Reduction of Motion Artifacts and Noise Using Independent Component Analysis in Task-Based Functional MRI for Preoperative Planning in Patients with Brain Tumor


AJNR Am J Neuroradiol 2017, 38 (2) 336-342
doi: https://doi.org/10.3174/ajnr.A4996
http://www.ajnr.org/content/38/2/336
ABSTRACT

BACKGROUND AND PURPOSE: Although it is a potentially powerful presurgical tool, fMRI can be fraught with artifacts, leading to interpretive errors, many of which are not fully accounted for in routinely applied correction methods. The purpose of this investigation was to evaluate the effects of data denoising by independent component analysis in patients undergoing preoperative evaluation for glioma resection compared with more routinely applied correction methods such as realignment or motion scrubbing.

MATERIALS AND METHODS: Thirty-five functional runs (both motor and language) in 12 consecutive patients with glioma were analyzed retrospectively by double-blind review. Data were processed and compared with the following: 1) realignment alone, 2) motion scrubbing, 3) independent component analysis denoising, and 4) both independent component analysis denoising and motion scrubbing. Primary outcome measures included a change in false-positives, false-negatives, z score, and diagnostic rating.

RESULTS: Independent component analysis denoising reduced false-positives in 63% of studies versus realignment alone. There was also an increase in the z score in areas of true activation in 71.4% of studies. Areas of new expected activation (previous false-negatives) were revealed in 34.4% of cases with independent component analysis denoising versus motion scrubbing or realignment alone. Of studies deemed nondiagnostic with realignment or motion scrubbing alone, 65% were considered diagnostic after independent component analysis denoising.

CONCLUSIONS: The addition of independent component analysis denoising of fMRI data in preoperative patients with glioma has a significant impact on data quality, resulting in reduced false-positives and an increase in true-positives compared with more commonly applied motion scrubbing or simple realignment methods.

ABBREVIATIONS: BOLD = blood oxygen level–dependent; ICA = independent component analysis; DVARS = root-mean-square of the derivatives of the differentiated time courses of every brain voxel for each acquired volume; MD = mean displacement; TC = task correlation.
It is essential to understand that these typical rigid-body motion-correction algorithms do not account for significant variance in the data produced by myriad motion-related artifacts.\footnote{E. H. M.} Two major culprits are those related to a changing susceptibility profile in the scanner and the effect of the spatial location of spin magnetization saturation, more commonly referred to as spin-history artifacts.\footnote{I. S. T.} Likewise, numerous other prospective and retrospective methods designed to account for motion-, physiologic-, and scanner-related effects have shown limited improvement.\footnote{I. S. T. M.}

The use of independent component analysis (ICA) for removing the nuisance effects in fMRI data has grown in popularity.\footnote{I. S. T. M.}

Fundamentally, ICA separates a mixture of signals in a dataset into its individual signal components. ICA is commonly explained with the analogy of a cocktail party, in which the recorded, indistinct sound from numerous conversations can be deconstructed into individual voices. ICA denoising has been shown to increase statistical significance and sensitivity, resulting in fewer false-positives and false-negatives.\footnote{I. S. T. M.}

The goal of our study was to evaluate the role of ICA denoising on fMRI data of patients undergoing preoperative planning for brain tumor resection.

\section*{MATERIALS AND METHODS}

The requirement for informed consent was waived in this Health Insurance Portability and Accountability Act–compliant retrospective study, which was approved by the University of Florida institutional review board. Thirty-five runs from 12 consecutive patients undergoing preoperative planning for glioma resection were retrospectively reviewed. Various language (semantic decision \cite{6/35}, sentence completion \cite{3/35}, word generation \cite{5/35}, picture naming \cite{3/35}) and motor (finger \cite{6/35}, foot \cite{4/35}, tongue \cite{2/35}, face \cite{6/35}) paradigms were used. See On-line Table 1 for details of the task designs. All patients were trained on the tasks by the administering physician (E. H. M.).

\section*{Image Acquisition}

MR imaging data were acquired on a 3T Verio scanner (Siemens, Erlangen, Germany) with a 12-channel head coil. In an attempt to minimize head motion, the head was tightly packed with moldable foam. The importance of reducing head motion was explained to patients. A volumetric T1-weighted MPRAGE sequence with a voxel size of $1 \times 1 \times 1$ mm, TR = 2530 ms, TE = 3.5 ms, TI = 1100 ms, and flip angle = $7^\circ$ was obtained. BOLD-EPI images used a $3.5 \times 3.5 \times 3.5$ mm voxel size with a 0.7-mm intersection gap, TR = 2000 ms, TE = 30 ms, and flip angle = $78^\circ$, with an oblique axial interleaved acquisition. All patients were monitored and scanned by the administering physician (E. H. M.).

\section*{Data Processing}

Echo-planar images were processed with the fMRIB Software Library (FSL, Version 5.0; \url{http://www.fmrib.ox.ac.uk/fsl}). Pre-processing of all datasets included realignment (Motion Correction using FMRIB’s Linear Image Registration Tool [MCFLIRT]; \url{http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/MCFLIRT}), section-timing correction, and spatial smoothing (5-mm full width at half maximum). The preprocessed data were then processed using 4 differing methods. First, the preprocessed, realigned data were processed without any further manipulation. Second, motion scrubbing was performed by regressing individual volumes that exceeded a strict threshold for excessive section-signal variation based primarily on mean displacement and root-mean-square of the derivatives of the differentiated time courses of every brain voxel for each acquired volume (DVARS), as described in Power et al.\footnote{E. H. M.} Third, ICA analysis was performed (Multivariate Exploratory Linear Optimized Decomposition into Independent Components [MELODIC]; \url{http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/MELODIC}), and nuisance components were manually identified by visual inspection of components by an experienced ICA user (E. H. M.) and were removed by using the fsl_regfilt command. The methodology of characterizing components as noise- or task-related signal is based on the methodology used in the excellent review presented in Kelly et al. Last, both ICA denoising and motion scrubbing were performed, as previously described.

Data were analyzed using the FSL General Linear Model (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/GLM) with cluster thresholding in FMRIB Expert Analysis Tool (FEAT; \url{http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FEAT}). Identical threshold levels were applied to all 4 processing methods to determine the effects of the 4 processing pipelines on statistical significance. Processed EPI data were then linearly registered to the T1-weighted structural images (FMRIB Linear Image Registration [FLIRT]; \url{http://www.fmrib.ox.ac.uk}). Data for all 4 processing methods were overlaid to detect differences in each methodology. Two neuroradiologists with functional imaging experience (E. H. M., I. S. T.) independently and blindly evaluated the overlaid data to assess the presence or absence of expected activation based on the task, amount of noise, and change in statistical significance ($z$ score variance) of the activated areas and whether any or all of the processing methods were considered diagnostic. Diagnostic studies were defined as showing all expected areas of activation for the given task regardless of tumor location, to minimize the subjectivity of the meaning of diagnostic versus nondiagnostic. The subjective measures of noise, new real areas of activation, and diagnostic ability for each method resulted in 560 individual binary data points, which were entered in to an assessment of reader agreement. Disagreements were settled by a third blinded reader (L. M. S.).

\section*{Data Analysis}

We extensively evaluated image quality and tabulated 3 primary motion parameters: mean displacement (MD), task correlation (TC), and DVARS. For MD, each functional run was placed in 1 of 3 categories from none/mild, moderate, to severe, as follows: <1-mm MD throughout the study, 1- to 2-mm MD in ≤4 spikes, or any MD of >2 mm or >1 mm for >4 spikes. Likewise, DVARS was categorized as the following: no spikes of >5% signal change from volume to volume, 1–5 spikes of >5%, or >5 spikes of >5%. TC was categorized as $r < 0.05$, $0.05 \leq r < 0.2$, or $r \geq 0.2$.

Due to the ordinal and nominal nature of the data, contingency analysis was used with motion parameters (ordinal) as independent measurements to evaluate their relationships to the nominal, binary (yes/no) response outcome measures. The Cochran-Armitage exact test for trend was appropriate to capture the power of the ordinal motion parameter measurements to assess statistical significance; the Fisher exact test was used when the motion parameter data were collapsed into binary variables.
RESULTS

The 2 readers’ independent evaluations agreed in 97% (546/560) of individual recorded data points; only 14 evaluations required a tie-breaker with a blinded third reader (On-line Table 2).

Motion Parameters

The individual motion parameters are illustrated in Fig 1. A severe rating was present in 1 of the 3 motion parameters in 60% (21/35) of studies and in >1 motion parameter in 17% (6/35) of studies. MD and DVARS were patient-specific: Half of patients had none/mild MD (χ² = 35.5, P = .034), while half of patients accounted for nearly all (17/20) moderate or severe DVARS (χ² = 39.3, P = .013). In contrast, task-correlated motion of >0.05 was observed in at least 1 paradigm in all 12 patients, with severe task-correlated motion distributed across 8 patients (χ² = 20.6, P = .547). While there were no statistical differences in motion parameters between language and motor studies, motor tasks accounted for 8 of the 12 severe TC motion studies (On-line Fig 1).

fMRI Statistical Analyses

ICA denoising resulted in demonstrable improvement in overall fMRI quality. ICA denoising decreased the level of noise (false-positives) in the activation maps in 63% (22/35) of cases compared with realignment alone. A reduction in noise after ICA denoising was not correlated with any individual motion parameter or combination thereof. ICA also improved the statistical significance in regions of expected activation in 71.4% (25/35) of cases compared with realignment or motion scrubbing alone. ICA improved fMRI statistics, as assessed by an increase in the z score in areas of expected activation, in almost all scans with high MD (7/8 and 3/3 cases of moderate and severe MD, respectively; z = −1.75, P = .040). Additionally, ICA improved fMRI statistics with moderate or high TC (z = −1.80, P = .036) and when MD and TC were considered together (z = −1.92, P = .049). New expected areas of activation (previous false-negatives) were present only after ICA denoising in 34.3% (12/35) of the datasets and were not correlated with any motion parameter. Motion scrubbing alone did not reveal new areas of expected activation (previous false-negatives) compared with realignment alone or ICA. The combination of ICA and motion scrubbing showed no difference from ICA alone.

Diagnostic Quality

Most important, ICA denoising improved the diagnostic value of the fMRI studies (Fig 2). After realignment alone, 10 of the 12 patients had at least 1 nondiagnostic scan (χ² = 14.9, P = .186; Fig 2, upper graph), with 3 patients having all nondiagnostic-quality scans. In all, 57% (20/35) of scans were deemed nondiagnostic when corrected with realignment only. ICA denoising rendered diagnostic 65% (13/20) of the previously nondiagnostic scans, increasing the overall percentage of diagnostic scans from 42.9% to 80.0% (χ² = 6.3, P < .001). Moreover, ICA denoising improved the diagnostic outcome for 11 of 12 patients, with 8 pa-
patients having all scans of diagnostic quality ($\chi^2 = 22.5, P = .021$; Fig 2, lower graph). Notably, all 7 scans (in 4 patients) that remained nondiagnostic after ICA denoising contained severe TC and/or DVARS ($z = 2.76, P = .004$). In contrast, motion scrubbing failed to improve any of these scans to the point of being diagnostic (Fig 2, second graph), which is in line with the failure of motion scrubbing to reveal any new areas of activation compared with motion correction alone.

Motion scrubbing alone increased the statistical significance in expected areas of activation in 20% (7/35) of the scans compared with realignment alone but only showed improved statistical significance versus ICA denoising in 1 case (with low MD, severe TC, and moderate DVARS). Motion scrubbing had the greatest effect in those scans with moderate-severe ratings in TC alone ($z = -1.75, P = .040$) or to a lesser extent when MD and TC were considered together ($z = -1.67, P = .054$). The addition of motion scrubbing to ICA denoising did not show any improvement over ICA denoising alone.

**DISCUSSION**

Based on the previously reported success of ICA in improving data quality in both task-based and resting-state fMRI studies, our goal was to investigate its utility in preoperative planning for patients undergoing resection for gliomas. We found improvement in statistical significance and decreased noise, resulting in fewer false-positives and, perhaps more important, an increase in true-positives, similar to findings in studies in patients without tumors.20,27,28 We also found that a proportion of studies, otherwise nondiagnostic when processed with traditional methods, were diagnostic after ICA denoising.

As previously discussed, underlying motion-related artifacts, physiologic noise, and scanner-related noise, which are ubiquitous in fMRI, have serious detrimental effects on fMRI interpretation. The first of these, subject motion, is becoming an increasingly recognized cause of fMRI misinterpretation. For example, a prominent theory in autism that touted decreased long-range connectivity was eventually shown to be related to increased head motion (Fig 2, second graph), which is in line with the failure of motion scrubbing to reveal any new areas of activation compared with motion correction alone.

Motion scrubbing alone increased the statistical significance in expected areas of activation in 20% (7/35) of the scans compared with realignment alone but only showed improved statistical significance versus ICA denoising in 1 case (with low MD, severe TC, and moderate DVARS). Motion scrubbing had the greatest effect in those scans with moderate-severe ratings in TC alone ($z = -1.75, P = .040$) or to a lesser extent when MD and TC were considered together ($z = -1.67, P = .054$). The addition of motion scrubbing to ICA denoising did not show any improvement over ICA denoising alone.

Motion scrubbing alone increased the statistical significance in expected areas of activation in 20% (7/35) of the scans compared with realignment alone but only showed improved statistical significance versus ICA denoising in 1 case (with low MD, severe TC, and moderate DVARS). Motion scrubbing had the greatest effect in those scans with moderate-severe ratings in TC alone ($z = -1.75, P = .040$) or to a lesser extent when MD and TC were considered together ($z = -1.67, P = .054$). The addition of motion scrubbing to ICA denoising did not show any improvement over ICA denoising alone.
increased radiofrequency pulses resulting in decreased temporal resolution. Furthermore, no single method accounts for all sources of noise present in the data.\textsuperscript{5,7,32,36}

Data-driven approaches, such as removal or regression of volumes with signal changes above a threshold (motion scrubbing) or regression of motion parameters, do not account for all sources of artifacts and have limited effect.\textsuperscript{5,25} The limited effect on quality improvement is evident when volumes are discarded on the basis of motion parameters alone, because motion artifacts can have serious detrimental effects, even with minor head motion (<1 mm).\textsuperscript{8} Thus, more thorough parameters are needed when using motion scrubbing, such as DVARS or frame-wise displacement, which can detect very subtle artifactual changes in signal.\textsuperscript{25} Furthermore, motion scrubbing results in decreased statistical power, particularly in data with numerous motion-affected volumes, while images with detrimental artifacts that do not meet the threshold for rejection still exist. Regression of the motion parameters can also have deleterious effects if the motion is correlated with the task and therefore must be used with extreme caution in task-based experiments.\textsuperscript{6} Even with these considerations, the overall improvement gained by these methods is minimal.\textsuperscript{5,12,25} This outcome is also likely due, in part, to the lack of accounting for physiologic and scanner-related noise.

There has been an increasing interest in the use of ICA to denoise fMRI data to improve the sensitivity and specificity of results. The methodology behind ICA is beyond the scope of this article (see Stone\textsuperscript{17} for an excellent introduction), but in short, ICA deconstructs an fMRI time-series into unique components (each having a spatial map and the corresponding time course), which are maximally spatially independent.\textsuperscript{19,22,28,37} These components can then be classified as either artifactual/noise versus task-related activity (see Kelly et al,\textsuperscript{26} for an excellent review). In addition to manual identification, numerous methods of automated classification have been described, primarily by using machine-learning algorithms, but a discussion of these methods is beyond the scope of this article.\textsuperscript{38–43} The artifactual components consist of a variety of physiologic processes, motion-related artifacts, scanner-related noise, aliasing, and so forth. Once identified, these artifactual components can be removed from the data set before statistical analysis, such as with the General Linear Model. The improved statistical modeling results in increased sensitivity and specificity, with reduction of false-positives and increased true-positives (Fig 3).\textsuperscript{20,27,28}

Previous studies evaluating the use of ICA in task-based presurgical fMRI have yielded mixed results.\textsuperscript{21,44,45} While these studies were the first to explore ICA in the presurgical setting, they relied on the extraction of a single component from the ICA analysis results that best modeled the task. This approach can be problematic because task-related activation can often be distributed across multiple components.\textsuperscript{46} An approach rendering the most accurate and reliable results is to remove those components reflecting structured or random noise during preprocessing and then to process the denoised dataset with statistical modeling. Additionally, these studies were performed at 1.5T, in which physiologic noise and susceptibility-related changes are substantially lower than at 3T, so the results may not be directly applicable to higher field strengths.

The results of this study highlight the benefits of ICA denoising in presurgical mapping for tumor resection, even at higher field strengths. The reduction in false-positives related to these nuisance variables also reduces the number of surgical false-positives. Furthermore, this reduction has the effect of providing the surgeon with “cleaner” activation maps. Additionally, we have demonstrated that in some cases, true-positive activation was only identified with the use of ICA denoising, the lack of which could have had detrimental effects on the operation (Fig 4). The
appearance of otherwise undetected activation is likely due to improved statistical significance achieved by the removal of overlapping nuisance signal or artifacts in the area of true activation (Fig 5). This effect was most dramatic in patients with a higher incidence of motion-induced artifacts. However, when motion was correlated with the task, the motion-induced signal changes were generally indistinguishable from the task itself and any improvement in data quality with ICA denoising, as well as motion scrubbing, was often negligible. Last, we have shown an overall improvement in statistical significance of activated voxels with the use of ICA denoising.

Notable limitations of the current study include its retrospective nature. Because limited evidence exists on the effects of ICA denoising in the surgical setting, we do not believe that the prospective use of this technique is yet warranted. In this retrospective design, a strong inherent bias is present when assessing whether the results would alter patient treatment or decision-making, and as such, this assessment was not included in the analysis. Additionally, precise intraoperative mapping data were not collected in a structured fashion; this omission limits the ability to directly correlate with surgical data in all patients. Likewise, an inherent limitation of many surgical mapping studies is the inability to test all activation sites from limited craniotomies. To address these limitations, we relied on a double-blind expert review of activation maps to assess activation in expected areas on the basis of the task administered. We believe these limitations do not detract from the objective of this study in identifying data-quality improvement with ICA denoising relative to other standard methods of fMRI data denoising. Future prospective studies should be aimed at evaluating the impact on decision-making and patient outcomes.

CONCLUSIONS

The addition of ICA denoising of fMRI data in preoperative patients with glioma has a significant impact in data quality, resulting in reduced false-positives and an increase in true-positives compared with more commonly applied motion scrubbing or simple realignment methods.

REFERENCES

Isolated assessment of translation or rotation severely underestimates the effects of subject motion in fMRI data. PlOse One 2014;9(e106498) CrossRef Medline


