Quantitative Quality Assurance Metrics in a High Angular Resolution Diffusion Imaging (HARDI) Multicenter Study

Xiaopeng Zhou[1], Ken Sakaie[1], Josef Debbins[2], Robert Fox[1], and Mark Lowe[1]
[1]The Cleveland Clinic, Cleveland, OH, United States, [2]Barrow Neurological Institute, Phoenix, AZ, United States

Target Audience – Researchers planning multicenter trials with advanced MR.

Purpose – To establish quantitative quality assurance (QA) metrics for HARDI in a multicenter phase II clinical trial. In a multicenter trial, QA metrics are necessary to provide confidence in data quality and to indicate when scanner repairs may be necessary. Qualitative assessment of image quality and artifact are necessary parts of the image analysis workflow. Quantitative assessment may provide the ability to detect subtle degradation in scanner performance prior to the appearance of artifact, offering an opportunity to repair the scanner proactively. An important result would be to avoid costly repeat scans or lost data in a multicenter trial. In order to establish bottom-line criteria, i.e. QA criteria that are based on measures that will directly affect study metrics, we propose a receiver operating characteristic (ROC) analysis to establish quantitative thresholds on signal to noise ratio (SNR).

Methods – Images are acquired from BIRN phantoms [1] on a monthly basis from each of 27 sites in the SPRINT-MS trial (11 Siemens TIM Trio, 6 Siemens Skyra, 1 GE Signa EXCITE, 7 GE Signa HDxt, 1 GE DISCOVERY MR750 and 1 GE DISCOVERY MR750w). In vivo images were acquired under a central IRB. The trial is performed within the NeuroNext (www.neuronext.org) network. Scans were acquired with 2.5mm isotropic resolution with 64 b=700sec/mm^2 (only 55 b=700sec/mm^2 were allowed on 1 GE scanner) and 8 b=0 volumes. SNR was determined on a voxel-by-voxel basis by taking the ratio of the mean and standard deviation among b=0 volumes for each scan. An ROC analysis [2] was performed to determine the loss of sensitivity in transverse diffusivity (TD), a marker for demyelination, associated with a given loss of SNR. In this way, we translate a readily measured parameter, SNR, into a quantity relevant to the study, sensitivity to changes in TD. For the ROC analysis, we simulated signal assuming the diffusion tensor model with values, for healthy controls, of 0.5x10^{-3}mm^2/sec for TD and 1.2x10^{-3}mm^2/sec for longitudinal diffusivity. Signal was also simulated for patients, assuming a 10% elevation in TD. Noise was added to the signal, with SNR varying from 5-100 in increments of 0.1 and 10000 noise realizations for each value of SNR. The signal was fit to the diffusion tensor model to generate a distribution, among noise realizations, of TD for control and patient. The cumulative sum of the distributions generated from the normalized histogram of TD associated with the patient is the true positive rate and that associated with the control is the false positive rate. Plotting true positive rate versus false positive rate yields the ROC curve. A 10% drop in sensitivity for a given SNR yields the threshold for acceptable scanner performance.

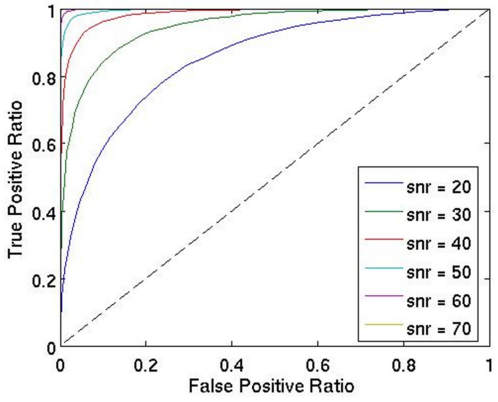

Fig.1 ROC curves for different SNR values.

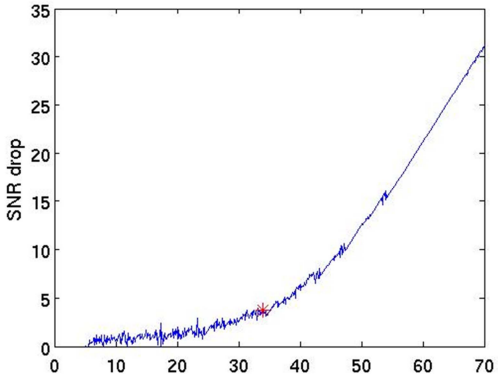

Fig.2 Drop in SNR that leads to a 10% loss of sensitivity as a function of baseline SNR.

Results – Fig. 1 shows ROC curves for different values of SNR. In vivo SNR is ~30 among the GE scanners, ~34 for the Trio scanners and ~60 for Skyra scanners. The SNR of phantom is similar to the in vivo SNR with GE and Trio scanners, but is much higher than the in vivo value for the Skyra scanners (~120). The drop of SNR that leads to a 10% loss in sensitivity to changes in TD for a given baseline SNR is shown in Fig.2. If, for example, a Trio with baseline SNR of 34 exhibits a drop in SNR of 3.7 one may suspect that the scanner requires service.

Discussion – This analysis provides a quantitative way to assess SNR in phantoms for QA purposes. As the SNR in the phantom measurements match that of in vivo measurements for the GE and Trio scanners, we can directly use the phantom measurements to indicate scanner problems. The disparity between SNR values from the phantom and in vivo measurements on the Skyras, however, is problematic. We will investigate the source of this disparity to make the phantom measurements more useful for QA. As the SPRINT-MS trial proceeds, we will determine if this analysis indeed predicts the onset of scanner performance degradation, a result that will be useful for this and future trials.

Conclusion – ROC analysis provides quantitative thresholds for SNR that are relevant to HARDI measurements. We expect this approach to improve the reliability of multicenter imaging trials.

References

[1]. Friedman, et al. JMRI 2006; 23(6):827-39. [2]. Zweig M, Campbell G. Clinical Chemistry. 1993; 39:461-77.