Estimating Test-Retest Reliability in Functional MR Imaging: three-state independence model

Yue Zhang[1], Xiaoying Wang[2], Jue Zhang[1], Xiaoping Hu[3], and Jing Fang[1]

[1]*College of Engineering, Peking University, Beijing, Beijing, China,* [2]*Department of Radiology, Peking University First Hospital, Beijing, Beijing, China,* [3]*Department of Biomedical Engineering, Georgia Institute of Technology / Emory University, Atlanta, Georgia, United States*

Introduction: The reliability is important for functional magnetic resonance imaging (fMRI) data. Previous study has developed a statistical model (independence model) to quantify the test-retest reliability [1,2]. While only two states (active or inactive) was considered for these studies [1,2]. More and more fMRI experiments have detected three state regions in the brain, including active, deactive and non-significant regions [3]. In order to quantify the test-retest reliability of the fMRI data with three state regions, we developed three-state independence model based on the independence model.

Materials and Methods:

The three-state independence model: Similar to the independence model, the data are assumed statistically independent across both voxels and replications, and the probability of individual voxel has the same distribution under the three-state independence model. The probability of one voxel classified active k1 times, deactive k2 times and non-significant M-k1-k2 times was defined:

$$P_{onepixel} = A1 * C_M^{k1} * C_{M-k1}^{k2} P_{11}^{k1} * P_{12}^{k2} * (1 - P_{11} - P_{12})^{M-k1-k2} + A2 * C_M^{k1} * C_{M-k1}^{k2} P_{21}^{k1} * P_{22}^{k2} * (1 - P_{21} - P_{22})^{M-k1-k2} + (1 - A1 - A2) * C_M^{k1} * C_{M-k1}^{k2} P_{31}^{k1} * P_{32}^{k2} * (1 - P_{31} - P_{32})^{M-k1-k2} \quad (1),$$

Where P_{ij} is the probability that i-state voxel is classified j-state.(where i,j=1,2,3, and 1 represents active, 2 represents deactive and 3 represents non-significant). A1 represents the proportion of truly active voxels in the brain, A2 represents the proportion of truly deactive voxels in the brain, M represents the times replicating an fMRI experiment. By the assumption of independent voxels, the likelihood function is a product of the likelihoods for all the individual voxels, and the log of the likelihood function for the data thus takes the followed form plus a constant that is independent of the parameters. Where $n_{k1,k2}$ is the number of voxels classified active k1 times, deactive k2 times and non-significant M-k1-k2 times.

$$L(P, A) = \sum_{k1=0}^{M} \sum_{k2=0}^{M-k1} n_{k1,k2} \ln[A1 * P_{11}^{k1} * P_{12}^{k2} * (1 - P_{11} - P_{12})^{M-k1-k2} + A2 * P_{21}^{k1} * P_{22}^{k2} * (1 - P_{21} - P_{22})^{M-k1-k2} + (1 - A1 - A2) * P_{31}^{k1} * P_{32}^{k2} * (1 - P_{31} - P_{32})^{M-k1-k2}] \quad (2)$$

The method of Maximum Likelihood (ML) was used to estimate the parameters in the model. And the threshold values for one voxel classified active or deactive was selected according to the ratio of all voxels in the brain, ranged from 1% to 30% with an incremental interval of 1% respectively (30*30=900 threshold values). Considering the P_{11},P_{22},P_{33} denote the "hit" probabilities, the sum of them is defined by "accurate ratio", i.e. "accurate ratio" = $P_{11}+P_{22}+P_{33}$, and we think "accurate ratio" is more large, the classification method is more reliable.

Subjects: One right-handed subjects (male, age 33 years) participated in this study. The participant received electrical acupoint stimulation (EAS) for ten times, and performed only one run each day, one week apart. Electroacupuncture was delivered with an MRI-compatible Hans 200 electric acupoint stimulation device (Nanjing Gensun Medical Technology Co. Ltd., Nanjing, China). EAS at 100 Hz was administered with a pair of adhesive skin electrodes placed on the LI4 (*Hegu*) acupoint of left hand. The intensity was adjusted to a maximal but comfortable level for EAS [3,4], with the intensity being 9 mA each run. Each fMRI scan lasted 4.5 minutes comprised of alternating 0.9-minute blocks of rest (A) and EAS (B) with the rest block first, i.e. A-B-A-B-A, resulting in 1.8-minute EAS and 2.7-minute rest.

Image Acquisition: All MRI experiments were performed on a General Electric 3T Signa system (GE Medical Systems, Waukesha, WI) with a standard head coil. Functional data were acquired using a double readout spiral-out sequence with simultaneous Gradient-echo cerebral blood flow (CBF) and blood oxygenation level dependent (BOLD) acquisitions, at short and long TEs, respectively [5,6]. Both readouts utilized slice thickness / gap (THK) of 8.0 / 2.0 mm with 3.6 x 3.6 mm^2 in-plane resolution, a 230 mm^2 field of view (FOV) with a 64 x 64 acquisition matrix, a repetition time (TR) of 3000 ms and a 90° flip angle. Following inversion, the time of the saturation pulse was 700 ms with an 800 ms delay between saturation and excitation. CBF/BOLD readouts were acquired at TEs of 3.1/30 ms, respectively, covering 12 axial slices of the whole cerebrum.

Data Analyses: For each run, data was first motion corrected. *Sinc* interpolation of the ASL time course was performed to obtain time-matched control and label images, following by subtraction to suppress BOLD contamination [7]. Each run's functional images were then coregistered with the corresponding anatomical images to facilitate transformation to Montreal Neurological Institute (MNI) space and resampling to isotropic 2 mm^3 voxels. The data was spatially smoothed with a Gaussian kernel of 8 mm at full-width half-maximum and analyzed using a random effect approach subsequently. A general linear model (GLM) with hemodynamic response function was applied to the smoothed data to model MR signal responses of each experimental paradigm for all runs individually. The data of BOLD were high-pass filtered with a cut-off of 240 s. Single run contrast maps were then analyzed by a second-level analyses using two-sampled t test to obtain group comparisons. A combined threshold of $p < 0.01$ for single voxel with a minimum cluster size of 10 voxels was used for the generation of the comparison maps.

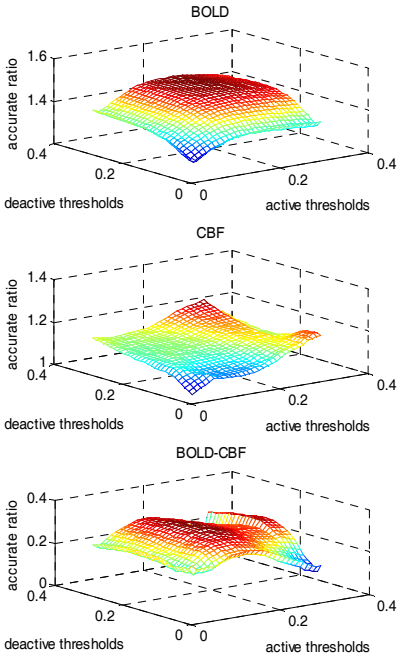

Fig2 Accurate ratio maps for BOLD,CBF and difference between them.

Results: All the three states (active, deactive and non-significant) regions were found in the group comparison maps for both CBF and BOLD (Fig1). For all the 900 threshold values, the "accurate ratio" of BOLD is larger than that of CBF. And the areas under the curve were 1.2963e+003 and 1.0405e+003 for the BOLD and CBF respectively.

Discussion and Conclusion: In summary, this study developed the three-state independence model to quantify the test-retest reliability of the fMRI data, which is more suitable for the experiments including three states (active, deactive and non-significant) regions. And the model was applied to acupuncture fMRI data, the results indicated that the reliability of BOLD was higher than that of CBF, this maybe comes from the bigger SNR (signal-to-noise ratio) for BOLD.

Reference: [1] CR Genovese et al., MRM 1997;38:497-507.[2] DC Noll et al., MRM 1997;38:508-517.[3]Y Zhang et al., Neurosci Lett 2012;530(1):12-17. [4] WT Zhang et al., Brain Res 2003 ;982 :168-178.[5] EC Wong et al., NMR Biomed 1997 ;10:237-249. [6] EC Wong et al., Magn Reson Med 1998 ;39:702-8. [7] TT Liu et al., Neuroimage 2005 ;24 : 207-15.

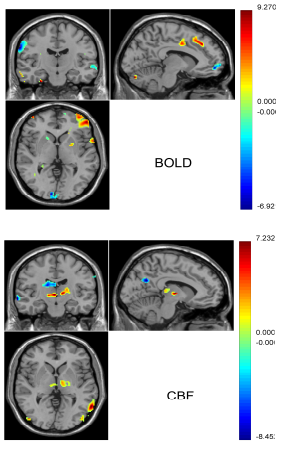

Fig1 group comparisons for BOLD and CBF.