Statistical Modeling of Longitudinal Total Choline Measurements during Chemotherapy

A. M. Brearley[1], L. E. Eberly[1], N. R. Mraz[2], M. T. Nelson[2], D. Yee[3], M. Garwood[4], and P. J. Bolan[4]

[1]Biostatistics, University of Minnesota, Minneapolis, MN, United States, [2]Radiology, University of Minnesota, Minneapolis, MN, United States, [3]Masonic Cancer Center, University of Minnesota, Minneapolis, MN, United States, [4]Center for MR Research / Radiology, University of Minnesota, Minneapolis, MN, United States

Introduction: There is great interest in using quantitative total choline measurements of *in vivo* breast cancer to monitor the response to chemotherapy. An earlier study has shown that choline levels drop within 1 day of beginning successful therapy (*1*), but other studies using more delayed follow-up MRS measurements (days or weeks post-treatment) have found that the specificity of choline changes is more variable (*2-4*). A current multi-site clinical trial, ACRIN 6657, is evaluating follow-up time points at 1-4 days after starting chemotherapy.

Designing these and future clinical trials requires more information about the time variation and reproducibility of total choline measurements during chemotherapy. In this work we describe a retrospective analysis of choline measurements acquired in breast cancer subjects undergoing neoadjuvant chemotherapy. We proposed a series of models to describe the time-course of the choline response during therapy, and selected an optimal model based on Akaike and Bayesian information criteria. The model fits, and the relative suitability of the various models, were then used to infer important characteristics of the choline response.

Methods: A retrospective analysis was performed on a completed clinical trial, in which breast cancer patients undergoing neoadjuvant chemotherapy were scanned before and during the first course of treatment. The targeted scan dates were -7, +1, +21, +42, and +63 days relative to the first treatment. The scans were performed on a 4 T MR system (Varian console, Siemens gradients, Oxford magnet) with custom-built unilateral breast coils. MR methods included dynamic contrast-enhanced imaging followed by quantitative single-voxel spectroscopy measurements (details in ref. *1*). Of the 74 subjects in the trial, 41 subjects had measureable total choline levels both pre- and post-therapy, and were included in this analysis. After excluding failed measurements there were 2-5 measurements per subject, for a total of 132 measurements acquired -34 to +71 days relative to the first treatment. Due to the small sample size, the entire dataset was treated as a single group without separation based on clinical response. The raw data are shown in figure 1 after transformation to natural log scale to reduce non-normality.

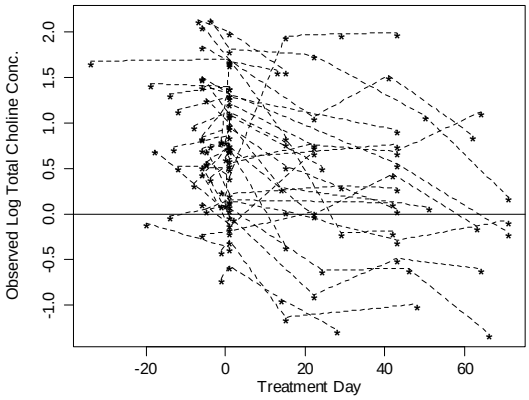

Fig. 1: The log-transformed time series data for all 132 measurements in 41 subjects.

We formulated a series of physically-plausible general linear mixed models to describe the total choline value as a function of treatment-day. The six model types are shown in figure 2. The complete set of covariates includes pre-treatment slope, y-intercept, step change, and slope change. The time lag between treatment and change was set to either 0, 2, or 15 days. Several variations of each model type were considered, using either fixed or subject-specific values for the covariates. In total, 43 models were tested. The data were fit to each model with the PROC MIXED function in SAS 9.1, using the restricted maximum likelihood criterion. The suitability of each model was evaluated using the Akaike Information Criterion (AIC), corrected AICc (*5*), the Bayes Information Criterion (BIC), and where possible, likelihood ratio tests (LRT). These tests compare models on the basis of goodness-of-fit while correcting for degrees of freedom, and are thus appropriate for selecting models that best describe the data without introducing extraneous parameters. The square root of intra-subject (residual) variance for the best model was taken as a best estimate of the measurement reliability.

Results and Discussion: The total choline distribution was approximately log-normal: before transformation the skewness was +1.4, after log transformation the skewness was reduced to -0.1. All modeling was performed on transformed data. Comparing the performance of the various models leads to the following observations:

1) All model variations with subject-specific y-intercepts had improved performance compared to models with fixed y-intercepts. This suggests that the absolute choline values are not suitable for modeling response; subject-by-subject normalization is needed.
2) Adding a step-change at day 0 (no lag) did not improve model fits, but a step change at day 2 did improve the model fits. This suggests that scanning 1 day post-therapy may not be an optimal study design. No significant difference was found between models with 2-day and 15-day lags; there was insufficient data to evaluate intermediate lag times.
3) Those models with post-therapy slopes were better than those without slopes, for both fixed and subject-specific slope models. Furthermore all models with post-therapy slopes had negative slope values, indicating that choline decreases in response to therapy.
4) In all models with a step drop and/or a post-therapy slope, the net total choline decrease after 60 days is comparable to the baseline total choline level. This indicates that on the whole, total choline drops to approximately zero at 60 days. This suggests that clinical trials should not be designed to measure total choline after two months of treatment.

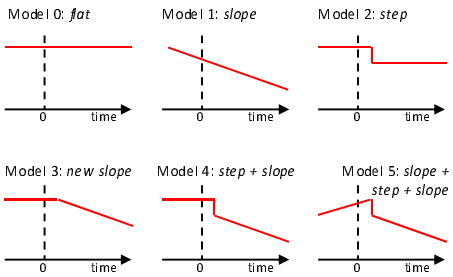

Fig. 2: The six models used for describing total choline response as a function of treatment-day. The vertical dashed line indicates the first day of chemotherapy.

Considering all comparisons, the most suitable model variant was type 4, with a step and slope change at day 2, and subject-specific steps, slopes, and y-intercepts. This model has an 83% likelihood of being the most the suitable, based on Akaike weightings. Using this model, we estimate the measurement reliability for a single measurement is 0.24 log choline units, which leads to asymmetric confidence intervals in real units. For a typical subject with a choline concentration of 2.1 mmol/kg, the 95% confidence interval would be (1.31, 3.42) mmol/kg.

Conclusion: A general linear mixed model for describing the time-course of total choline measurements during chemotherapy was developed and optimized using Akaike and Bayesian information criteria. The modeling process provides insight into the overall choline response that can be used for designing future clinical trials, and gives specific estimates of the individual subject's measurement variation and the overall decrease of choline during the course of therapy.

References: 1) Meisamy et al., Radiology 2004; 2) Tan et al., ISMRM 2006; 3) Baek et al., ISMRM 2007; 4) Danishad et al., ISMRM 2007; 5) Burnham and Anderson, *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*, New York, 2002.

Acknowledgements: NIH grants R01 CA120509, P41 RR08079, and P30 CA77398