Published in final edited form as:

J Neurosci Methods. 2010 June 15; 189(2): 233–245. doi:10.1016/j.jneumeth.2010.03.028.

Visual inspection of independent components: Defining a procedure for artifact removal from fMRI data

Robert E. Kelly Jr.^{a,*}, George S. Alexopoulos^a, Zhishun Wang, Ph.D.^e, Faith M. Gunning^a, Christopher F. Murphy^a, Sarah Shizuko Morimoto^a, Dora Kanellopoulos^a, Zhiru Jia^a, Kelvin O. Lim^d, and Matthew J. Hoptman^{b,c}

^aWeill Cornell Institute of Geriatric Psychiatry, Weill Cornell Medical College

^bDivision of Clinical Research, Nathan S. Kline Institute for Psychiatric Research

^cDepartment of Psychiatry, New York University School of Medicine

^dDepartment of Psychiatry, University of Minnesota

^eThe MRI Unit and The Division of Child and Adolescent Psychiatry, Columbia University and New York State Psychiatric Institute (NYSPI), New York, New York

Abstract

Artifacts in fMRI data, primarily those related to motion and physiological sources, negatively impact the functional signal-to-noise ratio in fMRI studies, even after conventional fMRI preprocessing. Independent component analysis' demonstrated capacity to separate sources of neural signal, structured noise, and random noise into separate components might be utilized in improved procedures to remove artifacts from fMRI data. Such procedures require a method for labeling independent components (ICs) as representing artifacts to be removed or neural signals of interest to be spared. Visual inspection is often considered an accurate method for such labeling as well as a standard to which automated labeling methods are compared. However, detailed descriptions of methods for visual inspection of ICs are lacking in the literature. Here we describe the details of, and the rationale for, an operationalized fMRI data denoising procedure that involves visual inspection of ICs (96% inter-rater agreement). We estimate that dozens of subjects/

© 2010 Elsevier B.V. All rights reserved.

*Corresponding author: Weill Cornell Institute of Geriatric Psychiatry, 21 Bloomingdale Road, White Plains, N.Y. 10605, USA. Telephone +1 (914) 997-4028. Fax +1 (914) 682-6979. rek2005@med.cornell.edu .

George S. Alexopoulos, M.D., Weill Cornell Institute of Geriatric Psychiatry, 21 Bloomingdale Road, White Plains, New York 10605, USA

Zhishun Wang, Ph.D., Columbia University and New York State Psychiatric Institute, 1051 Riverside Drive, New York, NY 10032, USA

Faith M. Gunning, Ph.D., Weill Cornell Institute of Geriatric Psychiatry, 21 Bloomingdale Road, White Plains, New York 10605, USA

Christopher F. Murphy, Ph.D., Weill Cornell Institute of Geriatric Psychiatry, 21 Bloomingdale Road, White Plains, New York 10605, USA

Sarah Shizuko Morimoto, Psy.D., Weill Cornell Institute of Geriatric Psychiatry, 21 Bloomingdale Road, White Plains, New York 10605, USA

Dora Kanellopoulos, Weill Cornell Institute of Geriatric Psychiatry, 21 Bloomingdale Road, White Plains, New York 10605, USA

Zhiru Jia, Ph.D., Weill Cornell Institute of Geriatric Psychiatry, 21 Bloomingdale Road, White Plains, New York 10605, USA

Kelvin O. Lim, M.D., Department of Psychiatry, University of Minnesota, 717 Delaware Street SE, Minneapolis, Minnesota 55414, USA

Matthew J. Hoptman, Ph.D, Division of Clinical Research, Nathan S. Kline Institute for Psychiatric Research, 140 Old Orangeburg Road, Orangeburg, New York 10962, USA

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

sessions can be processed within a few hours using the described method of visual inspection. Our hope is that continued scientific discussion of and testing of visual inspection methods will lead to the development of improved, cost-effective fMRI denoising procedures.

Keywords

fMRI; independent component analysis (ICA); denoising; visual inspection; artifacts; structured noise; independent component (IC) labeling

Introduction

Structured noise from numerous sources (Biswal et al., 1996; Friston et al., 1996; Chen and Zhu, 1997; Birn et al., 1998; Dagli et al., 1999; Glover et al., 2000; Raj et al., 2001; Windischberger et al., 2002; Beauchamp, 2003) and random (Gaussian) noise compromise the functional signal-to-noise ratio and the sensitivity and specificity of analytical results derived from brain blood-oxygenation-level dependent (BOLD) functional magnetic resonance imaging (fMRI) data (Thomas et al., 2002; Huettel et al., 2004a; Raichle and Snyder, 2007). Some structured noise remains in the data after traditional corrections are applied, such as slice-timing correction, rigid-body motion correction, high-pass temporal filtering, and spatial smoothing (Hu et al., 1995; Grooten et al., 2000; Andersson et al., 2001; Raj et al., 2001; Birn et al., 2004).

Independent component analysis (ICA) has been used in denoising procedures to improve the sensitivity and specificity of results derived from fMRI data, beyond those obtained with traditional preprocessing (Stone et al., 2002; Thomas et al., 2002; Kochiyama et al., 2005; McKeown et al., 2005; Zou et al., 2009). ICA produces spatiotemporal components (pairs of time courses and spatial maps) through linear decomposition of fMRI data (McKeown et al., 1998). ICA denoising procedures have involved some method for determining (labeling) which independent components (ICs) represent noise (N-ICs) and which represent neural signals of interest (S-ICs). For some studies, the S-ICs have been the denoised end results of interest (Calhoun et al., 2001a; Moritz et al., 2003; Greicius et al., 2007; Stevens et al., 2007; Sui et al., 2009). For other studies, denoising of the fMRI data has been performed as an extension of data preprocessing by 1) filtering the N-ICs from the preprocessed data using the N-IC time courses as nuisance variables with linear regression (Zou et al., 2009); 2) reconstructing the fMRI data from the S-ICs alone (i.e., matrix multiplication of the S-IC time courses and spatial maps) (Thomas et al., 2002; Kochiyama et al., 2005; Perlberg et al., 2007; Tohka et al., 2008); or 3) projecting (least squares regression) the fMRI data into the linear subspace spanned by the S-ICs (McKeown, 2000; McKeown et al., 2005). The success of these denoising methods depends upon the accuracy of labeling the ICs, but potential complications in the labeling process are that some ICs appear to represent a synthesis of artifactual and neurally derived signals (McKeown et al., 1998; Thomas et al., 2002; Birn et al., 2008) and it is not clear in every case how ICs should be labeled.

A number of approaches for labeling ICs have been described. These approaches can be divided according to 1) whether ICA is performed on individual fMRI data runs (McKeown, 2000; Calhoun et al., 2001a; Thomas et al., 2002; Moritz et al., 2003; Kochiyama et al., 2005; McKeown et al., 2005; Greicius et al., 2007; Perlberg et al., 2007; Tohka et al., 2008) or is performed through a single, group ICA on all fMRI data runs together (Stevens et al., 2007; Sui et al., 2009; Zou et al., 2009); 2) whether the approaches are completely automated (McKeown, 2000; Thomas et al., 2002; Kochiyama et al., 2005; McKeown et al., 2005; Greicius et al., 2007; Perlberg et al., 2007; Stevens et al., 2007; Tohka et al., 2008; Sui et al., 2009) or are "manual" (McKeown et al., 1998; Calhoun et al., 2001a; Moritz et al.,

2003; Zou et al., 2009), requiring some element of visual inspection and human decision making; 3) whether the approaches are completely data driven (Calhoun et al., 2001a; Greicius et al., 2007; Perlberg et al., 2007; Stevens et al., 2007; Tohka et al., 2008; Sui et al., 2009; Zou et al., 2009) or require task-related temporal or spatial (brain areas affected) information (McKeown, 2000; Thomas et al., 2002; Moritz et al., 2003; Kochiyama et al., 2005; McKeown et al., 2005); and 4) whether the approaches are based on IC temporal (McKeown, 2000; Thomas et al., 2002; Kochiyama et al., 2005; McKeown et al., 2005; Greicius et al., 2007; Perlberg et al., 2007) or spatial (Calhoun et al., 2001a; Stevens et al., 2007; Sui et al., 2009; Zou et al., 2009) characteristics, or both (McKeown et al., 1998; Moritz et al., 2003; Tohka et al., 2008). Characteristics of IC time courses and their associated Fourier frequency spectrums that have been used to distinguish N-ICs from S-ICs include abrupt, large shifts (time course "spikes") (McKeown et al., 1998; Tohka et al., 2008); oscillating, "quasi-periodic" pattern (McKeown et al., 1998); similarity to white noise (Thomas et al., 2002; Tohka et al., 2008); similarity to time courses from regions of the brain where neural activity does not occur (ventricles and vasculature) (Perlberg et al., 2007); similarity to task-related activity (McKeown, 2000; Thomas et al., 2002; Moritz et al., 2003; Kochiyama et al., 2005; McKeown et al., 2005); heteroscedasticity in residuals from regressing IC time courses against a task-related time course (Kochiyama et al., 2005); and relative amount of power at frequencies considered typical for artifacts (Thomas et al., 2002; Greicius et al., 2007). IC spatial characteristics used for labeling include degree of association with cerebrospinal fluid, white matter, grey matter, and/or blood vessels (Stevens et al., 2007; Sui et al., 2009; Zou et al., 2009); extent of component variance in brain periphery (McKeown et al., 1998; Tohka et al., 2008); degree of clustering and degree of scattering of thresholded voxels in IC spatial maps (McKeown et al., 1998; Sui et al., 2009); and correspondence with constellations of brain regions known to perform particular functions (Calhoun et al., 2001a; Moritz et al., 2003).

Three recently reported methods for automated labeling of ICs (Perlberg et al., 2007; Tohka et al., 2008; Sui et al., 2009) were validated in part by comparing the results of automated labeling with those derived from visual inspection. These investigators considered visual inspection by "experts" sufficiently accurate to be used as "a gold standard to assess the quality of the automatic selection procedure" (Perlberg et al., 2007). However, no operationalized procedure for labeling ICs with visual inspection was described, and such descriptions are lacking in the literature. Perhaps the best description of visual inspection of ICs is found in McKeown et al. (1998), which gives examples of the appearance of spatial and temporal patterns in components that might suggest the presence of artifacts or random noise; but no detailed procedure or guidelines are provided for how to appropriately label components in every case and no data are provided concerning the reliability or accuracy of visual inspection as a method for identifying artifacts.

Detailed descriptions of procedures for visual inspection, ICA-based denoising (VIID) are needed to facilitate further exploration of the potential advantages and disadvantages of using such procedures for denoising fMRI data. For example, it might be useful to explore whether the high levels of accuracy reported for different automated methods, using visual inspection as the standard for comparison, would be as high if automated labeling results were compared to visual inspection results at a location with slightly different visual inspection traditions or philosophies. Such exploration would require a more detailed methodology description than "visual inspection was performed." Here, we provide an example of a detailed description of visual inspection of ICs as part of a procedure for denoising fMRI data.

Materials and methods

Overview

This procedure derives a set of ICs, each representing a separate portion of the total variance in the fMRI data, using spatial ICA. We assume that the sources of variance for each component are a mixture of neural signals of interest (NSI), neural signals of no interest (NSNI, e.g., activity related to a brain function or region not being studied), structured noise, and random noise. The primary goal of this procedure is to reduce noise in the fMRI data, while preserving as much NSI as possible. This goal is accomplished by selecting components considered to represent predominantly noise (N-ICs; all others are designated S-ICs), and “filtering” them from the fMRI data. Our rule of thumb is to eliminate components considered to represent less than 10% NSI, but this figure is arbitrary and can be adjusted up or down depending upon how much NSI one is willing to sacrifice in order to remove noise. Removal of NSNI is optional, by treating NSNI as noise, but for the purposes of this demonstration all neurally-derived signals are considered to be NSI.

The decision to label a component N-IC (or S-IC) is based primarily upon visual inspection of the thresholded component spatial map. Components are labeled N-IC when they show predominantly (90% or more) “activation” or “deactivation” in peripheral areas or in a spotty or speckled pattern, seemingly scattered at random over a large section (roughly $\frac{1}{4}$ or more) of the brain without regard for functional-anatomical boundaries. Conversely, components are labeled S-IC when at least 10% of the activations/deactivations are found in small (roughly 25 voxels of resolution $4 \times 4 \times 4$ mm) to larger grey-matter clusters localized to small regions of the brain rather than being diffusely scattered across large regions or found in the periphery. In cases where there is doubt about whether more than 90% of a component represents noise, the general rule is to label it S-IC; but if it also seems likely that the component represents at least 90% noise, then the component is labeled N-IC if and only if one or more of the following secondary criteria apply:

- A. High frequencies: More than 50% of the power in the Fourier frequency spectrum of the component’s time course lies above 0.1 Hz. This cutoff frequency is appropriate for resting-state functional connectivity studies (Cordes et al., 2001; Greicius et al., 2007) and may be modified as needed for studies focusing on a different frequency range of neural signals.
- B. Spikes: One or more large (greater than five standard deviations), abrupt (over six seconds or less) changes in the normalized time course.
- C. Sawtooth pattern: Sharply and regularly alternating up-and-down time course.
- D. Sinus co-activation: Roughly ten or more thresholded voxels present in the superior sagittal sinus.

These component labeling rules are used in a two-pass, data denoising system involving the following steps.

1. fMRI data are pre-processed by conventional means, in this illustration with brain extraction, slice-timing correction, motion correction, high-pass temporal filtering, spatial smoothing, and normalization to a standard brain atlas.
2. Individual subject ICs are generated with spatial ICA.
3. ICs are labeled N-IC or S-IC.
4. The N-ICs are filtered from the original fMRI data using the time courses of the N-ICs with linear regression (Alternatively, reconstructing the fMRI data from the S-

ICs alone or projecting the fMRI data into the linear subspace spanned by the S-ICs could be performed).

5. Intensity normalization across volumes of fMRI data is performed, to filter out signals that affect the entire brain simultaneously.
6. A single, group ICA is performed on fMRI data that is temporally concatenated across subjects.
7. Group ICs are labeled N-IC or S-IC.
8. If group ICs are the object of the study, no further processing is needed: The S-ICs are the desired results of the denoising process. Otherwise, subject-specific time courses and spatial maps corresponding to each group IC can be derived through a process involving back-reconstruction from the aggregate mixing matrix as described in Calhoun et al. (2001b). Alternatively, the subject-specific components can be derived using dual regression (Beckmann et al., 2009), which involves regressing the individual subject fMRI datasets (output from Step 5) against the group component spatial maps to obtain subject-specific component time courses; followed by regressing the individual subject fMRI datasets against the subject-specific component time courses to obtain subject-specific component spatial maps. The subject-specific components corresponding to the N-ICs are then filtered from the fMRI data as in Step 4.

Steps 2–4 represent the first pass of the procedure, while Steps 6–8 represent the second.

Participants

The resting-state datasets for the IC examples below were derived from eighteen elderly (age > 60) Caucasian participants. Ten of the participants met DSM-IV criteria for unipolar major depression without psychotic features and had a score of 18 or greater on the 24-item Hamilton Depression Rating Scale (HDRS) within two weeks of scanning. The remaining eight were psychiatrically normal (assessed with the Structured Clinical Interview for DSM-III-R, SCID-R) control subjects, recruited through advertisement. The study was approved by Institutional Review Boards of Weill Cornell Medical College and the Nathan Kline Institute. All participants signed informed consent.

For assessment of improvement in sensitivity for detecting brain activation after VIID, data were used from a separate, task-driven study involving a right-handed adult male participant who was recruited from among students in an fMRI lab course at Columbia University. The study was approved by an Institutional Review Board at Columbia University.

Experimental paradigm

For the resting-state data, participants were instructed to close their eyes and relax without falling asleep and without focusing on any particular topic; a single, six-minute run of fMRI data was collected from each participant.

For the task-driven data, six, six-minute consecutive runs were collected while the participant viewed grayscale images of either faces or names of people who were strangers, friends, famous people, or the participant's parents. The same set of 40 images was shown in each run, in pseudo-randomized order. Each image was shown for an average of 9 seconds (random durations of 7.5 – 10.5 seconds), with a blank screen for 0.5 seconds between images. The images were shaded green or blue, and the participant was instructed to press a button in response to the hue, in order to ensure that he was paying attention to the images. Further details of this experimental paradigm are described in Kelly Jr. et al. (in press).

Data acquisition

The resting-state brain scans were acquired with the 1.5T Siemens Vision Scanner housed at the Nathan Kline Institute Center for Advanced Brain Imaging. For each participant, 180 contiguous BOLD contrast volumes were acquired in a single-shot, multi-slice echoplanar acquisition (TR = 2000 ms, TE = 35 ms, flip angle = 90, matrix = 64×64 , FOV = 224 mm, 22 slices, slice thickness = 5 mm, no gap, voxel size = $3.5 \times 3.5 \times 5$ mm); the first 10 volumes were discarded (to allow the scanner time to reach steady-state). High-resolution whole brain images were also acquired using three-dimensional T1-weighted magnetization-prepared rapid gradient echo (MPRAGE), with TR = 11.6 ms, TE = 4.9 ms, FOV = 320 mm, matrix = 256×256 , flip-angle = 8 degrees, slice thickness = 1.25 mm, number of slices = 172, no gap, and effective TI = 1017.2 ms.

Brain images for the task-driven experiment were acquired using a 1.5T General Electric (GE) TwinSpeed MRI scanner with a standard GE birdcage head coil, housed at the Functional Magnetic Resonance Imaging Center of Columbia University. Functional scans were performed using EPI-BOLD (TR = 2000 ms; TE = 38 ms; flip angle = 90 degrees; FOV = 192 mm; 64×64 matrix; 29 oblique-axial slices 4.5 mm thick; skip 0 mm; interleaved acquisition; voxel size = $3 \times 3 \times 4.5$ mm). Immediately after the functional scans, a 10-minute, high resolution, T1-weighted structural MRI image was acquired using the 3D SPGR sequence (186 slices; 256×256 ; FOV = 256 mm; voxel size = $1 \times 1 \times 1$ mm).

Image preprocessing

fMRI image preprocessing was carried out using FSL (Version 4.1), FMRIB's Software Library (<http://www.fmrib.ox.ac.uk/fsl>) (Smith et al., 2004; Woolrich et al., 2009), involving non-brain removal using BET (Smith, 2002); motion correction with MCFLIRT (Jenkinson et al., 2002); slice-timing correction for interleaved acquisitions using Fourier-space time-series phase-shifting; highpass temporal filtering using Gaussian-weighted least-squares straight line fitting (with $\sigma = 50$ seconds for resting-state data, and $\sigma = 27.5$ seconds for task-driven data); spatial smoothing using a Gaussian kernel with full-width half-maximum 8 mm; co-registration to high-resolution T1-weighted images; and normalization to standard space (Montreal Neurological Institute atlas, using resolutions of $4 \times 4 \times 4$ mm for resting state and $3 \times 3 \times 3$ mm for task-driven data, which were respectively the smallest voxel sizes that did not result in memory overflow errors on our 32-bit architecture computers while performing the group ICAs) using combined affine and non-linear registration (FSL FNIRT, with warp resolution = 10 mm) for resting-state data, and affine registration alone (FSL FLIRT) for task-driven data.

Data denoising

Steps 2–8 of the denoising process were performed as follows. Individual ICAs for all 18 subjects were applied to the preprocessed fMRI datasets, using Probabilistic Independent Component Analysis (Beckmann and Smith, 2004) as implemented in FSL's MELODIC (Multivariate Exploratory Linear Decomposition into Independent Components) Version 3.09. This process included voxel-wise variance normalization, whitening, projection into an N-dimensional subspace using Principal Component Analysis (PCA), and decomposition into time courses and corresponding spatial maps (spatiotemporal ICA components) by optimizing for non-Gaussian spatial source distributions using a fixed-point iteration technique (Hyvarinen, 1999). Estimated component maps were divided by the standard deviation of the residual noise and thresholded (alternative hypothesis test at $p > 0.95$) by fitting a mixture model to the histogram of intensity values. N was estimated using the FSL default, Laplace approximation (Minka, 2000; Beckmann and Smith, 2004). The rater (RK) was provided with the output from MELODIC for each component, including spatial maps, variance-normalized time courses, and Fourier power spectra of the time courses (Fig. 1).

For each spatial map, statistics were provided concerning what portion of thresholded voxels lay in the periphery of the brain (F) or within non-peripheral cerebrospinal fluid (CSF), white matter (WM), or grey matter (GM), in order to facilitate determination of what percentage of thresholded voxels might lie in functionally relevant grey matter clusters. These statistics were calculated with FSL utilities `fslmaths` and `fslstats` from individual subject spatial maps for F, CSF, WM, and GM, derived as follows. 1) CSF, WM, and GM spatial maps were generated from segmentation of individual high-resolution images with FSL FAST (Zhang et al., 2001); 2) Voxels falling outside of these regions were labeled F; and 3) Voxels were relabeled from CSF, WM, or GM to F for voxels whose mean fMRI data value was less than half of the mean value for the whole brain.

Each of the ICs were labeled N-IC or S-IC. The time courses for the N-ICs were regressed out of the fMRI data using the `fsl_regfilt` utility. Each volume of the fMRI data was intensity normalized with FSL FEAT. A single set of group ICs were generated from these data using MELODIC, as described above, but with the multi-subject temporal concatenation option. The group ICs were labeled.

Assessment of training requirements and inter-rater reliability

A co-author (SM) labeled a subset of the ICs from the first pass for assessment of inter-rater reliability, after calibrating labeling methods with the first author using ICs (106 in all) from the 4 subjects that were selected for Examples 1–6. The subset consisted of all ICs (124 in all) for 5 subjects randomly selected from the remaining 14 subjects. SM had not seen any of the ICs from this subset prior to the labeling exercise.

Assessment of improvement in sensitivity of fMRI data analysis on task-based data

For each task-driven run, a 3D spatial map representing brain areas that were more active during presentation of face images than during presentation of name images (or blank background) was obtained using the general linear model (GLM) by regressing the fMRI data against a hemodynamic response time course corresponding to times when face images were shown. FSL FEAT with FILM local autocorrelation correction (Woolrich et al., 2001) was used to generate Gaussianised, z-statistic images reflecting parameter estimates in relation to their standard errors and degrees of freedom. The hemodynamic response time courses for each run were generated by convolving the experimentally-derived time courses (for showing faces) against a Gaussian function with peak lag = 5 seconds and $\sigma = 2.8$ seconds. A combined, six-run fixed-effects analysis was also performed with FEAT. Z-statistic images were processed by maximum-height thresholding voxels based on Gaussian Random Field Theory (GRFT) to a corrected, one-tailed significance of $p \leq 0.05$ (Worsley, 2004).

3D spatial maps so obtained were compared before and after the first and second passes of the denoising process, and after applying intensity normalization (IN) without ICA-based denoising. For each run, voxels near threshold ($4.3 < z\text{-score} < 5.3$) prior to denoising were identified, and the mean z-score for these voxels was determined. Differences in mean z-scores (paired by run) were considered statistically significant when $p < 0.05$, evaluated with the Wilcoxon signed-rank test (2-tailed). Our primary hypothesis was that z-scores would increase after VIID. Standard errors for differences in six-run means were used to generate 95% confidence intervals (CI). Finally, improvement in sensitivity was assessed qualitatively by comparing the robustness of the activation patterns in the individual-run and combined-run maps, particularly in the right temporal occipital fusiform gyrus (rTOF), where we anticipated some activation in response to viewing faces (Kanwisher et al., 1997; O'Craven and Kanwisher, 2000; Downing et al., 2006). Statistical testing was performed with SPSS 12.0 for Windows.

Demonstration with synthetic data

To verify that the VIID procedure is capable of removing artifacts while leaving neural signals of interest intact, we tested it using fMRI data seeded with six artificial ICs. The ICs represented artifacts (N-ICs 1–3) or neurally-derived signals (S-ICs 1–3) as follows.

N-IC1: Global signal over 30% of voxels; time course dominated by high frequencies (>0.1 Hz).

N-IC2: Global signal over 10% of voxels, involving superior sagittal sinus; time course with periodic pattern that might resemble aliasing of high-frequency structured noise.

N-IC3: Motion-related pattern; time course with prominent spike.

S-IC1: Default-mode network; time course typical of default mode pattern in resting state.

S-IC2: Visual cortex; task-related time course convolved with hemodynamic response function.

S-IC3: Auditory cortex; task-related time course convolved with hemodynamic response function.

These synthetic signal sources were added to a "componentless" resting-state fMRI data run that consisted of Gaussian noise and signals too weak to be detected by ICA. The "componentless" fMRI data run was prepared by randomly selecting one of the remaining nine resting-state runs that had not been used for the inter-rater reliability study; partially preprocessing with non-brain removal, slice-timing correction, and motion correction; followed by filtering out the 22 ICs previously derived in the main experiment of this study (corresponding to 60% of the total variance in the fMRI data run), by regressing out the 22 IC time courses from the fMRI data using `fsl_regfilt`. The first 70 volumes were then discarded, leaving 100 volumes (to make all signal sources and "componentless" data the same length). After highpass temporal filtering ($\sigma = 50$ seconds) and spatial smoothing (8mm) of the filtered data, a new ICA revealed four small components (9% of total variance), which we regressed out of the data using the same procedure, resulting in our "componentless" fMRI data run.

Each synthetic signal source was the outer product of a time course (values at 100 time points) and a spatial map (intensity values for each voxel in the brain) to yield a set of 100 volumes spanning 200 seconds (TR=2 seconds, consistent with our resting-state data). Time courses were demeaned and scaled to unit variance after shortening them when necessary to 100 time points by removing time points from the beginning. Each spatial map was the product of a) an initial spatial map, with voxel values limited to the range -1 to $+1$ (with two exceptions noted below), and b) a spatial map with voxel values corresponding to 1% of the mean MR signal intensities of the "componentless" fMRI data run. IC mixture model probability maps were used when initial spatial maps were derived from ICs. Negative values in initial spatial maps represented brain regions anticorrelated with regions having positive values.

The time courses for N-ICs 1–3 (Figs. 9a–c) were drawn from among the eight signal sources in the simulated fMRI dataset found at http://mlsp.umbc.edu/simulated_fmri_data.html. The time course for S-IC1 (Fig. 9d) was derived from an IC selected from the eight resting-state runs that had not been used for other purposes in the experiments described here. The selected IC was the one whose thresholded spatial map had the greatest overlap with the "standard template" of the default mode network used in Greicius et al. (2007). The time courses for S-IC2 and S-IC3 (Figs. 9e–f)

were randomly selected from the six *a priori* time courses (convolved with hemodynamic response function) used to analyze our task-related fMRI data.

The initial spatial map for N-IC1 (Fig. 9a) was a binary map of randomly distributed spots covering 30% of the brain's voxels. The spots were of random shapes and sizes formed from 27-voxel cubes where each voxel in the cube had a probability of 0.5 of having the value 1. The same algorithm was used to generate the initial spatial map for N-IC2 (Fig. 9b), which was a binary map of spots covering 10% of the brain. To this map we added "activation" in 50% of the superior sagittal sinus from the plane $z = 4$ and above in MNI space. This was done by hand-drawing a mask representing the superior sagittal sinus and filling it randomly with spots until 50% of the superior sagittal sinus mask was covered. We gave each "active" voxel in the superior sagittal sinus the value 5 rather than 1 to make it easier to detect after spatial smoothing. The initial spatial map for N-IC3 (Fig. 9c) was derived from the spatial IC pattern that resulted from shifting the brain in our data diagonally by 1 voxel (1-voxel shifts in the +x, +y, and +z directions, after a randomly selected time point) as was suggested in McKeown et al. (1998). We elected not to actually shift the brain location in space to create this IC because of the unrealistically large portion of the total variance that would have been represented by such an IC, and because we wanted to carefully track any differences that might arise between the added synthetic signals and how they would be represented with ICA. For S-IC1 we used the corresponding spatial map (Fig. 9d) from the IC whose time course was selected for S-IC1. For S-IC2 and S-IC3 (Figs. 9e–f) we chose an initial spatial map corresponding respectively to the bilateral probability map for Lateral Occipital Cortex, Superior Division and Superior Temporal Gyrus, Posterior Division from the Harvard-Oxford Cortical Structural Atlas (HOCSA) provided with FSL. We scaled the S-IC2 map to a maximum value of 1 and the S-IC3 map to a maximum value of 2. Double the usual value was used in the latter case to compensate for the fact that S-IC3's spatial map covered a smaller region of the brain than the other components, and therefore risked not creating enough net variance to be detected by ICA.

The synthetic signal sources were added to the "componentless" fMRI data run; then high-pass temporal filtering, spatial smoothing, and ICA were performed. The resulting ICs were visually inspected to evaluate how well the synthetic signal sources were represented. The N-ICs were identified (by SM, who had not seen the data beforehand), and removed with `fsl_regfilt` from the fully preprocessed fMRI data. ICA was performed again and the ICs visually inspected to evaluate how well the N-ICs had been removed from the data, while preserving the S-ICs.

Finally, task-related GLM was performed before and after removal of the N-ICs to examine the effects of denoising on GLM-derived spatial maps. The GLM-based procedure used on our task-based data was applied, using the time courses for S-IC2 and S-IC3, and fMRI data before and after denoising.

Results

For the resting-state data, the individual-ICA pass resulted in 21 to 37 components for each subject, for a total of 464 components; 292 (63%) were labeled N-IC. In the second pass, 2 (13%) of 16 components were labeled N-IC. Inter-rater agreement was 96%, with Cohen's $\kappa = 0.91$. The total time required by SM for calibration sessions was 165 minutes; and for the labeling procedure 45 minutes, corresponding to an average of 9 minutes per subject and 22 seconds per IC. Little time was required to move from one IC to the next using FSL's web browser interface, so approximately 20 seconds per IC were devoted to meticulously applying the labeling rules while examining the FSL-provided axial slices, time courses, and

frequency spectra; and statistics on paper concerning the percentage of components in CSF, WM, GM, and F.

Several illustrative examples are provided below. Examples 1–2 and 6–8 are typical of most cases, showing patterns that with visual inspection can be labeled easily and quickly, within seconds. Examples 3–5 show more difficult cases where less than 100% inter-rater agreement might be expected. Examples 1–6 were taken from the procedure's first pass, and Examples 7–8 from the second pass. The spatial maps are displayed in radiological convention, showing a full set of axial slices in Fig. 1 and every other slice in Figs 2–12.

Example 1

Fig. 1 depicts a component that was labeled N-IC due to the classic, ring-like pattern in the periphery of the brain. Some “deactivation” along the edge of the brain opposite to regions of “activation” or, as in this case, adjacent to regions of “activation” is commonly seen in this type of pattern. Eighty-five percent of thresholded voxels lie in peripheral regions of the brain, while only 6% lie in non-peripheral grey matter.

Example 2

In Fig. 2, thresholded voxels are tightly clustered bilaterally in frontal polar regions, but most of those voxels lie in the periphery (54%) or CSF (13%, including CSF between brain and skull) with only 20% in GM. This component was labeled N-IC because nearly all thresholded voxels in GM were considered part of a peripheral pattern of “activation,” and therefore did not count toward the 10% GM required for an S-IC label. Fig. 2 shows an example of a saw-tooth time course pattern, which was not considered in labeling this component.

Example 3

The spatial map in Fig. 3 shows tightly clustered thresholded voxels spanning a large portion of the frontal lobes. The thresholded voxels are not diffusely scattered over the brain, but the huge frontal cluster occupies an improbably large region of the brain with no regard for grey matter boundaries, indicating that this cluster probably does not reflect neural activity. This component could be labeled differently depending on the rater. A rater who feels certain that this pattern represents an artifact would label it N-IC, while one who feels very uncertain would label it S-IC. Raters falling between these extremes would classify the component as N-IC due to the large spike in the time course.

Example 4

“Speckled” patterns were not seen in these data, presumably because of the extent of spatial smoothing, but “spotty” patterns with small clusters diffusely spread over the brain were frequently seen. One such example is shown in Fig. 4, which raters might label differently because although there is a large central cluster in CSF and the pattern is a diffusely scattered one, 32% of the thresholded voxels lie within GM. It is conceivable, though perhaps unlikely, that the component represents a synthesis of structured noise, random noise, and/or neural activity (>10%). Raters unable to decide upon classification based on the spatial map would label this component N-IC because most of the power in the frequency spectrum of the time course lies above 0.1 Hz.

Example 5

The spatial map in Fig. 5 shows a diffuse and somewhat spotty pattern that is not scattered across a large region of the brain. Fifty percent of thresholded voxels lies in peripheral brain areas and 23 percent in CSF, while 21 percent lies in GM. A rater who attributes the entire

pattern to structured or random noise would label this component N-IC, while a rater who judges that half or more of the thresholded voxels in GM probably represent neural signal would label the component S-IC. Uncertain raters leaning toward labeling the component N-IC would do so upon making note of the considerable presence of thresholded voxels in the superior sagittal sinus.

Example 6

Fig. 6 depicts a component whose spatial map shows approximately half of thresholded voxels in small to larger grey-matter clusters. The clusters are limited to smaller regions of the brain rather than diffusely scattered. No further information is needed to label this component S-IC.

Example 7

The group ICA component in Fig. 7 was labeled N-IC because thresholded voxels were nearly exclusively located in a large central white matter cluster.

Example 8

The group ICA component in Fig. 8 was labeled S-IC due to thresholded voxels being tightly clustered in predominantly grey matter regions. This component resembles the left parietal-dorsolateral prefrontal cortex component "IC 16" described in Chen et al. (2008), one of nine ICs consistently derived from group independent component analyses of resting-state fMRI data.

For the task-driven data, the first pass resulted in 15 to 29 components for each run, for a total of 143 components; 80 (56%) were labeled N-IC and regressed out of their corresponding fMRI data. In the second pass, none of the 8 group ICs were labeled N-IC, so the only procedure performed between the first pass and the end of the second pass was intensity normalization (IN). The mean z-score in the targeted regions increased from 4.75 to 5.21 (9.7%, CI 3.5–15.9%) after the first pass and to 5.63 (18.4%, CI 5.6–31.2%) after the second pass. After IN alone, without ICA-based denoising, the mean z-score increased to 5.33 (12.2%, CI 2.0–22.5%). Mean z-scores after the combination of first-pass VIID with IN were significantly greater than 1) before denoising; as well as 2) after first-pass VIID; or 3) IN alone. Without denoising, no thresholded voxels were found in the rTOF for any of the runs. After first-pass VIID, thresholded voxels were found in two runs and after the second pass, in four runs. Finally, in the combined-run map, an arm of activated voxels could be seen extending from the visual cortex rostrally into the rTOF before denoising. This arm appeared more robust after first-pass VIID and most robust after first-pass VIID + IN. Of the 97 voxels (3×3×3 mm) in the rTOF (defined as voxels whose probability of belonging to the rTOF was ≥ 0.5 according to the HOCSA), the number of thresholded voxels increased respectively from 13 to 32 to 47.

Finally, our demonstration using synthetic data resulted in six ICs (Fig. 10), each matching one of the seeded synthetic ICs (Fig. 9). The percentage of total fMRI variance accounted for by each IC was 15% (S-IC1), 14% (N-IC1), 9% (N-IC2), 6% (N-IC3), 6% (S-IC2), and 4% (S-IC3): in all, 55% of the total variance. Some mild distortion of the time courses was evident, but apart from the effects of spatial smoothing, the spatial maps were nearly identical to their corresponding seed ICs (Reversing the polarity of both an IC's spatial map and time course, as happened with N-IC3, is an equivalent way of expressing the data). All six ICs were correctly labeled. After filtering out the N-ICs, ICA resulted in three ICs (Fig. 11), one corresponding to each of the S-ICs. The S-IC spatial maps and time courses from the second ICA were nearly identical to those from the first, while the spatial maps from the GLM analyses (Fig. 12) were more robust after denoising than before, showing more of the

seeded voxels. These findings demonstrate that ICs corresponding to artifacts can be filtered from the data without significantly distorting the remaining signals, resulting in noticeable improvements in the sensitivity and accuracy of GLM-based analyses.

Discussion

The main purpose of this article is to advocate for detailed descriptions of what is meant by visual inspection of IC spatial maps and/or time courses, in articles that use this term. Such descriptions are needed to facilitate the exploration of visual inspection as part of fMRI data denoising procedures. The level of detail in such descriptions must be sufficient to allow reproduction by independent investigators, meaningful comparison of findings, and scientific discussion of the potential uses for and advantages of one method of visual inspection over another. We provide an example of a description of an operationalized denoising procedure that includes visual inspection of ICs, which we believe is detailed enough to meet this requirement. The procedure involves a novel combination of methods and conceptual framework that allows for the possibility that some ICs might represent a combination of noise and neural signals of interest, rather than representing exclusively one category or the other. Two independent trained raters achieved 96% ($\kappa = 0.91$) agreement using this procedure; and testing the procedure on a separate, task-related dataset resulted in statistically significant (18% increase in z-score) and visually noticeable improvements in sensitivity for detecting responses in brain areas relevant to the experimental paradigm.

Our goal was to define a data-driven VIID procedure that combines elements from previously described ICA-based denoising approaches and can be performed with a minimum of training and labeling time, using widely available brain imaging software packages. The procedure focuses primarily on the appearance of IC spatial maps because we found that IC spatial maps are often highly suggestive of artifacts or of neural activity corresponding to known brain functions (constellations of brain areas used for particular brain functions). The use of peripheral “activation” to identify artifacts was suggested in McKeown et al. (1998) and implemented in Tohka et al. (2008). Considering diffuse, speckled (or spotty) patterns as indications of noise was suggested in McKeown et al. (1998); and conversely, considering tightly clustered, not diffusely spread spatial patterns as indications of neural signal was discussed in Sui et al. (2009). Utilization of information concerning what portion of spatial maps lies in CSF, GM, and WM for automated methods of IC-labeling was described in Stevens et al. (2007) and Sui et al. (2009). Using “activation” in the superior sagittal sinus (Criterion D) as an indication of artifacts is similar to the focus on vasculature in Zou et al. (2009). Such activation has been hypothesized to be due to breathing-related changes in central venous pressure (Windischberger et al., 2002), and has been found to correlate with the cardiac cycle (Dagli et al., 1999). The sinus co-activation criterion was secondary because we noticed that sometimes such activation was present in the spatial maps of components that in all other respects seemed to reflect neural activity. Only three temporal aspects of ICs were considered in the VIID procedure, to be applied in secondary criteria when labeling based on spatial characteristics was inconclusive. The high-frequencies criterion (A) is exactly as was implemented in Greicius et al. (2004; 2007); the “spikes” criterion (B) is similar to criteria suggested in McKeown et al. (1998) and implemented in Tohka et al. (2008); and we included the saw-tooth pattern criterion (C), similar to the “quasiperiodic” pattern suggested in McKeown et al. (1998), because we hypothesized that saw-tooth temporal patterns are a sign of aliasing of cardiac and/or respiratory signals whose frequencies are faster than the Nyquist frequency for our experiment (0.25 Hz) (Huettel et al., 2004a; Beckmann et al., 2005). The main reason for focusing only secondarily on IC temporal characteristics was that frequency ranges of artifactual and neural activity sometimes overlap: Some artifacts manifest themselves at frequencies typically dominated by neural activity (Beckmann et al., 2005; Birn et al.,

2006), and some neural activity occurs at relatively higher frequencies, as illustrated by EEG data (Luck, 2005). Such overlap in frequencies complicates the determination of how much of component variance is due to artifacts vs. neural signal. In addition, we found that some of the methods for labeling ICs based on temporal characteristics required programs that were not available in the FSL software package.

We introduced a two-pass system because of perceived advantages with individual ICs over group ICs and vice-versa. Individual ICs allow comparison with a more precise segmentation of high-resolution images; while group ICs may reflect signals that might only be revealed after combining data from all study participants. In some cases, the group ICAs might not be necessary depending upon how much of the variance from artifacts is removed in the first pass of denoising, as illustrated in the denoising of our task-related data. We performed intensity normalization between the first and second passes because the (spatial) ICA constraint of spatial independence reduces the likelihood that ICA would detect components affecting most of the brain (Thomas et al., 2002), so some form of global brain signal removal may enhance ICA-based denoising, as was the case with denoising of our task-related data. We did not regress out global brain signal (e.g., CSF, WM, GM, or whole-brain signal) as in Fox et al. (2005) because we were concerned that in cases where the data variance is dominated by neural signals of interest, regressing out global brain signals might remove much of neural signals (Petersson et al., 1999; Birn et al., 2006), more so than would be the case with intensity normalization; however for "noisy" data, regressing out global brain signals should be an acceptable alternative to intensity normalization.

The rationale for further exploration of VIID procedures includes the following. 1) The current study and others have demonstrated the potential improvement in fMRI data analysis sensitivity that may be obtained by adding ICA-based denoising to conventional fMRI data preprocessing. 2) Procedures for manual labeling need to be defined and validated, in order for validation of automated methods based on manual methods to be meaningful. 3) Improved methods for IC labeling with visual inspection may model enhancements to IC labeling with automated methods. 4) Perhaps the most compelling reason for exploring IC labeling with visual inspection is its presumed accuracy, despite potential limitations regarding rater expertise, time expended by raters, and rater subjectivity. The situation is analogous to that with visual inspection for determination of regions of interest (ROIs) in fMRI and other brain imaging studies. The amount of training and time required for manual drawing of ROIs can be prohibitive. According to one estimate, at least one month of training is required, followed by hours to weeks to manually label the anatomy of a single brain (Klein et al., 2009). However, in spite of such costs and the availability of automated methods for drawing ROIs, some studies still utilize manual drawing of ROIs (McKeown and Hanlon, 2004; Pereira et al., 2007; Wager et al., 2008), presumably because of anticipated improvement in drawing accuracy with manual methods (Huettel et al., 2004c). Although human subjectivity can decrease reliability for manual drawing of ROIs or visual inspection of ICs, this decrease would be inconsequential if it could be demonstrated that the overall accuracy for manual methods is greater than that for automated methods. Also, a potential selection bias due to the human element can frequently be neutralized by blinding the rater with respect to knowledge of subject and/or scanning session characteristics.

We believe that the demands of time for training and performance of IC labeling with visual inspection are small enough that visual inspection may be preferred over automated methods in many applications involving ICA-based denoising. Raters performing IC labeling should have some knowledge of the general locations of brain CSF, WM, and GM, but detailed knowledge of comparative brain anatomy as required for drawing ROIs is not necessary. The results from our inter-rater reliability study indicate that training of raters need not take more than a few hours, and once raters are trained, approximately 5–10 minutes (possibly

fewer) are required per subject, depending upon the number of IC's generated per subject and the rater's experience with the visual inspection labeling procedure. Most components can be classified quickly through pattern recognition. The process of visual inspection can also save time by serving a dual role in facilitating fMRI quality assurance (Huettel et al., 2004b) through identifying artifacts, including those related to scanner malfunction.

Our example description of fMRI data denoising through visual inspection of ICs is not intended as an optimal, finished proposal, but as one of many conceivable proposals, and a starting point for discussion of what elements should be included in a VIID procedure. In addition to the minor modifications to the procedure that we have indicated, many potential enhancements can be considered. To provide three examples: 1) Methods involving task-related information (if available) can be added to the VIID procedure (McKeown, 2000; Calhoun et al., 2001a; Thomas et al., 2002; Moritz et al., 2003; Kochiyama et al., 2005; McKeown et al., 2005). 2) Combining the proposed visual inspection method with an automated component clustering method such as Partner-Matching (PM) (Wang and Peterson, 2008) might save time and mental effort in the labeling process by organizing components across subjects or sessions into groups according to spatial similarity. One could begin manual labeling on clusters of homologous ICs identified by PM as being highly reliable (most significantly reproducible), representing artifacts or neural activity that are common across subjects or sessions. Then, one could visually inspect the remaining clusters of components that are less significantly reproducible across subjects or sessions, which may represent components of artifacts or neural activity that exist only in some individuals. Conversely, the proposed visual inspection method may be applied to verification of the automated labeling of highly reliable clusters of similar components identified by PM, to explore the possibility that such components, spatial similarities notwithstanding, might appear to differ somewhat in degree of contribution from artifactual vs. neural signals. 3) We have observed that ICs that are reproducible from one run to the next or that correlate well with an experimental task appear to maintain the characteristic form of their spatial maps regardless of what z-score is used for thresholding. In contrast, the spatial maps of some other components dwindle in size and disappear quickly as thresholding z-scores are increased. We speculate that this property might be useful in distinguishing components representing predominantly neural signals from those representing predominantly random noise. Automated methods could be used to generate statistics such as the slope of the log of the number of thresholded voxels as a function of thresholding z-score. Such statistics could then be incorporated into a visual inspection procedure. This example illustrates the general principle that automated methods can be used to enhance visual inspection by providing the rater with precise statistics that would otherwise be time-consuming or impossible for the rater to approximate.

We envision that the development of VIID methods might proceed as follows. Potentially useful elements that may be incorporated in visual inspection procedures would be described and motivated. Based on face validity considerations, one or more VIID procedures would be selected for evaluation of reliability and validity. Testing of inter-rater reliability would be straightforward. However, validity testing would be complicated by the absence of a definitive standard for what constitutes a "correct" IC. With manual drawing of ROIs the structures shown in brain images can ultimately be compared to those seen in cadavers, and well-established comparative neuroanatomy considerations can help determine the accuracy of ROI drawings. With ICA-based denoising no physical entity corresponding to an IC can be examined. Spatially independent components are an abstraction of the fMRI data that does not take into account non-linear effects and that is based upon the assumption that the complexities of brain function can be modeled simply, with no more than a few dozen ICs that are assumed to reflect perfectly temporally synchronized brain activity (McKeown and Sejnowski, 1998; Thomas et al., 2002). Thus, validation efforts cannot determine the

“correctness” with which ICs reflect brain function. Instead, VIID validation must proceed by circuitous routes. For example, artificial signal can be added to fMRI data to gain insight into how such additions might affect the generated set of ICs. Experiments of this kind might elucidate the conditions under which an IC might represent a synthesis of neural signal, structured noise, and/or random noise.

Ultimately, the most important indicator of the success of a VIID procedure is the improvement in functional signal-to-noise ratio that results from its application. A variety of methods can be used to assess the improvement in sensitivity and/or specificity of fMRI data analysis after denoising (Biswal et al., 1996; Liu et al., 2001; Stone et al., 2002; Thomas et al., 2002; Kochiyama et al., 2005; McKeown et al., 2005; Gretton et al., 2006). If an improvement in receiver operating characteristics is observed after a modification of a VIID procedure, we can judge that the modification has a positive effect on the process, even if we cannot demonstrate that the accuracy of generating ICs or of labeling them has been enhanced. Thus, it should be possible to improve VIID procedures through a process of trial and error, systematically including or excluding various elements of the procedures. We hope that such an iterative process will result in successively better denoising procedures and elucidate the potential advantages and disadvantages of IC labeling with visual inspection methods compared with automated methods.

In conclusion, the use of visual inspection to label ICs has been reported as part of procedures for denoising fMRI data and as a standard of comparison for automated labeling methods. In order for such studies to be reproducible, detailed descriptions of visual inspection procedures are needed. We address this need by providing an example of an operationalized VIID procedure and demonstrate its reliability, costs in terms of time for training and IC labeling, and capacity for improving the sensitivity of results from fMRI data analysis. In addition to serving as a procedure that can readily be implemented using a standard brain imaging software package (FSL, among others) we hope that the procedure will serve as a starting point for discussion of what elements should be included in a VIID procedure, and will encourage investigators to document what steps are involved in their visual inspection procedures.

Acknowledgments

This work was supported by NIMH grants R01 MH065653, P30 MH085943, T32 MH019132 (to George S. Alexopoulos, M.D.), K23 MH067702 (to Christopher F. Murphy, Ph.D.), K23 MH074818 (to Faith M. Gunning, Ph.D.), the Sanchez and TRU Foundations, and Forest Pharmaceuticals, Inc. Dr. Alexopoulos has received research grants by Forest Pharmaceuticals, Inc. and Cephalon and participated in scientific advisory board meetings of Forest Pharmaceuticals. He has given lectures supported by Forest, Bristol Meyers, Janssen, and Lilly and has received support from Comprehensive Neuroscience, Inc. for the development of treatment guidelines in late-life psychiatric disorders. All other authors report no competing interests. The authors thank Michael Greicius, M.D. for the default mode network spatial map used in our demonstration with synthetic data. We also thank Raj Sangoi (RT) (R) (MR) for his work as Chief MRI Research Technologist, Nathan Kline Institute; and at Columbia University, Stephen Dashnaw for his work as Imaging Supervisor, and Joy Hirsch, Ph.D. for fMRI instruction and use of her imaging lab.

References

- Andersson JL, Hutton C, Ashburner J, Turner R, Friston K. Modeling geometric deformations in EPI time series. *Neuroimage*. 2001; 13:903–919. [PubMed: 11304086]
- Beauchamp MS. Detection of eye movements from fMRI data. *Magn Reson Med*. 2003; 49:376–380. [PubMed: 12541259]
- Beckmann CF, DeLuca M, Devlin JT, Smith SM. Investigations into resting-state connectivity using independent component analysis. *Philos Trans R Soc Lond B Biol Sci*. 2005; 360:1001–1013. [PubMed: 16087444]

- Beckmann, CF.; Mackay, CE.; Filippini, N.; Smith, SM. Group comparison of resting-state fMRI data using multi-subject ICA and dual regression; 15th Annual Meeting of Organization for Human Brain Mapping, poster 441 SU-AM; 2009.
- Beckmann CF, Smith SM. Probabilistic independent component analysis for functional magnetic resonance imaging. *IEEE Trans Med Imaging*. 2004; 23:137–152. [PubMed: 14964560]
- Birn RM, Bandettini PA, Cox RW, Jesmanowicz A, Shaker R. Magnetic field changes in the human brain due to swallowing or speaking. *Magn Reson Med*. 1998; 40:55–60. [PubMed: 9660553]
- Birn RM, Cox RW, Bandettini PA. Experimental designs and processing strategies for fMRI studies involving overt verbal responses. *Neuroimage*. 2004; 23:1046–10458. [PubMed: 15528105]
- Birn RM, Diamond JB, Smith MA, Bandettini PA. Separating respiratory-variation-related fluctuations from neuronal-activity-related fluctuations in fMRI. *Neuroimage*. 2006; 31:1536–1548. [PubMed: 16632379]
- Birn RM, Murphy K, Bandettini PA. The effect of respiration variations on independent component analysis results of resting state functional connectivity. *Hum Brain Mapp*. 2008; 29:740–750. [PubMed: 18438886]
- Biswal B, DeYoe AE, Hyde JS. Reduction of physiological fluctuations in fMRI using digital filters. *Magn Reson Med*. 1996; 35:107–113. [PubMed: 8771028]
- Calhoun VD, Adali T, McGinty VB, Pekar JJ, Watson TD, Pearlson GD. fMRI activation in a visual-perception task: network of areas detected using the general linear model and independent components analysis. *Neuroimage*. 2001a; 14:1080–1088. [PubMed: 11697939]
- Calhoun VD, Adali T, Pearlson GD, Pekar JJ. A method for making group inferences from functional MRI data using independent component analysis. *Hum Brain Mapp*. 2001b; 14:140–151. [PubMed: 11559959]
- Chen S, Ross TJ, Zhan W, Myers CS, Chuang KS, Heishman SJ, Stein EA, Yang Y. Group independent component analysis reveals consistent resting-state networks across multiple sessions. *Brain Res*. 2008; 1239:141–151. [PubMed: 18789314]
- Chen W, Zhu XH. Suppression of physiological eye movement artifacts in functional MRI using slab presaturation. *Magn Reson Med*. 1997; 38:546–550. [PubMed: 9324320]
- Cordes D, Haughton VM, Arfanakis K, Carew JD, Turski PA, Moritz CH, Quigley MA, Meyerand ME. Frequencies contributing to functional connectivity in the cerebral cortex in "resting-state" data. *AJNR Am J Neuroradiol*. 2001; 22:1326–1333. [PubMed: 11498421]
- Dagli MS, Ingelholm JE, Haxby JV. Localization of cardiac-induced signal change in fMRI. *Neuroimage*. 1999; 9:407–415. [PubMed: 10191169]
- Downing PE, Chan AW, Peelen MV, Dodds CM, Kanwisher N. Domain specificity in visual cortex. *Cereb Cortex*. 2006; 16:1453–1461. [PubMed: 16339084]
- Fox MD, Snyder AZ, Vincent JL, Corbetta M, Van Essen DC, Raichle ME. The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proc Natl Acad Sci U S A*. 2005; 102:9673–9678. [PubMed: 15976020]
- Friston KJ, Williams S, Howard R, Frackowiak RS, Turner R. Movement-related effects in fMRI time-series. *Magn Reson Med*. 1996; 35:346–355. [PubMed: 8699946]
- Glover GH, Li TQ, Ress D. Image-based method for retrospective correction of physiological motion effects in fMRI: RETROICOR. *Magn Reson Med*. 2000; 44:162–167. [PubMed: 10893535]
- Greicius MD, Flores BH, Menon V, Glover GH, Solvason HB, Kenna H, Reiss AL, Schatzberg AF. Resting-state functional connectivity in major depression: abnormally increased contributions from subgenual cingulate cortex and thalamus. *Biol Psychiatry*. 2007; 62:429–437. [PubMed: 17210143]
- Greicius MD, Srivastava G, Reiss AL, Menon V. Default-mode network activity distinguishes Alzheimer's disease from healthy aging: evidence from functional MRI. *Proc Natl Acad Sci U S A*. 2004; 101:4637–4642. [PubMed: 15070770]
- Gretton A, Belitski A, Murayama Y, Scholkopf B, Logothetis N. The effect of artifacts on dependence measurement in fMRI. *Magn Reson Imaging*. 2006; 24:401–409. [PubMed: 16677946]
- Grooten S, Hutton C, Ashburner J, Howseman AM, Josephs O, Rees G, Friston KJ, Turner R. Characterization and correction of interpolation effects in the realignment of fMRI time series. *Neuroimage*. 2000; 11:49–57. [PubMed: 10686116]

- Hu X, Le TH, Parrish T, Erhard P. Retrospective estimation and correction of physiological fluctuation in functional MRI. *Magn Reson Med*. 1995; 34:201–212. [PubMed: 7476079]
- Huettel, SA.; Song, AW.; McCarthy, G. Functional magnetic resonance imaging. Vol. 229. Sunderland, Mass: Sinauer Associates; 2004a. p. 75
- Huettel, SA.; Song, AW.; McCarthy, G. Functional magnetic resonance imaging. Sunderland, Mass: Sinauer Associates; 2004b. p. 254–256.
- Huettel, SA.; Song, AW.; McCarthy, G. Functional magnetic resonance imaging. Sunderland, Mass: Sinauer Associates; 2004c. p. 350
- Hyvarinen A. Fast and robust fixed-point algorithms for independent component analysis. *IEEE Trans Neural Netw*. 1999; 10:626–634. [PubMed: 18252563]
- Jenkinson M, Bannister P, Brady M, Smith S. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage*. 2002; 17:825–841. [PubMed: 12377157]
- Kanwisher N, McDermott J, Chun MM. The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J Neurosci*. 1997; 17:4302–4311. [PubMed: 9151747]
- Kelly RE Jr, Wang Z, Alexopoulos GS, Gunning FM, Murphy CF, Morimoto SS, Kanellopoulos D, Jia Z, Lim KO, Hoptman MJ. Hybrid ICA-seed-based methods for fMRI functional connectivity assessment: a feasibility study. *Int J Biomed Imaging*. in press.
- Klein A, Andersson J, Ardekani BA, Ashburner J, Avants B, Chiang MC, Christensen GE, Collins DL, Gee J, Hellier P, Song JH, Jenkinson M, Lepage C, Rueckert D, Thompson P, Vercauteren T, Woods RP, Mann JJ, Parsey RV. Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. *Neuroimage*. 2009; 46:786–802. [PubMed: 19195496]
- Kochiyama T, Morita T, Okada T, Yonekura Y, Matsumura M, Sadato N. Removing the effects of task-related motion using independent-component analysis. *Neuroimage*. 2005; 25:802–814. [PubMed: 15808981]
- Liu TT, Frank LR, Wong EC, Buxton RB. Detection power, estimation efficiency, and predictability in event-related fMRI. *Neuroimage*. 2001; 13:759–773. [PubMed: 11305903]
- Luck, SJ. An introduction to the event-related potential technique. Cambridge, MA: MIT Press; 2005.
- McKeown M, Hu YJ, Jane Wang Z. ICA Denoising for Event-Related fMRI Studies. *Conf Proc IEEE Eng Med Biol Soc*. 2005; 1:157–161. [PubMed: 17282135]
- McKeown MJ. Detection of consistently task-related activations in fMRI data with hybrid independent component analysis. *Neuroimage*. 2000; 11:24–35. [PubMed: 10686114]
- McKeown MJ, Hanlon CA. A post-processing/region of interest (ROI) method for discriminating patterns of activity in statistical maps of fMRI data. *J Neurosci Methods*. 2004; 135:137–147. [PubMed: 15020098]
- McKeown MJ, Makeig S, Brown GG, Jung TP, Kindermann SS, Bell AJ, Sejnowski TJ. Analysis of fMRI data by blind separation into independent spatial components. *Hum Brain Mapp*. 1998; 6:160–188. [PubMed: 9673671]
- McKeown MJ, Sejnowski TJ. Independent component analysis of fMRI data: examining the assumptions. *Hum Brain Mapp*. 1998; 6:368–372. [PubMed: 9788074]
- Minka, TP. Automatic choice of dimensionality for PCA. MIT Media Laboratory Vision and Modeling Group. 2000. Technical Report 514
- Moritz CH, Rogers BP, Meyerand ME. Power spectrum ranked independent component analysis of a periodic fMRI complex motor paradigm. *Hum Brain Mapp*. 2003; 18:111–122. [PubMed: 12518291]
- O'Craven KM, Kanwisher N. Mental imagery of faces and places activates corresponding stimulus-specific brain regions. *J Cogn Neurosci*. 2000; 12:1013–1023. [PubMed: 11177421]
- Pereira AC, Huddleston DE, Brickman AM, Sosunov AA, Hen R, McKhann GM, Sloan R, Gage FH, Brown TR, Small SA. An in vivo correlate of exercise-induced neurogenesis in the adult dentate gyrus. *Proc Natl Acad Sci U S A*. 2007; 104:5638–5643. [PubMed: 17374720]
- Perlberg V, Bellec P, Anton JL, Pelegrini-Issac M, Doyon J, Benali H. CORSICA: correction of structured noise in fMRI by automatic identification of ICA components. *Magn Reson Imaging*. 2007; 25:35–46. [PubMed: 17222713]

- Petersson KM, Nichols TE, Poline JB, Holmes AP. Statistical limitations in functional neuroimaging. I. Non-inferential methods and statistical models. *Philos Trans R Soc Lond B Biol Sci.* 1999; 354:1239–1260. [PubMed: 10466149]
- Raichle ME, Snyder AZ. A default mode of brain function: a brief history of an evolving idea. *Neuroimage.* 2007; 37:1083–1090. [PubMed: 17719799]
- Raj D, Anderson AW, Gore JC. Respiratory effects in human functional magnetic resonance imaging due to bulk susceptibility changes. *Phys Med Biol.* 2001; 46:3331–3340. [PubMed: 11768509]
- Smith SM. Fast robust automated brain extraction. *Hum Brain Mapp.* 2002; 17:143–155. [PubMed: 12391568]
- Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TE, Johansen-Berg H, Bannister PR, De Luca M, Drobnjak I, Flitney DE, Niazy RK, Saunders J, Vickers J, Zhang Y, De Stefano N, Brady JM, Matthews PM. Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage.* 2004; 23 Suppl 1:S208–S219. [PubMed: 15501092]
- Stevens MC, Kiehl KA, Pearlson G, Calhoun VD. Functional neural circuits for mental timekeeping. *Hum Brain Mapp.* 2007; 28:394–408. [PubMed: 16944489]
- Stone JV, Porrill J, Porter NR, Wilkinson ID. Spatiotemporal independent component analysis of event-related fMRI data using skewed probability density functions. *Neuroimage.* 2002; 15:407–421. [PubMed: 11798275]
- Sui J, Adali T, Pearlson GD, Calhoun VD. An ICA-based method for the identification of optimal FMRI features and components using combined group-discriminative techniques. *Neuroimage.* 2009; 46:73–86. [PubMed: 19457398]
- Thomas CG, Harshman RA, Menon RS. Noise reduction in BOLD-based fMRI using component analysis. *Neuroimage.* 2002; 17:1521–1537. [PubMed: 12414291]
- Tohka J, Foerde K, Aron AR, Tom SM, Toga AW, Poldrack RA. Automatic independent component labeling for artifact removal in fMRI. *Neuroimage.* 2008; 39:1227–1245. [PubMed: 18042495]
- Wager TD, Davidson ML, Hughes BL, Lindquist MA, Ochsner KN. Prefrontal-subcortical pathways mediating successful emotion regulation. *Neuron.* 2008; 59:1037–1050. [PubMed: 18817740]
- Wang Z, Peterson BS. Partner-matching for the automated identification of reproducible ICA components from fMRI datasets: algorithm and validation. *Hum Brain Mapp.* 2008; 29:875–893. [PubMed: 18058813]
- Windischberger C, Langenberger H, Sycha T, Tschernko EM, Fuchsjaeger-Mayerl G, Schmetterer L, Moser E. On the origin of respiratory artifacts in BOLD-EPI of the human brain. *Magn Reson Imaging.* 2002; 20:575–582. [PubMed: 12467863]
- Woolrich MW, Jbabdi S, Patenaude B, Chappell M, Makni S, Behrens T, Beckmann C, Jenkinson M, Smith SM. Bayesian analysis of neuroimaging data in FSL. *Neuroimage.* 2009; 45:S173–S186. [PubMed: 19059349]
- Woolrich MW, Ripley BD, Brady M, Smith SM. Temporal autocorrelation in univariate linear modeling of FMRI data. *Neuroimage.* 2001; 14:1370–1386. [PubMed: 11707093]
- Worsley, KJ. Statistical analysis of activation images, Chapter 14. In: Jefferard, P.; Matthews, PM.; Smith, SM., editors. *Functional MRI: an introduction to methods.* Oxford: Oxford University Press; 2004. p. 251–270.
- Zhang Y, Brady M, Smith S. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Trans Med Imaging.* 2001; 20:45–57. [PubMed: 11293691]
- Zou Q, Wu CW, Stein EA, Zang Y, Yang Y. Static and dynamic characteristics of cerebral blood flow during the resting state. *Neuroimage.* 2009; 48:515–524. [PubMed: 19607928]

Fig. 1.
Ring pattern considered mainly due to structured noise.

Fig. 2.
Tightly clumped frontal cluster mainly in fringe areas. Timecourse shows sawtooth pattern.

Fig. 3.
Large rostral cluster. Timecourse shows deep “spike.”

Fig. 4.
Spotty, diffusely distributed pattern over most of the brain. Most of the power spectrum lies above 0.1 Hz.

Fig. 5.
Spotty pattern with small clusters in grey matter and in the superior sagittal sinus.

Fig. 6.
Image showing some grey matter clumps consistent with brain function (default network).

Fig. 7.
Group ICA spatial pattern consistent with location of white matter.

Fig. 8.
Group ICA component with thresholded voxels tightly clustered in predominantly grey matter regions.

Fig. 9.

Synthetic components: 1% of the mean MR signal (individually for each voxel) times the outer products of the above time courses and spatial maps was added to the “componentless” fMRI data. The maps and time courses correspond to N-IC1 (a), N-IC2 (b), N-IC3 (c), S-IC1 (d), S-IC2 (e), and S-IC3 (f), described in the text.

Fig. 10. IC spatial maps and time courses corresponding to the seeded ICs, after high-pass temporal filtering, spatial smoothing, and ICA: N-IC1 (a), N-IC2 (b), N-IC3 (c), S-IC1 (d), S-IC2 (e), and S-IC3 (f).

Fig. 11. IC spatial maps and time courses derived from a repeat ICA after filtering out the N-ICs. The N-ICs were no longer represented, while the S-ICs (a-c) were nearly identical to those from the ICA prior to filtering out the N-ICs.

Fig. 12.

Task-related GLM spatial maps and time courses for S-IC2 (a–b) and S-IC3 (c–d) before (a, c) and after (b, d) filtering out the N-ICs. The blue lines show the time course regressor for each GLM analysis; and the red lines, the time course of the voxel with best fit to each model. The spatial maps reveal more of the seeded voxels after denoising.