Automatic quality assessment in structural brain magnetic resonance imaging

B. Mortamet[1], M. A. Bernstein[2], C. R. Jack Jr[2], J. L. Gunter[2], C. Ward[2], P. J. Britson[2], R. Meuli[3], J-P. Thiran[4], and G. Krueger[1]

[1]Advanced Clinical Imaging Technology, Siemens Suisse SA, Healthcare Sector IM&WS - CIBM, Lausanne, Switzerland, [2]Mayo Clinic, Rochester, MN, United States, [3]Centre Hospitalier Universitaire Vaudois and University of Lausanne, [4]Ecole Polytechnique Fédérale de Lausanne (EPFL), Signal Processing Laboratory (LTS5), Lausanne, Switzerland

Introduction: Quality assessment of magnetic resonance imaging is of great importance to derive reliable diagnostic information. With the growing number of clinical research and multicenter neuroimaging studies as well as the promising use of computer-aided diagnosis, automated and objective measures of quality are needed to replace visual inspection that is time-consuming and has imperfect repeatability. To address this need, we propose a new method which derives quality indices from air background analysis of magnitude MR images. This region corresponds typically to >30% of the total image volume and provides sufficient information to detect image degradation from several sources such as patient head movement, residual magnetization from incomplete spoiling, blurring, ghosting. For verification and fine-tuning of the method, we evaluate the effectiveness of two different indices to correctly classify 749 3D head scans as either high- or low-quality.

Materials & Methods: In this study, 749 T1-weighted structural head MRI data (188 subjects, 72.5±17.5 years old) were obtained from the Alzheimer's Disease Neuroimaging Initiative ADNI database (*www.adni-info.org*) and acquired on 36 clinical scanners (Siemens Medical Solutions, Erlangen, Germany) using 3D MPRAGE pulse-sequences (TR-TI=2.3-0.9/2.4-1s, 1.25x1.25x1.2/1x1x1.2 in-plane resolution FA=8/9° for 1.5T/3T acquisitions) operating with various software and hardware combinations at 1.5 and 3T [1]. Automatic quality control is achieved in a three-step process. Firstly, an atlas-based algorithm extracts background region over which analyses are performed. Secondly, artifacts regions are isolated by means of morphological operations and a quality index QI_1 is computed as the proportion of artifactual voxels relative to the background size (see solid red line in Fig1). Thirdly, a chi distribution [2] is fitted to the background noise histogram (without artifactual voxels detected in Step 2) using maximum likelihood estimation. A quality index QI_2 is computed by adding up the goodness-of-fit (absolute error between histogram and fit) and artifactual voxels from QI_1 (see dot line in Fig1). To investigate the predictive performance of the proposed indices, results are compared to qualitative grades assigned by the ADNI quality control center (taken as gold standard). Sensitivity and specificity are considered as the true positives (high-quality prediction) rate and the true negatives (low-quality prediction) rate, respectively. Receiver operating characteristics (ROC) curve (see Fig2) represents the range of combinations of sensitivity and specificity achievable over the range of possible cutoff points for our quality indices (*QIs*). ROC is used (a) to evaluate the performance of each QI measured by area under the curve (AUC, see Fig2) and (b) to compare the discriminative abilities of QIs in order to identify the preferred one. *QIs* cutoff values are determined by equalizing sensitivity and specificity. Analysis of variance is performed to detect significant mean differences among quality groups (low-/high-quality) resulting from classification based on either QI_1 or QI_2.

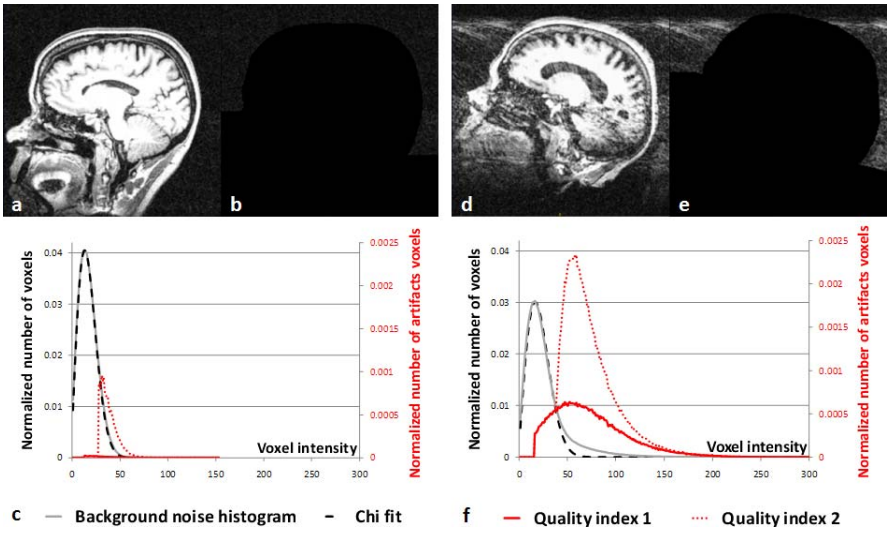
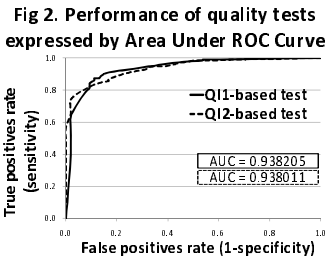


Fig 1. Magnitude sagittal slices of high- (a) and low- (d) quality datasets (Allegra, single channel RX coil) with corresponding background masks (b-e), intensity distributions (c-f)



Results: Both quality indices exhibited excellent prediction performances (AUC>0.9). Trading sensitivity and specificity at equal rates (87.19 & 85.18% for QI_1 & QI_2) provided optimal thresholds of 5.06e-3 & 5.7e-2. ANOVA revealed significant differences among the two quality groups for each *QI* ($p < 0.001$). Overall, the model (used for QI_2) fitted the data well and our quality indices QI_1 and QI_2 appeared to be both accurate and consistent.

Discussion: In this work, a novel method is proposed to measure MR image quality in an automatic manner by analyzing the air background of magnitude images. The two quality measures presented here both perform equally well and have proven to be very efficient and robust in predicting overall image quality in a large and heterogeneous data collection when compared to grading of an experienced reader. The results suggest that the proposed method may have a great potential in clinical research studies and routine clinical practice. It could offer the ability to rule-out the need for a repeat-scan while the patient is still in the MR bore. Because of its automation, quality assessment is highly reproducible (if the cutoff point is held fixed). The performance of our quality indices could even be further improved by differentiating artifacts subtypes with regard to their spatial location, checking protocols or analyzing k-space or brain tissue intensity distribution. Initial results show for example, that our atlas-based background segmentation technique performs well in detecting poor slice positioning such as nose-wrap, which typically results in a poor quality rating by the reader. Since the QI_2 index is based on a noise model, extension to data that have been processed with filters and parallel reconstruction is expected to result in reduced performance. As initial investigations exhibit, the QI_1-based test, however, is expected to be directly extendable to data with other contrasts, using parallel-imaging techniques or corrected for various MR-related intensity inhomogeneities or geometrical distortions since the 3D connected structure of artifacts is not altered.

References: [1] Jack CR, Jr. et al. The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI methods. J Magn Reson Imaging 2008;27(4):685-691. [2] Constantinides CD et al. Signal-to-noise measurements in magnitude images from NMR phased arrays. MRM 1997;38(5):852-857.

Acknowledgement: data collection and sharing for this study was funded by the ADNI (PI: Michael Weiner; NIH grant U01 AG024904)