments. This is not a valid assumption to make in cardiac ^{31}P MRS because the signal to noise quality can change due to factors such as shim, coil loading (tune and match), subject motion.

Using cardiac ^{31}P MRS data as an example application, this work demonstrates an intuitive approach for weighting data according to the variance associated with fitted peak amplitudes and specifying a probability for each outcome. The Cramér-Rao lower bound (CRLB) of variance is widely used as a measure of attainable precision by MRS fitting algorithms where standard deviations can not be obtained by repeated measurement.

Methods

31P spectroscopy: 11 fit, healthy male student volunteers were scanned in a Siemens Tim Trio 3T (Erlangen, Germany). Data were acquired with AW-CSI protocol described [Tyler, NMR Biomed. 2008]. An optimized RF pulse, centred between the γ- and α-ATP resonance frequencies ensured uniform excitation of all peaks, enabling PCr/ATP ratios to be calculated using an average of all three ATP peaks (Figure 1). Correction factors for NAD, NOE and RF saturation identified in previous experiments, and blood correction factors [Neubauer, Circ. 1992] were applied. Non-localized inversion recovery spectra were used in conjunction with a calculated RF field profile to determine subject specific flip angles at each chosen voxel.

Fitting: Spectra from three mid ventricular voxels per subject (Figure 2) and their spectral sum were fit using jMRUI [Naressi, Comp Biol Med. 2001] using the AMARES [Vanhamme, JMR 1997] fitting algorithm with standard prior knowledge for all expected peaks.

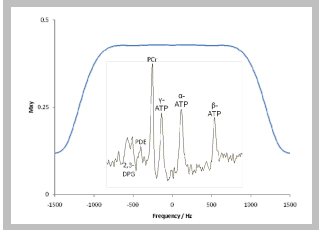

Fig1: Profile of the optimized RF pulse provides uniform excitation of the measured peaks.

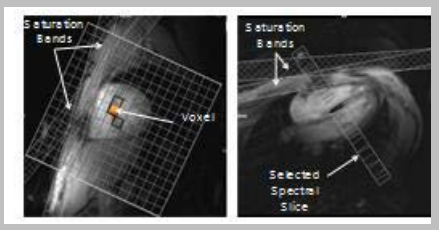

Fig2: Orientation of saturation bands and CSI matrix shown in 2 chamber cardiac short axis and 4 chamber cardiac horizontal long axis views.

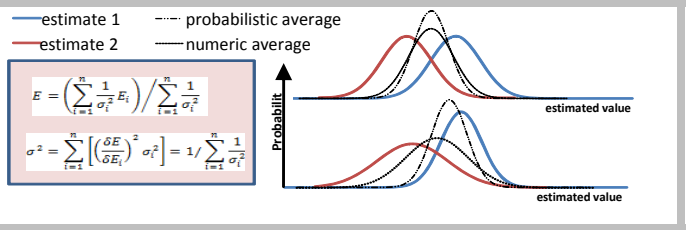

Fig3 top: when both estimates have the same variance the resulting PDF (dashed) has expected value centred between the two estimates. **Bottom:** unequal variances results in an expected value closer to the sample with the lowest variance.

Probabilistic averaging: The ratio of peak areas fitted to the β-, γ-, and α-ATP peaks are assumed to be unity since additional contributions to their resonances from other phosphates are small, or accounted for (20% of α-ATP for NAD contribution). Thus, ATP amplitudes and corresponding CRLB's calculated by AMARES fitting provide three measures of the same molecule. Likewise, the three voxels of interest provide three measures of the same PCr/ATP in mid septal myocardium. The task of integrating these measures is now that of finding the maximum likelihood estimator for their average value E, and calculating the variance of that estimate σ^2 (Figure 3).

This is essentially the propagation of errors for a function of several measured quantities formulated by Gauss in 1809, applied to the AMARES fitted peak amplitudes and their corresponding CRLB (which assumes Gaussian distribution of error characteristics) to specify a probability density function (PDF).

Results and discussion

Figure 4 displays the final outcome of treating 11 healthy volunteer datasets with numeric averaging versus probabilistic averaging. Considering first the data from individual subjects:

- Variance of every subjects PCr/ATP has been reduced resulting in increased calculated probability that each represents their true value.
- Inter subject PCr/ATP spread of values demonstrated in this group of healthy volunteers is appreciably reduced by probabilistic averaging whilst the average value remains statistically unchanged.
- These factors lead to a notably improved confidence in the overall group average PCr/ATP (solid lines) for the probabilistic method.

The process of treating data by considering probability density is illustrated in Figure 3 where it may be seen that a data point with high variance is deemed to have less to say about the true value of the datum than a low variance measure of the same factor.

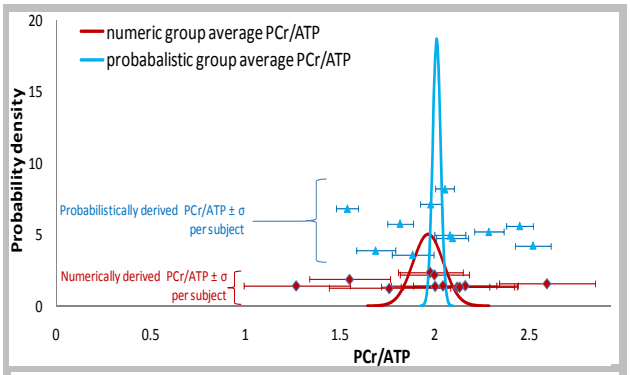

Figure 4: solid lines represent comparison of PCr/ATP probability density functions for standard numeric and probabilistic group averages. These averages are derived by:

red) spectral sum of 3 voxels (in jMRUI) and numeric average of 3 ATP resonances; red diamonds show per-subject PCr/ATP ± σ plotted against their calculated probability

blue) probabilistic average of 3 ATP resonances, followed by probabilistic average of data from 3 voxels; blue triangles show per-subject PCr/ATP ± σ plotted against probability

Conclusion

Maximum likelihood estimation is an intuitive approach to data involving an assessment of the probability that a data point represents the true value. Here, we use the Cramér-Rao lower bound of variance generated by the AMARES fitting algorithm to weight our use of each peak amplitude according to this measure of confidence.

This probabilistic approach leads to a more accurate and intuitive parameter estimation, increasing the confidence in the final result by using all available information. By processing cardiac 31P MRS data in this way, noisy measures contribute less to the final estimate than cleaner, more confident data. The outcome is one of appreciably reduced variance, increased confidence and a reduction in potential for bias when processed with common statistical tests.