NeuroImage 50 (2010) 479–490

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Group specific optimisation of fMRI processing steps for child and adult data

J.W. Evans [a,c], R.M. Todd [b], M.J. Taylor [a,c], S.C. Strother [a,d,e,*]

[a] Institute of Medical Science, University of Toronto, Toronto, Canada
[b] Ontario Institute for Studies in Education,University of Toronto, Toronto, Canada
[c] Research Institute and Diagnostic Imaging, The Hospital for Sick Children, Toronto, Canada
[d] Rotman Research Institute, Baycrest, Toronto, Canada
[e] Medical Biophysics, University of Toronto, Canada

ARTICLE INFO

Article history:
Received 20 July 2009
Revised 11 November 2009
Accepted 15 November 2009
Available online 3 December 2009

Keywords:
fMRI
Child
Adult
Preprocessing
Motion parameter regression
Motion correction

ABSTRACT

Motion is a major issue in functional magnetic resonance imaging (fMRI) dataseries and causes artifacts or increased overall noise obscuring signals of interest. It is particularly important to be able to control for and correct these artifacts when dealing with child data. We analysed the data from 35 children (4–8 years old) and 13 adults (18–30 years old) during an emotional face paradigm. The children were split into low and high motion groups on the basis of having less or more than an estimated maximal movement of one voxel (3.75 mm) and one degree of rotation in any motion direction between any pair of scans in the run. Several different preprocessing steps were evaluated for their ability to correct for the excess motion using agnostic canonical variates analysis (aCVA) in the NPAIRS (Nonparametric, Prediction, Activation, Influence, Reproducibility, re-Sampling) framework. The adult dataset was reasonably stable whereas the motion-prone child datasets benefited greatly from motion parameter regression (MPR). Motion parameter regression had a strong beneficial impact on all datasets, a result that was largely unaffected by other preprocessing choices; however, motion correction on its own did not have as much impact. The low motion child group subjected to MPR had reproducibility values at par with those of the adult group, but needed almost twice as many subjects to achieve this result, indicating weaker responses in young children. The aCVA showed greater sensitivity to the task response pattern than the mixed effects general linear model (mGLM) in the expected face processing regions, although the mGLM showed more responses in some other areas. This work illustrates that preprocessing choices must be made in a group-specific fashion to optimise fMRI results.

Introduction

Motion is a major issue in functional magnetic resonance imaging (fMRI) dataseries and causes problems of two types: artifacts that mimic 'true' activations, in the case of task-correlated motion, or increased overall noise in the dataset obscuring the signal of interest, in the case of uncorrelated motion (Hajnal et al., 1994; Thacker et al., 1999). It is of particular importance to be able to control for and correct these artifacts when dealing with child data as children are likely to be less compliant at lying still than their adult counterparts. Retrospective studies of large pediatric cohorts studying language development have found significant interactions among age, task and motion (Byars et al., 2002; Szaflarski et al., 2006; Yuan et al., 2008). Similar differences in head motion characteristics between subject groups have also been noted in a study of a battery of motor tasks in stroke patients and adult controls (Seto et al., 2001).

While it is possible to use restraint techniques such as bite bars in adults, it is more common to use the easily implemented compressive padding that restricts head motion in less compliant or paediatric populations. Prospective motion correction can also be used to correct for motion during the scan (Thesen et al., 2000; Zaitsev et al., 2006), but these techniques are not implemented on many scanners in current use. In paediatric subjects pre-scan motion training in a mock scanner with automated motion tracking giving feedback to the child has also been used to try and improve behaviour (Kotsoni et al., 2006); it is, however, no guarantee of compliance during the actual scan, and greatly increases the time required of the children and their parent to be in the scanning suite, which in and of itself can reduce cooperation.

Motion correction tools based on post-acquisition image intensity-based registration are included in all fMRI analysis software packages and also serve to estimate the motion parameters of the brain (usually in 6 directions assuming rigid body motion) during the scan. Accurate estimation of these parameters has been verified in both phantom and human experiments and the resulting corrections have been shown to improve the magnitude and cluster size of detectable activations. Although the correction algorithms vary, most produce similar results in general linear model (GLM) analyses (Oakes et al., 2005; Ardekani

* Corresponding author. Institute of Medical Science, University of Toronto, Toronto, Canada.
E-mail address: sstrother@rotman-baycrest.on.ca (S.C. Strother).

1053-8119/$ – see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2009.11.039

et al., 2001). There is some evidence showing that mutual information-based measures are more robust, particularly in datasets with very little motion where artifacts from the correction algorithm may occur (Freire and Mangin, 2001). Even after motion correction in datasets with little motion, residual movement-related effects remain in the fMRI timeseries perhaps due to interpolation effects in the data, spin-history effects (Grootoonk et al., 2000; Friston et al., 1996; Jiang et al., 1995) or field-inhomogeneity interactions (Andersson et al., 2001).

In the absence of significant task-correlated motion, the inclusion of the estimated motion parameters as regressors in the standard GLM analysis improves sensitivity to the task, and decreases spurious activations and inter-subject variability (Lund et al., 2005, 2006; Johnstone et al., 2006). To facilitate this correction, Johnstone et al. (2006) suggested a procedure to identify motion artifacts in datasets using the raw, motion-corrected (MC), and motion parameter regressed (MPR) images by identifying the voxels which decrease in value after either type of correction.

Yet, not all amounts of motion can be modelled by retrospective correction techniques, which has led to current standards of acceptable motion levels that are often set at a fraction of a voxel in size. However, children frequently move more than this somewhat arbitrarily set standard. This results in the loss of valuable data from the young participants as well as the large commitment of resources dedicated to the collection of neuroimaging results in children. Also, restricting acquired fMRI datasets to children who are compliant to adult motion standards may bias towards a subset of the population that is less active or more mature than their peers or that have different behavioural patterns and skew the results towards an atypical group of young children.

In the current analyses, children with both high and low estimates of motion were analysed to evaluate the effects of a range of motion on the resulting fMRI data. Several different preprocessing steps were evaluated for their ability to correct for the excess motion: no processing, motion correction (MC), estimated motion parameter regression (MPR), detrending (DTR), MPR and DTR, and MC and MPR and DTR, using canonical variates analysis (CVA) with agnostic class labels (Strother et al., 2004). CVA is a multivariate statistical technique that investigates the relation between two sets of variables to extract the patterns common to both sets (e.g. a set of fMRI scans and a set of indicator variables reflecting the different experimental task conditions, or classes, for the scans in the CVA model). The classes in the model can be chosen to group the data according to task condition (as in a general linear model (GLM) design matrix) or it is also possible to label each time point or scan as a separate class. In the latter case, the CVA is effectively 'agnostic' to the experimental task structure as no labels are applied, which is called agnostic CVA (aCVA) in this paper (Kjems et al., 2002). The assumption underlying such an aCVA is that there are multiple observations (e.g. subjects) that share a common but unknown temporal response to the common experimental manipulation. The data were also processed using GLM analysis to evaluate the corresponding changes in a standard univariate model.

It is critical to be able to determine if one or more of the various preprocessing pipelines significantly improves correction for motion. Quantitative evaluation of these various pipelines options can be performed using the NPAIRS (Nonparametric Prediction, Activation, Influence, and Reproducibility reSampling) data analysis framework (Strother et al., 1997, 2004; Kjems et al., 2002). In the NPAIRS software package the neuroimaging data are split randomly into two equal-sized groups of subjects. The chosen analysis pipeline is then independently performed on each half of the data and this splitting and analysis process is repeated (i.e., split-half resampling), fifty times. The agreement between the two results is quantified by calculating a correlation coefficient between each pair of independent, split-half analysis eigenimages. This measure is then a metric of spatial reproducibility, and a global, spatial signal-to-noise ratio (SNR) for the performed analysis. Furthermore, a statistical parametric map (SPM) is robustly generated on a standard *Z*-score scale from the multiple canonical eigenimages produced with split-half resampling.

We address three key points: the optimisation of preprocessing choice with respect to motion likelihood, a quantitative evaluation of the amount of tolerable motion in a dataset and a comparison of multivariate and univariate models with respect to motion sensitivity. These results should serve as a guideline for making more quantitative decisions as to which pipeline option should be chosen and which subjects can be included in analyses when different head motion patterns between groups are of concern.

Methods

Subjects

Thirty-five children (4–8 years, mean 6.4 years, 22 female) and thirteen adults (18–30 years, mean 24.2 years, 7 female) participated in the study. All participants were healthy, with normal or corrected vision, without any history of neurological or psychiatric disorders. This study was approved by the Hospital for Sick Children's Research Ethics Board; the adult subjects signed informed consent, parents signed for their children and children gave informed verbal assent.

Pre-scan

Each child was given half an hour of instruction on the task prior to their scan. This involved being familiarised with scanner sounds and practising the task with exemplar faces on a laptop. The child also practised lying still while simulating pressing a button. An adult watched the child and alerted the child to unintentional movements that occurred with the button press, and asked the child to repeat the press until the child was visibly still.

Task

The fMRI task was a passive viewing block design with 12 epochs each 16s long. Face blocks were interleaved with scrambled face blocks, with each image displayed for 1s. Scrambled faces were created from the face images by randomising their phase. Three blocks

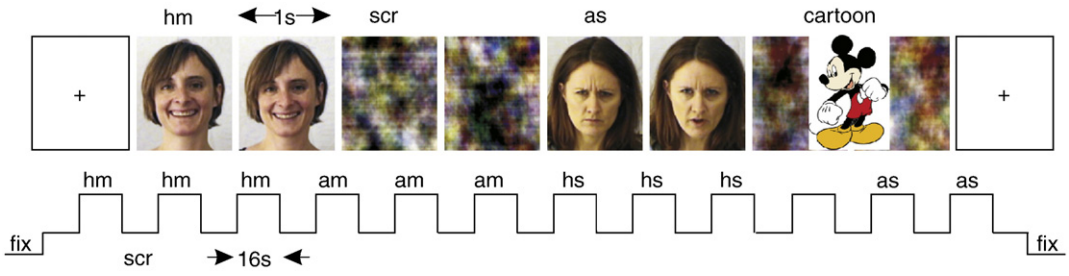

Fig. 1. Block design paradigm and example stimuli. The images show examples of faces with happy and angry expressions and a sample cartoon. The faces were flanked by a fixation cross at the beginning and end of the run. The timing diagram for the paradigm is below. Cartoons were shown at random intervals throughout the task. Happy mother (hm), angry mother (am), happy stranger (hs), angry stranger (as), scrambled face (scr), fixation (fix).

each of four face types were presented: happy mother, angry mother, happy stranger, angry stranger (see Fig. 1). Occasionally, a well known cartoon (e.g. Mickey Mouse) would be presented for 500 ms and the subject pressed a response button to maintain attention.

The paradigm was presented on a computer that was triggered by the start of the imaging sequence using Presentation software (Neurobehavioral Systems, Albany, CA). The stimuli were delivered to the subjects in the magnet using the CinemaVision goggles and headphone system (Resonance Technologies, Northridge, CA). Subject responses to cartoon figures were recorded using an MRI compatible keypad placed under the subject's right hand (Lumitouch, Burnaby, BC, Canada).

Scanning

Image acquisition was performed on a 1.5T GE Signa Excite HD system (GE Medical Systems, Milwaukee, WI) using a standard quadrature head coil. Foam padding was used to constrain the subject's head lightly in the head coil as necessary. Generally no additional restraints were used for our child subjects although, in a few cases, a velcro band was placed around their shins or a strip of tape across their forehead, to remind them to remain still. First a high-resolution whole brain anatomical scan was acquired using an axial fast SPGR sequence (106 slices with TE/TR/flip angle = 8 ms/30 ms/30°; voxel size $1\times1\times1.5$ mm, matrix 256×256). Subsequently, the functional images were acquired using a T2* weighted axial spiral in/out sequence (24 slices withTE/TR/flip angle = 40 ms/2000 ms/90°; voxel size = $3.75\times3.75\times5$ mm, matrix 64×64).

Individual analyses

All individual analyses and preprocessing were carried out using AFNI (Cox, 1996), Analysis of Functional Neuroimages, 11 October 2008 compile date.

Preprocessing

Common preprocessing steps included spatial normalisation and smoothing. A 4 mm full-width-at-half-maximum Gaussian kernel was used to smooth the functional data. Each subject's high resolution anatomical data were warped onto the ICBM 452 atlas using a single twelve parameter affine transform. This transform was then applied to the motion corrected fMRI dataset after visual verification that alignment between the fMRI and anatomical images was good. The displacement between the anatomical and fMRI scan was corrected manually if any was observed.

The nine pipelines evaluated were defined as follows:

(1) no additional processing, beyond that described above (none);
(2) Motion Correction (MC): data were motion corrected using 3dvolreg with Fourier interpolation with registration to the twentieth image of the time series. Motion parameters for each of the six directions (I/S, A/P, R/L, roll, pitch, yaw) were also estimated from this procedure;
(3) Motion parameter regression (MPR): motion parameter estimates for each of the six directions from the motion correction step were included as nuisance regressors;
(4) Polynomial Detrending (DTR): up to third order Legendre polynomials were included as additional regressors;
(5) MC + MPR + DTR;
(6) MPR + DTR;
(7) MPR + Slice Timing Correction (TS): in addition to MPR, the voxel time series was corrected for the interleaved acquisition slice timing by using 3dTshift.;
(8) MPR + first derivative of task (MPR + Diff1): in addition to the motion parameters, the first derivative of the task timeseries (a twelve epoch boxcar function convolved with the AFNI canonical gamma function) was included as a nuisance regressor;
(9) MPR + second derivative of task (MPR + Diff2): in addition to the motion parameters, the second derivative of the task timeseries (a twelve epoch boxcar function convolved with the AFNI canonical gamma function) was included as a nuisance regressor.

Motion

All subject data were motion corrected in AFNI (Cox, 1996) with 3dvolreg using Fourier interpolation and registration to the twentieth image of the time series. The twentieth image was chosen as a neutral reference that was unbiased by the initial scanner equilibration intensity changes and was unaffected by much motion since the beginning time series was where most children were relatively still. The threshold criterion for acceptable motion was chosen to be the estimated maximum displacement in any direction. This maximum excursion was calculated using the difference between the largest positive and negative frame-to-frame movement from the reference time point and was obtained separately for each direction. It should be noted that the displacement of the voxels within a brain volume will not be the same if the head motion includes a rotation. Therefore in the case of rotations, the maximum displacement of a voxel was at the edge of the brain, assuming an average adult brain circumference of about 50 cm.

Based on the motion threshold criterion, the children were separated into 2 groups: a low motion (LM) group with maximal movement of less than one voxel (3.75 mm) or less than one degree of rotation in any direction between any pair of scans per run (with the majority having less than a half a voxel of motion, 22 subjects), and a high motion group (HM, 11 subjects) with greater than one voxel of motion or one degree of rotation in any direction. Two subjects with extreme motion (>5 mm) were excluded. These thresholds were chosen to maximise the number of children included in the analyses while minimising the impact of motion differences.

Group analyses

The children were analysed separately as low (LM, 22 subjects) and high (HM, 11 subjects) motion groups defined above, and the entire group was also analysed (All, 33 subjects). The adult group included all 13 adult subjects.

Agnostic CVA and NPAIRS

The data from each of the pipelines were analysed using agnostic canonical variates analysis (aCVA) in the NPAIRS split-half resampling framework in IDL (ITT Visual Information Solutions, Boulder, CO) to quantitatively evaluate the effects of the processing on the signal and noise structure of the data (Strother et al., 1997; Liow et al., 2000; Kjems et al., 2002; Strother et al., 2004). A detailed description of the NPAIRS framework can be found in the appendix, while a summary is described below. A Java version of this software is available at: http://code.google.com/p/plsnpairs/.

In the NPAIRS framework (illustrated in Fig. 2), the dimensionality of the mean-centered (across time per subject) data was first reduced by performing a principal components analysis (PCA) decomposition and retaining the top 30% of the components (PCA1). This served both to reduce computational load in the further calculations and acted also as a "denoising" step. The group of data was then broken into the two split-half groups and a second PCA was performed on each of the groups separately (PCA2). This second PCA step is necessary to produce invertible data matrices needed in the aCVA model. Using the adult group as an example to summarise the above steps, there were 13 subjects (n) $\times$ 192 time points (t) ($\times\sim$ 30 000 voxels in the brain data matrix, v), which generated 2496 ($n\times t$)

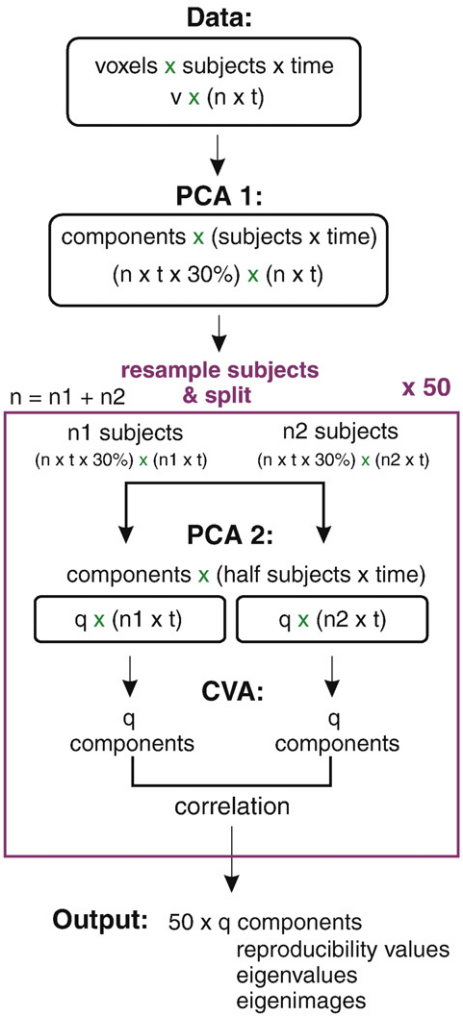

Fig. 2. NPAIRS framework schematic illustrating the data processing steps. The resampling step, repeated fifty times, is a shuffling of subjects which is then split into the two split-half groups of sizes n1 and n2 (for an even number of subjects $n1 = n2 = n/2$). The number of components retained from the second PCA was chosen to be the number, q, that produced the first maximum of the median reproducibility of the task component for separate NPAIRS analyses over a range of values of q (see Optimisation of number of principal components).

components after the first PCA. After keeping 30% of these components, 748 are passed on to the second PCA on the split-half groups with 6 and 7 subjects in each of them respectively, creating data matrices with sizes of 1152×748 and 1344×748 that were used in the aCVA.

In the aCVA, each of the time points was treated as a separate class and therefore the process ignored the experimental task labels. The canonical variables (CVs) generated from this analysis explained the major sources of temporally consistent covariance in the data. This enabled data-driven time series that were consistent across subjects to be discovered in a low-dimensional subspace. Each CV has an associated eigenvalue that is a measure of the ratio of the between-class (across time) over within-class (across subjects) variance or temporal signal-to-noise-ratio (SNR) of that variable. The % variance accounted for (VAF) by each CV can be calculated as the eigenvalue divided by the sum of the eigenvalues over all CVs. The spatial regions associated with the CV are expressed through canonical eigenimages (CEs). In the NPAIRS framework, the data are split into two halves each of which is subjected to the same analysis (aCVA in this case). The independent eigenimages from each split-half dataset are z-normalised (i.e., mean subtracted and divided by standard deviation), and a scatter plot of all paired voxel values is summarised by a reproducibility correlation coefficient, r. Following a PCA of the scatter plot, a principle signal axis ($e_{SIGNAL} = (1 + r)$), and an uncorrelated minor noise axis ($(e_{NOISE} = (1 - r))$ are defined. All voxel pairs are projected onto the principal signal axis and their values normalised by the standard deviation ($SD = \sqrt{e_{NOISE}}$) of the minor noise axis. This normalisation by the uncorrelated noise SD produces an approximately z-scaled SPM from the principal axis projection of each split-half pair. The final z-scored SPM is the consensus average of the z-scored voxels from each of the (50) split-half resamplings. This consensus averaging has two effects: (1) it produces a better approximation to a z-score scale by the central limit theorem, and (2) the result is robust to heterogeneity in the original sample (a strongly mismatched pair of split-half data sets tends to produce a larger variance on the uncorrelated minor noise axis and with rescaling of the signal axis this downweights the resulting approximate z-scores so they contribute less to the consensus average) (Strother et al., 2002). The median reproducibility value for each set of fifty splits was maximised to optimise both the number of principal components (PC) used as a basis for the aCVA, as well as to decide the best analysis pipeline for each group.

Mixed effects analysis

The data for each subject and each preprocessing pipeline were also analysed using the general linear model (mGLM) in AFNI. The task regressor was a twelve epoch boxcar function convolved with the AFNI canonical gamma function. The t-statistic and beta-weight map were calculated for each subject, then a two-level mixed effects group analysis was performed to calculate the mean group map for each processing pipeline.

Regions of interest (ROI)

The fusiform gyrus mask was created from outlines manually drawn on each subject's MRI and included the overlap of 95% of the individual fusiforms from the group of adults and a matched number of children selected from the low motion group across the 4- to 8-year age range. The individual fusiform outlines were delineated medially by the collateral sulcus and laterally by the occipital–temporal sulcus. The anterior boundary was the emergence of the gyrus around the mammillary bodies, and the posterior boundary was the parietal-occipital sulcus (Naidich et al., 1987). The primary visual cortex was taken to encompass the gyri immediately above and below the calcarine sulcus, with the anterior boundary being the parieto-occipital sulcus. The white matter mask was calculated from the

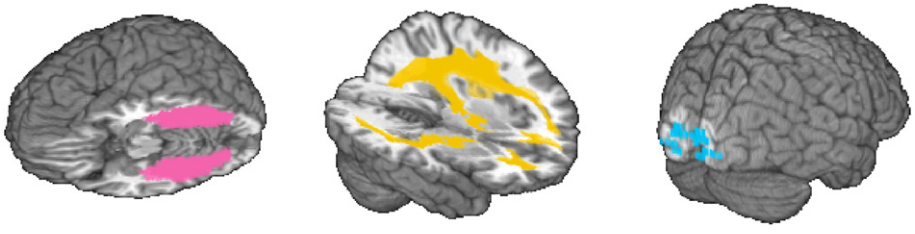

Fig. 3. Regions of interest: left: fusiform gyri, middle: 90% probability white matter, right: primary visual cortex.

ICBM 452 probabilistic white matter atlas at a threshold of 90% probability of white matter. The ROIs are depicted in Fig. 3.

Results

The motion characteristics of the adult and child groups are presented below, followed by the optimisation of the reproducibility metric that was used in the analysis pipeline summary. The effect of the motion correction (MC) and motion parameter regression (MPR) pipelines were then compared in the univariate GLM analysis as this is the most common method of analysis for fMRI data. Finally, a comparison of the multivariate aCVA with the GLM was made. In both these last analyses, voxelwise plots were employed to view the effect on the entire brain with the fusiform and primary visual cortex regions highlighted as they were regions that are predicted to respond to the stimuli with white matter voxels shown to indicate baseline response.

Motion

Shown in Fig. 4 are box-whisker plots of the distribution of values for maximum excursion, root mean square (RMS), and stimulus-correlated motion parameters for each direction and each group. The box is bounded by the first and third quartiles. The line (whiskers) stretches from the minimum and maximum data values with the points on the line indicating the values that lie outside the first and third quartiles. The stimulus correlated movement was calculated by computing the correlation coefficient between the estimated motion parameter time series and a boxcar (square wave) train of twelve epochs (periods) representing all face conditions, convolved with the canonical AFNI haemodynamic response.

Between group statistics were calculated to compare the adults with the low motion children and between the LM and HM children, and were calculated using a Wilcoxon rank-sum test. The motion differences between adults and the LM children were significant at $p<0.05$ for the directions of I/S (maximum excursion, $p=0.01$, RMS $p=0.04$) and A/P (RMS, $p=0.04$) translations and its associated rotation (yaw, RMS, $p=0.03$) but not at $p<0.01$. There were no substantial differences between adults and LM children in stimulus correlated motion. There were significant ($p<0.05$) differences between the HM and LM children for all types of motion except stimulus correlated motion. The R/L translations and roll rotations were the only directions of motion that had insignificant differences between the two groups of children.

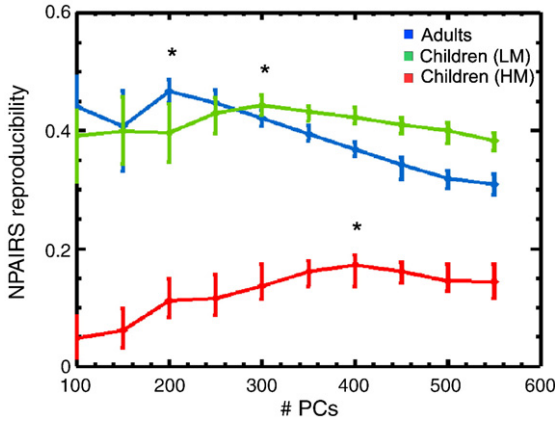

Fig. 5. Optimisation of the number of PCs for representative groups using the NPAIRS reproducibility metric for the MPR + Dtr pipeline from fifty split-half resampling values, with maxima indicated by an asterisk (*). The bars represent the distribution range of the split-half reproducibility values.

Optimisation of number of principal components

The number of principal components used for the aCVA split-half basis was optimised by choosing the value of q which produced the

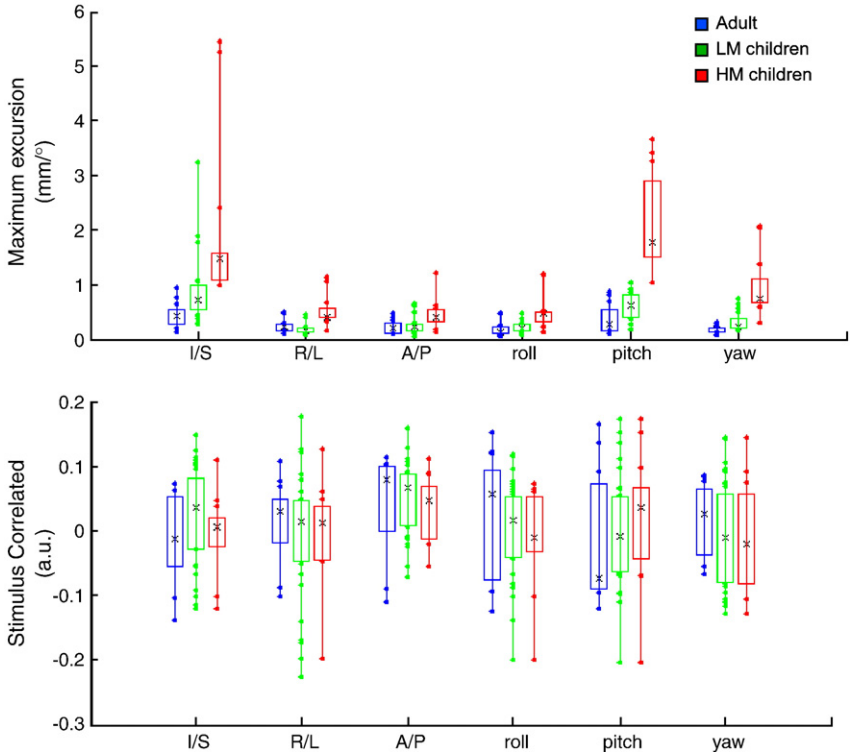

Fig. 4. Distribution of motion values over all directions for adults and both child groups for maximum excursion (top) and task correlated motion (bottom). The boxes are bounded by the first and third quartiles, the middle 50% of the data. The lines (whiskers) stretch from the minimum and maximum data values with the points on the line indicating the values that lie outside the first and third quartiles. The *x* indicates the median value.

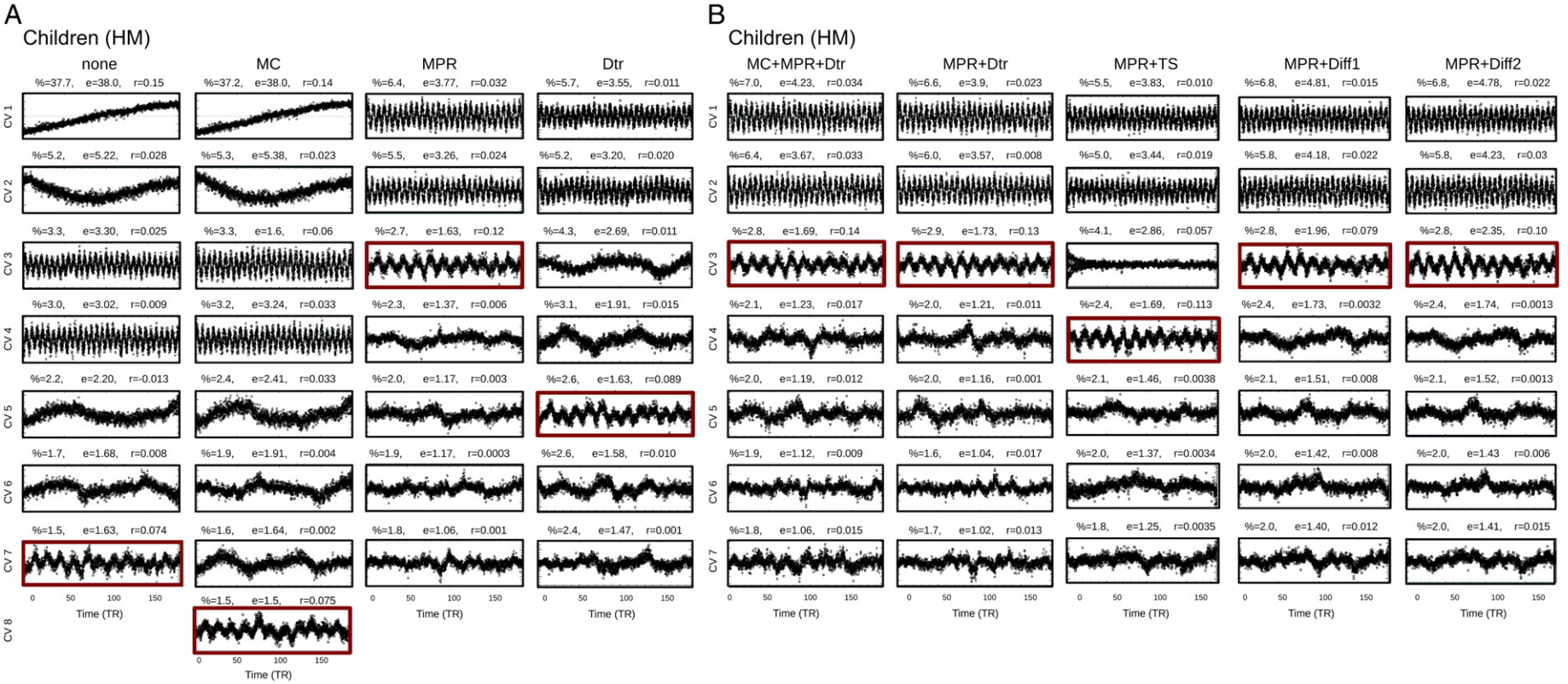

Fig. 6. (A) Effect of the preprocessing pipelines on the aCVA components for the HM children group along with their corresponding VAF (%), eigenvalues (e), and reproducibility (r) values. The task component is highlighted in the red box. Each timepoint on the horizontal axis of each panel is a scan with the interval between scans TR = 2 s. The vertical axis of each panel is unlabeled as the units are arbitrary. (B) Effect of five preprocessing pipelines on the aCVA components for the HM children group along with their corresponding VAF (%), eigenvalues (e), and reproducibility (r) values. The task component is highlighted in the red box. Each timepoint on the horizontal axis of each panel is a scan with the interval between scans TR=2 s. The vertical axis of each panel is unlabeled as the units are arbitrary.

first maximum of the median reproducibility value of the task component over the range of $q = 50$–550 components, where q was incremented in steps of fifty (eleven separate analyses). This procedure was performed separately for each group and pipeline. The task component was identified as the canonical time series that maximised the correlation with the block-design reference function for each NPAIRS analysis at every value of q. Fig. 5 shows a representative plot for the MPR + DTR pipeline for the task component reproducibility over the entire analysis range. Note that the profile for each group is somewhat different with adults peaking at a lower number of PCs (200) followed by the LM children (300), and finally by the HM (400) children. This trend was consistent over all pipelines and suggests that a larger number of PCs were required to capture the residual variance in the child data, particularly for more complex motion. All of the following aCVA properties are presented on their optimised PC basis.

aCVA components

The effect of each of the nine pipelines on the signal and noise structure of the data for the high motion group is shown in Fig. 6A and 6B for 7–8 aCVA components along with corresponding VAF (%), reproducibility (r), and eigenvalues (e) values (i.e. temporal SNR). Not surprisingly, every processing step increased the reproducibility and temporal SNR of the task component as compared to those of the raw unprocessed data. However, motion correction only slightly affected the components (column 2, Fig. 6A); particularly the linear trend and quadratic components which are likely motion residuals. Detrending did remove the prominent first and second order trends but left prominent higher order components (i.e. components 3 and 4, 4th column , Fig. 6A. Motion parameter regression successfully removed all significant noise components leaving the task components improved in both reproducibility and temporal SNR (column 3 Fig. 6A). The best combinations of VAF, eigenvalue and reproducibility are obtained for processing combinations of MC+MPR+Dtr and MPR+Dtr in Fig. 6B (columns 1 and 2) with the third, task-related component. The first two components remaining after MPR, or detrending (see the first two CVs in columns 3 and 4 of Figs. 6A and all columns in Fig. 6B still accounted for an amount of variance (VAF ~ 6%), which was greater than that of the task component (VAF ~ 2.6% – 2.7%). Both components 1 and 2 were at a frequency of 0.08 Hz, one with an even and the other with an odd phase. The presence of these components most likely indicates noise contamination of the time series. However, the reproducibility of these first two components was very low in all groups suggesting that they were spatially inconsistent. The reproducibility of the task component was much greater than any other components suggesting it was the strongest spatio-temporal pattern shared among all subjects. The above observations apply to all three groups.

Effect of motion correction (MC) and parameter regression (MPR) on the mGLM

There was extensive loss of high signal values in the high motion children, which was not substantially improved by the motion parameter regression preprocessing. Voxels in the white matter should be largely unaffected by the task and are indicative of an approximate lower bound on the noise found in all voxels. In Fig. 7, the white matter voxels (orange) illustrate that there was a positive systematic shift of values in the MC pipeline (x-axis) with respect to MPR (y-axis) in both the adults and LM children; however, in the HM children this shift was not evident. The values in fusiform (pink) showed a trend for slightly larger values, a shift to the right of the line of identity, using MC in the adults and LM children but not in the HM children. This trend was not clearly seen in the visual cortex (blue) giving the appearance of a clockwise rotation about the negative values for the adults and LM children, which was reversed in the HM children (i.e. a slope slightly greater than unity indicating larger values with MPR).

Analysis pipeline summary for the aCVA

Fig. 8 summarises the changes in reproducibility, eigenvalue (temporal SNR) and VAF as a function of each of the analysis pipelines for each group. The adult group's task component reproducibility was mildly affected by the different processing choices while the children's groups were strongly affected by pipeline choice. It should be noted that there was a larger (almost twice the) number of children in the LM group (22) compared to adults (13), and that a larger number of subjects should cause an increase in the reproducibility value (Strother et al., 1997), indicating a less consistent spatial signal contributed per child. The temporal SNRs of the adult components were greater than both child groups which also indicated that there was a less reliable temporal signal per child.

The spread of the reproducibility values across fifty splits (the bars in the top left graph of Fig. 8) also indicated greater group homogeneity in the adults compared to all groups of children regardless of the preprocessing pipeline. MPR pipelines also reduced the range of reproducibility values by about a half in the LM and HM children and adults compared to MC or no preprocessing, further indicating their importance. For all groups, MC alone produced less temporal SNR than MPR alone, and in most groups, SNRs from (MPR + DTR) were marginally better than from (MC + MPR + DTR). Motion had a large impact on reproducibility, or spatial SNR, in the HM children's group regardless of the pipeline used. Since none of the pipelines could recover the severe motion artifacts, the reproducibility of the analyses containing all children was lower than that of the LM children. The LM and HM children had similar temporal task SNRs (eigenvalues, Fig. 8)

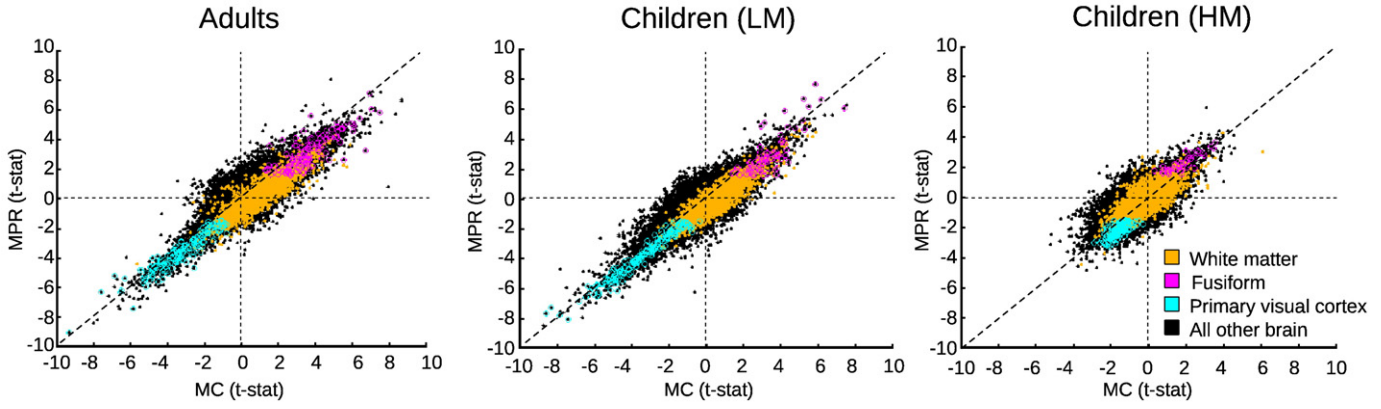

Fig. 7. Voxel-wise plots of the MC (x-axis) vs. MPR (y-axis) pipeline results for the GLM analysis for the adults and LM and HM children. Each point in the centre plot represents the same voxel in both maps, the pink represents voxels in the fusiform, blue the visual cortex and the orange are white matter (90% probability). The line of identity is also plotted.

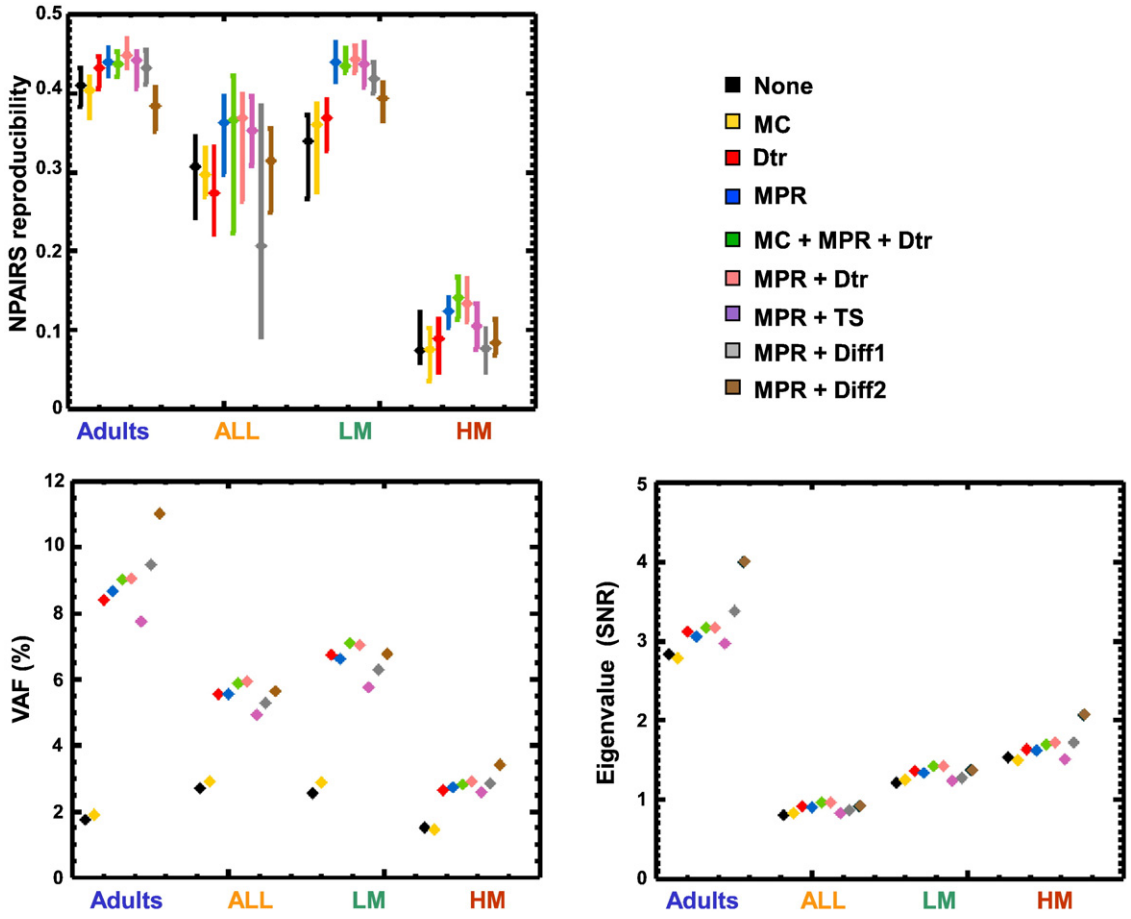

Fig. 8. Reproducibility, VAF and eigenvalues for the task component from the aCVA analysis of each of the pipelines (none, MC, MPR, MC + MPR + DTR, MPR + DTR) for each group: Adults (13 subjects), All children (33 subjects), LM children (22 subjects), HM children (11 subjects). The error bars in the top left graph represent the distribution range of the split-half reproducibility values.

despite greatly different reproducibility values. The VAF values further indicated that motion correction makes only minor reductions in the temporal noise structure of the data. Although the total amount of variance explained by the CVs varied somewhat with processing pipeline choice, it was suggestive that motion residuals take up larger and larger fractions of the data for the children compared to the adults, with the HM group predictably being most affected.

Comparison of the univariate (mGLM) and multivariate (aCVA) models

A comparison of the mGLM map and the aCVA task spatial component is shown in Fig. 9 for the adults and high and low motion child groups on axial brain slices for the aCVA (right) and mGLM (left). The *t*-statistic score of the mGLM was histogram-equalised to match the range of the aCVA *z*-score for this comparison (Hansen et al., 2001). This process was required as the *t*-statistic and *z*-score values map to different *p*-values. Histogram-equalization is similar to image contrast adjustment but operates on a *p*-value associated with the statistic in the image. When the *p*-value associated with the *t*-statistic is obtained and the corresponding *z*-score is found, this generates a mapping of *t*-statistics to *z*-scores that preserves the voxels' *p*-value rankings and is then applied to the image (Hansen et al., 2001).

The fusiform values were larger using the aCVA (pink dots in scatter plots), particularly for adults and LM children (Fig. 9). However, the visual areas (blue dots in scatter plots) showed a large spread with a split between which model yielded higher values, although in general, this may be due to different noise models as the values appear roughly evenly distributed about the line of identity. For all groups, white matter voxels indicated that the mGLM had a greater noise per voxel than the aCVA (i.e. elliptical shape of white matter scatter plot values with major axis slope greater than one). The high motion children showed extensive loss in signal values for both the aCVA and mGLM with a corresponding loss of response locations above the set threshold of $z = 2$.

In terms of the intensity values, there was good agreement between the maps with the aCVA generally having larger volumes of higher intensity than the mGLM. The extent of spatial overlap between the two models can be better seen in Fig. 10. There was extensive agreement in the areas of the fusiform and the visual cortex that typically respond to faces for all groups, with the aCVA having larger response volumes in these areas. Large differences between the models were found primarily in the frontal lobes and in the superior regions of the brain, where the mGLM had many small areas that were above threshold but the aCVA did not. In the adults, regions that were specific to the mGLM were the left anterior cingulate, right middle frontal gyrus, right inferior frontal gyrus, posterior cingular gyrus and bilateral precuneus. All these regions are implicated in the extended model for emotional face processing (Gobbini and Haxby, 2006). There was also an asymmetry seen in the aCVA intensity values in the adults (see Fig. 9, where the left fusiform shows larger values than the right, which was not seen in the mGLM). In the low motion children, there was an asymmetry in the response regions with the aCVA, whereas the mGLM indicated more bilateral responses in regions of the middle frontal gyrus, middle temporal gyrus and superior temporal gyrus. In the high motion

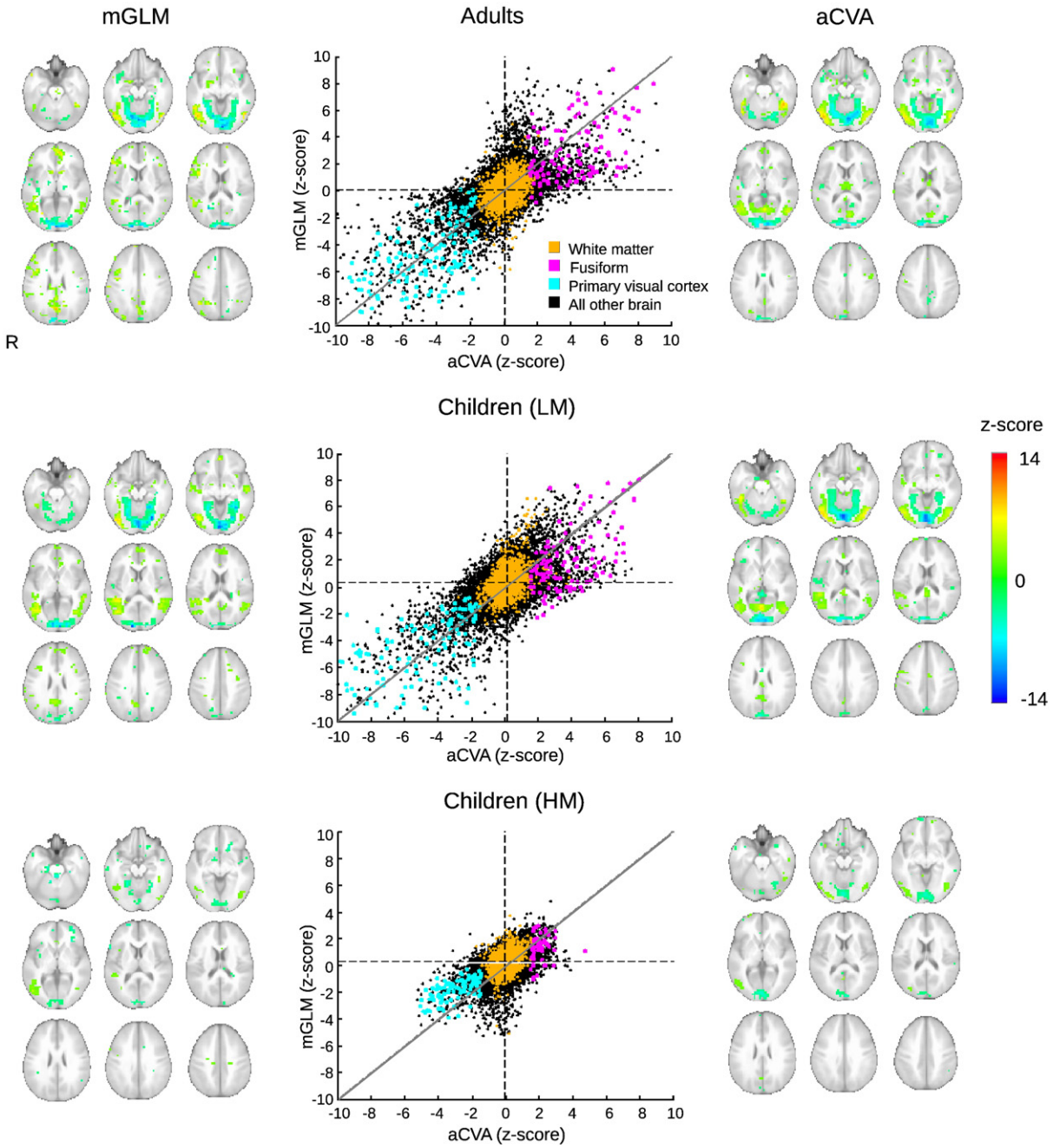

Fig. 9. Spatial components of the mGLM and aCVA shown on axial brain slices on the ICBM 452 atlas for each group (columns 1 and 3) at a threshold of $z = 2$. Images are shown in the radiological convention where left is on the right. Voxelwise plots in the middle column compare the z-scores of the two analyses. Each point in the plot represents the same voxel in both maps, the pink represents voxels in the fusiform, blue the visual cortex and the orange are white matter (90% probability).

children, the mGLM showed responses in the left middle frontal gyrus, and left anterior cingulate. Additionally, a large region anterior to the visual cortex (cuneus) had a large positive response in the aCVA that was not evident in the mGLM in either the adults and or low motion children.

Discussion

aCVA components

It is clear that post-processing motion algorithms can only account for some of the signal artifacts due to motion. In this study we have illustrated a way of quantifying whether the effects of motion in a dataset have been removed to a satisfactory level. The NPAIRS reproducibility metric is able to quantify the changes to the task spatial pattern, from different preprocessing patterns for group maps. The use of the maximum extent of motion was a good criterion to separate out motion that was unable to be corrected by the motion algorithms in the children's groups. The LM children's group had reproducibility values that became comparable with that of the adults after correction with MPR. The inclusion of the HM children with uncorrectable motion as seen by the low reproducibility values, decreased the reproducibility values of the entire group of children when they were analysed together. The eigenvalues of the LM child group suggested that there was a difference in the temporal SNR of the task component. This could either arise from further residual motion corruption or from a difference in physiological noise contributions, as has been shown by Thomason

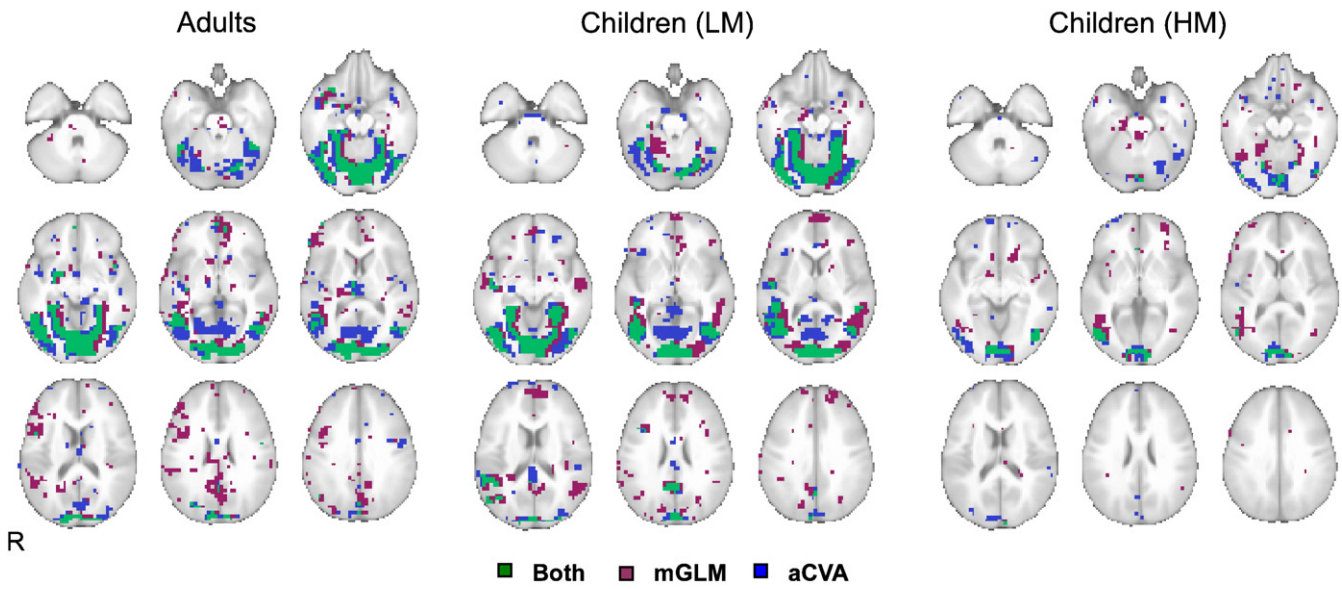

Fig. 10. Maps for each of the adults and child (HM, LM) groups illustrating the overlap (green) for the mGLM (burgundy) and aCVA (blue).

et al. (2005). The adult dataset values appear to be reasonably insensitive to pipeline choice in terms of the reproducibility of the task component.

Despite exploring several possibilities to explain the higher frequency components (first and second rows in Fig. 6B,) that, together, explain more signal variance than the task component, their cause remains unclear. In further analyses, slice timing correction was applied as a preprocessing step to account for the fact that the functional images were acquired in an interleaved fashion; however, these components were not removed (Fig. 6B column three and the reproducibility , eigenvalue and VAF of the task component was), and the reproducibility of the task component was reduced relative to an MPR + DTR pipeline (column 2). The first and second derivative of the task were added as regression components to the preprocessing pipeline to account for any latency differences in the haemodynamic response over the brain (Henson et al., 2002), and again the higher components persisted in the aCVA (columns 4 and 5, Fig. 6B). Increased eigenvalues and VAF in the adults compared to the children suggest that they are a function of an external effect (since the task component also increases) as opposed to an effect of motion. An investigation of the power spectrum of the time series data from individual subjects suggests some brain areas, like the precuneus for example, show a strong higher frequency component which is sometimes larger than the task frequency in non-core face processing regions. Investigations of this component suggest that it is caused by a technical issue in the spiral-in out sequence and can be removed with a notch filter. These results demonstrate a clear advantage of aCVA over GLM and potentially other analysis approaches, for fMRI data analysis by not only discovering this effect, but simultaneously separating it from the task component through the orthogonality constraint on the aCVA components. Note that these may be temporally nonorthogonal components in the original PCA subspace on which the aCVA is calculated because aCVA involves a general eigendecomposition of the data matrix defined by (between-class covariance)/(within-class covariance).

Pipeline choices

In the aCVA, the MPR pipelines have the greatest effect on improving the reproducibility of datasets while MC has little effect either alone or in conjunction with MPR (see Fig. 8). The components of the HM children illustrated in Fig. 6A showed that MPR (column 3) is able to reduce the noise components contained in the raw data (column 1 while improving %VAF and reproducibility). However, MC (column 2) is not able to completely remove even the linear trend found in the data, which is the first component of a PCA of the motion components.

In Fig. 7 it is clear that MC also adds a positive bias to the data with respect to the MPR processing pipeline in the mGLM. This speaks to the induction of artifacts from the intensity-based motion correction algorithm (Johnstone et al., 2006; Freire and Lee, 2001). Freire et al. (2001) suggest that more robust mutual information-based algorithms would not be affected by the task-induced intensity changes and render a better correction; however, this is not included in standard packages. Johnstone et al. (2006) also suggested that the MC and MPR spatial maps can be used to separate areas of artifact from regions of activation by observing the change in intensity value. However, since MC seems by itself to induce an overall increase in *t*-values, their suggested algorithm for detecting the false activations cannot be used in this case (since decreases in *t*-values are expected for motion corrupted voxels after MC in their model). Based on these results, MPR should be used in the preprocessing pipeline to improve the spatial SNR of the data. Motion correction can also be added to the preprocessing steps; however, on its own it does not greatly improve the spatial or time series SNR of the data, even for relatively low motion child or adult subjects. Varying the amount of spatial smoothing is another potential way to improve spatial SNR of the datasets (Strother et al., 2004). Preliminary investigations of different sized smoothing kernels on this dataset (0, 4, 8 mm) indicated that some improvement in spatial smoothing may be obtained by optimising kernel size. However, it does not affect the relative improvement of MPR over MC.

Comparison of the univariate (mGLM) and multivariate (aCVA) models

Since the true activation map is not known, the differences between the mGLM and aCVA models are difficult to quantify. Overall, the consensus map, or average map between the two can been shown to be closer to the "true" map (Hansen et al., 2001). That said, there is considerable agreement between the two maps in the expected regions of response: the fusiform, primary visual cortex, superior temporal gyrus, insula and medial temporal lobe (see Fig. 10). Thus, the strongest areas of activation are well represented

for all groups in both models. There are some notable deviations between the two models particularly in the frontal and superior regions and the left medial temporal lobe where the mGLM appears to have greater sensitivity. It is possible either that these regions have a different temporal signature (haemodynamic response difference) placing them in an altogether different components in the aCVA, which is not model driven, with significantly lower reproducibility and eigenvalues or that they are related to susceptibility artifacts. The reduced reproducibility of the noise components as compared to that of the task suggests that these regions of activation are not stable across subjects, meaning that the activations are more likely related to susceptibility artifacts than to actual task activation. Conservative conclusions about the regions responding to this task should then be taken as those that are common between the two models.

An additional difference between the two models is the way in which they compute their task effect maps. The mGLM calculates the estimated model parameter (for the condition of faces) on a voxel-wise basis using local noise estimates whereas the aCVA provides inferences about distributed brain responses. Furthermore, the mGLM is sensitive to both changes in local signal and noise values but the aCVA is more sensitive to consistent signal changes across the entire brain volume as the *z*-scored canonical eigenimages are produced using a robust global (i.e., pooled over all voxels) noise estimate. The aCVA explicitly models task and confounding effects (such as global physiological changes) at both the voxel level as well as the interaction among other voxels. As such, it is possible that the aCVA is a more robust model in terms of comparisons across varied subject populations such as adults and children.

Conclusions

Motion parameter regression is a benefit to all datasets whereas motion correction on its own does not have as much impact and may be detrimental. The maximal excursion from each direction is a good metric to choose the amount of motion to be allowed in a dataset. The low motion child group subjected to MPR had reproducibility values at par with those of the adult group, but needed almost twice as many subjects to achieve this result, indicating fundamentally weaker responses per child as compared to adults. The aCVA showed greater sensitivity to the task response pattern than the mGLM in the expected face processing regions, although the mGLM showed more responses in some other areas. In addition, aCVA has advantages over mGLM, and potentially other analysis approaches for fMRI data analysis by not only detecting strong artifactual signals, but simultaneously separating them from the task component through the orthogonality constraint on the aCVA components. The use of an average or an overlap map of response areas between both models may be useful identifying robust regions that consistently respond to the task regardless of very different modelling assumptions.

Acknowledgments

The research reviewed in this paper was supported in part by CIHR grants MOP-84483 (SCS) and MOP-81161 (MJT). JWE was supported by a Research Training Competition (RESTRACOMP) graduate scholarship at The Hospital for Sick Children, Toronto. SCS gratefully acknowledges support of the Heart & Stroke Foundation of Ontario through the Centre for Stroke Recovery.

Appendix. The NPAIRS framework

In the NPAIRS framework, the data matrix S of v voxels by s scans ($s = n \times t$, where n is the number of subjects and t the number of timepoints) is first split into two subjects groups $S = [S_1, S_2]$, and this splitting process is repeated 50 times. These matrices are large, and it would be computationally demanding (or impossible) to retain all voxels for all iterations in the process. Fortunately, it is possible to reduce the number of components (or dimensionality) of these matrices by first approximating them using a basis obtained using PCA before calculating the CVA, which improves computation time. The NPAIRS analysis takes about 2–5 h to complete on a 2.66 GHz processor (with 4 GB of memory) for the group of 13 adults given as an example in Methods for 100–400 CVA components, respectively, out of a possible 748.

The NPAIRS analysis procedure can be itemised as follows (more details on the process can be found in (Strother et al., 2002)):

(1) Compute the full PCA (1) on the data matrix S of all scans: $S = U\Lambda V^t$, which can then be represented by the reduced basis set as $X^* = (U^*)^t S = \Lambda^*(V^*)^t$. In the current analysis the top 30% of the PCA components were retained to form the matrix X^* (size $n \times t \times 30\%$ by $n \times t$.)
(2) Partition X^* into two independent split-half groups across the subjects yielding matrices X_i (size $n \times t \times 30\%$ by $n_i \times t$) for $i = 1,2$.
(3) Compute a smaller PCA (2) on X_1 and X_2 separately and retain q components. Considering X_1 as an example, $X_1 = Y_1\Lambda_1 V_1^t$, $X_1^* = (Y_1^*)^t X_1 = \Lambda_1^*(V_1^*)^t$, which has size q by $n_1 \times t$.
(4) Perform CVA on X_1^* and X_2^*. Considering the X_1^* matrix, there is one scan, x_i, per column and $x_{g,i}$ is the scan in the ith column that has label $g = 1 \ldots G$, where G is the number of groups. The $q \times q$ within-group, W_1, and between-group, B_1, variance-covariance matrices are given by

$$W_1 = \sum_{i,g} \left(x_{g,i} - \bar{x}_g\right)\left(x_{g,i} - \bar{x}_g\right)^t$$

$$B_1 = \sum_{g} n_g \left(\bar{x}_g - \bar{x}\right)\left(\bar{x}_g - \bar{x}\right)^t$$

where the group of scans labelled g has a mean given by $\bar{x}_g = 1/n_g \sum_{g,j} \bar{x}_{g,j}$ and where $\bar{x} = 1/n \sum_i^n \bar{x}_i$ is the average over all scans in the split-half.
The CVA seeks to maximise the ratio

$$\frac{c_{1,1}^t B c_{1,1}}{c_{1,1}^t W c_{1,1}}$$

to identify the first component. Finding the canonical variates results in calculating the eigenvalues and eigenvectors of the matrix $W_1^{-1} B_1$ and the canonical variable time series ($n_1 \times t$ scans) can then be calculated from $Z_{1,1} = X_1^{*t} c_{1,1}$, for the first component for X_1^*. The eigenvalues represent the amount of variance represented by the canonical component, and the total amount of variance for $W_1^{-1} B_1$ is the sum of all the eigenvalues.
Canonical eigenimages can be calculated by

$$C_{1,i} = U^* Y_1^* c_{1,i}$$

where Y_1^* is the matrix from the PCA (2) in step 3 above, U^* is from PCA (1) in step 1 above, and $i = 1 \ldots q$.
(5) Compute reproducibility between the cannonical eigenimages ($C_{1,i}$ and $C_{2,i}$, $i = 1 \ldots q$) from the split-halves (see note below on component ordering).
(6) Repeat steps 2–5 for each split-half group, i.e. 50 times.

The number of components, q, retained in the second PCA is variable as it is unknown how many variance components actually compose the signal and noise space. To choose the optimal value for q this entire process was repeated for values of $q = 50$–550, in steps of 50, out of a possible of 748. The first value of q which maximised the

median reproducibility of the task component was selected as the optimal number of PCA components to retain for the CVA (which produces the q canonical variates).

The components of the aCVA are only defined up to a sign and their order may vary between pairs of split-half groups. In order to account for this, an aCVA decomposition is done on the entire dataset and the order of these components is used as a reference. The canonical variates from each split half are then compared to the reference set to get the maximal correlation (reordering of the components to have the same order as the reference) and a positive correlation (possible change of sign). This is a restricted version of the procrustes matching procedure described in Milan and Whittaker (1995).

References

Andersson, J.L., Hutton, C., Ashburner, J., Turner, R., Friston, K., 2001. Modeling geometric deformations in EPI time series. NeuroImage 13, 903–913.

Ardekani, B.A., Bachman, A.H., Helpern, J.A., Sep 2001. A quantitative comparison of motion detection algorithms in fMRI. Magn. Reson. Imaging 19 (7), 959–963.

Byars, A.W., Holland, S.K., Strawsburg, R.H., Bommer, W., Dunn, R.S., Schmithorst, V.J., Plante, E., Dec 2002. Practical aspects of conducting large-scale functional magnetic resonance imaging studies in children. J. Child Neurol. 17 (12), 885–890.

Cox, R.W., Jun 1996. Afni: software for analysis and visualization of functional magnetic resonance neuroimages. Comput. Biomed. Res. 29 (3), 162–173.

Freire, A., Lee, K., Dec 2001. Face recognition in 4- to 7-year-olds: processing of configural, featural, and paraphernalia information. J. Exp. Child Psychol. 80 (4), 347–371.

Freire, L., Mangin, J.F., Sep 2001. Motion correction algorithms may create spurious brain activations in the absence of subject motion. NeuroImage 14 (3), 709–722.

Friston, K.J., Price, C.J., Fletcher, P., Moore, C., Frackowiak, R.S., Dolan, R.J., Oct 1996. The trouble with cognitive subtraction. NeuroImage 4 (2), 97–104.

Gobbini, M.I., Haxby, J.V., Dec 2006. Neural response to the visual familiarity of faces. Brain Res. Bull. 71 (1-3), 76–82.

Grootoonk, S., Hutton, C., Ashburner, J., Howseman, A.M., Josephs, O., Rees, G., Friston, K.J., Turner, R., Jan 2000. Characterization and correction of interpolation effects in the realignment of fMRI time series. NeuroImage 11 (1), 49–57.

Hajnal, J.V., Myers, R., Oatridge, A., Schwieso, J.E., Young, I.R., Bydder, G.M., Mar 1994. Artifacts due to stimulus correlated motion in functional imaging of the brain. Magn. Reson. Med. 31 (3), 283–291.

Hansen, L.K., Nielsen, F.A., Strother, S.C., Lange, N., Jun 2001. Consensus inference in neuroimaging. NeuroImage 13 (6 Pt 1), 1212–1218.

Henson, R.N.A., Price, C.J., Rugg, M.D., Turner, R., Friston, K.J., Jan 2002. Detecting latency differences in event-related BOLD responses: application to words versus non-words and initial versus repeated face presentations. NeuroImage 15 (1), 83–97.

Jiang, A., Kennedy, D., Baker, J.R., Weisskoff, R.M., Tootell, R., Woods, R., Benson, R.R., Kwong, K.K., Brady, T.J., Rosen, B., Belliveau, J.W., 1995. Motion detection and correction in functional MR imaging. Hum. Brain Mapp. 3, 224–235.

Johnstone, T., Walsh, K.S.O., Greischar, L.L., Alexander, A.L., Fox, A.S., Davidson, R.J., Oakes, T.R., Oct 2006. Motion correction and the use of motion covariates in multiple-subject fMRI analysis. Hum. Brain Mapp. 27 (10), 779–788.

Kjems, U., Hansen, L.K., Anderson, J., Frutiger, S., Muley, S., Sidtis, J., Rottenberg, D., Strother, S.C., Apr 2002. The quantitative evaluation of functional neuroimaging experiments: mutual information learning curves. NeuroImage 15 (4), 772–786.

Kotsoni, E., Byrd, D., Casey, B.J., Jun 2006. Special considerations for functional magnetic resonance imaging of pediatric populations. J. Magn. Reson. Imaging 23 (6), 877–886.

Liow, J.S., Anderson, J.R., Strother, S., Jun 2000. Comparing reconstruction algorithms using a multi-variate analysis. IEEE Trans. Nucl. Sci. 47 (3, Part 3), 1136–1142.

Lund, T.E., Norgaard, M.D., Rostrup, E., Rowe, J.B., Paulson, O.B., Jul 2005. Motion or activity: their role in intra- and inter-subject variation in fMRI. Neuroimage 26 (3), 960–964.

Lund, T.E., Madsen, K.H., Sidaros, K., Luo, W.-L., Nichols, T.E., Jan 2006. Non-white noise in fMRI: does modelling have an impact. Neuroimage 29 (1), 54–66.

Milan, L., Whittaker, J., 1995. Application of the parametric bootstrap to models that incorporate a singular value decomposition. J. R. Stat. Soc., Ser. C Appl. Stat. 44, 31–49.

Naidich, T.P., Daniels, D.L., Haughton, V.M., Pech, P., Williams, A., Pojunas, K., Palacios, E., 1987. Hippocampal formation and related structures of the limbic lobe: anatomic-MR correlation. Part II. Sagittal sections. Radiology 162, 755–761.

Oakes, T.R., Johnstone, T., Walsh, K.S.O., Greischar, L.L., Alexander, A.L., Fox, A.S., Davidson, R.J., Nov 2005. Comparison of fMRI motion correction software tools. NeuroImage 28 (3), 529–543.

Seto, E., Sela, G., McIlroy, W.E., Black, S.E., Staines, W.R., Bronskill, M.J., McIntosh, A.R., Graham, S.J., Aug 2001. Quantifying head motion associated with motor tasks used in fMRI. NeuroImage 14 (2), 284–297.

Strother, S., Conte, S.L., Hansen, L.K., Anderson, J., Zhang, J., Pulapura, S., Rottenberg, D., 2004. Optimizing the fMRI data-processing pipeline using prediction and reproducibility performance metrics: I. a preliminary group analysis. Neuroimage 23 (Suppl 1), S196–S207.

Strother, S., Lange, N., Anderson, J., Schaper, K., Rehm, K., Hansen, L., Rottenberg, D., 1997. Activation pattern reproducibility: measuring the effects of group size and data analysis models. Hum. Brain Mapp. 5 (4), 312–316.

Strother, S.C., Anderson, J., Hansen, L.K., Kjems, U., Kustra, R., Sidtis, J., Frutiger, S., Muley, S., LaConte, S., Rottenberg, D., Apr 2002. The quantitative evaluation of functional neuroimaging experiments: the npairs data analysis framework. NeuroImage 15 (4), 747–771.

Szaflarski, J.P., Schmithorst, V.J., Altaye, M., Byars, A.W., Ret, J., Plante, E., Holland, S.K., May 2006. A longitudinal functional magnetic resonance imaging study of language development in children 5 to 11 years old. Ann. Neurol. 59 (5), 796–807.

Thesen, S., Heid, O., Mueller, E., Schad, L.R., 2000. Prospective acquisition correction for head motion with image-based tracking for real-time fMRI. Magn. Reson. Med. 44, 457–465.

Thacker, N.A., Burton, E., Lacey, A.J., Jackson, A., Aug 1999. The effects of motion on parametric fMRI analysis techniques. Physiol. Meas. 20 (3), 251–263.

Thomason, M.E., Burrows, B.E., Gabrieli, J.D.E., Glover, G.H., Apr 2005. Breath holding reveals differences in fMRI BOLD signal in children and adults. NeuroImage 25 (3), 824–837.

Yuan, W., Altaye, M., Ret, J., Schmithorst, V., Byars, A.W., Plante, E., Holland, S.K., Jul 2008. Quantification of head motion in children during various fMRI language tasks. Hum. Brain Mapp. 30, 1481–1489.

Zaitsev, M., Dold, C., Sakas, G., Heinnig, J., Speck, O., 2006. Magnetic resonance imaging of freely moving objects: prospective real-time motion correction using an external optical motion tracking system. 31(3) 1038–1050.