Comparison of Independent Component Analysis and Conventional Hypothesis-Driven Analysis for Clinical Functional MR Image Processing

Michelle A. Quigley, Victor M. Haughton, John Carew, Dietmar Cordes, Chad H. Moritz and M. Elizabeth Meyerand

AJNR Am J Neuroradiol 2002, 23 (1) 49-58
http://www.ajnr.org/content/23/1/49
Comparison of Independent Component Analysis and Conventional Hypothesis-Driven Analysis for Clinical Functional MR Image Processing

Michelle A. Quigley, Victor M. Haughton, John Carew, Dietmar Cordes, Chad H. Moritz, and M. Elizabeth Meyerand

BACKGROUND AND PURPOSE: With independent component analysis (ICA), regions of activation can be identified on functional MR (fMR) images without a priori knowledge of expected hemodynamic responses. The purpose of this study was to compare the results of fMR imaging data processed with spatial ICA with results obtained with conventional hypothesis-driven analysis.

METHODS: Eleven patients with focal cerebral lesions and one with agenesis of the corpus callosum were enrolled. Each patient performed text-listening, finger-tapping, and word-generation tasks. Conventional activation maps were generated by fitting time courses of each voxel to a boxcar reference function. Maps were created from the same image data with ICA techniques. To compare the maps, a concurrence ratio (CR) (number of voxels activated on both maps to number of voxels activated on either map) was calculated.

RESULTS: In the ICA analysis, maps with appropriate spatial and temporal features for auditory, sensorimotor, or language cortices were identified in most patients. Images processed with ICA resembled images processed with conventional means. In patients who moved or performed the task incorrectly, ICA produced a map that resembled the expected activation pattern but differed from the conventional image. CRs averaged 70% for all comparisons in the 12 patients.

CONCLUSION: fMR imaging maps for auditory, sensorimotor, and language tasks produced with ICA and conventional techniques were similar in most cases. Differences were consistent with the particular characteristics of the method. In data sets corrupted by motion or incorrect task performance, ICA may produce more accurate maps.

Independent component analysis (ICA) is a blind source separation method that does not require knowledge of the expected hemodynamic response to identify activation with functional MR (fMR) imaging data sets. It can be used to identify task-related components in the data that are due to different spatial hemodynamic responses without the necessity of assuming a reference function, as in conventional regression analyses (1–3). ICA, as formulated by Bell and Sejnowski (1), is a method of factoring measured signals into a set of signals that are statistically independent by using a feed-forward neural network to estimate the independent components in the data. ICA can be used to effectively identify activated voxels in synthetic fMR imaging data sets and data sets acquired in healthy subjects (4, 5). However, to our knowledge, ICA has not been extensively tested in patients.

The purpose of this study was to compare fMR images processed with spatial ICA and fMR imaging processed with conventional hypothesis-driven techniques in a series of patients referred by neurosurgeons for presurgical mapping. The advantage of ICA over conventional hypothesis-driven postprocessing methods is that it can be used to distinguish hemodynamic responses even if the response is not known. It may allow distinction of responses that do not approximate or are not fully correlated with the expected response. Therefore, in patients who do not

Received October 1, 1999; accepted after revision August 22, 2001.

From the Departments of Radiology (M.A.Q., V.M.H., C.H.M.) and Medical Physics (J.C., D.C., M.E.M.), University of Wisconsin, Madison.

Address reprint requests to M. Elizabeth Meyerand, PhD, Department of Medical Physics, 1530 Medical Sciences Center, 1300 University Avenue, Madison, WI 53706-1532.

© American Society of Neuroradiology
remained still during data acquisition or perform the tasks according to instructions. ICA may provide satisfactory maps when conventional hypothesis-driven methods fail.

Methods

For this study, data acquired in 12 patients referred for fMR imaging for neurosurgical planning were retrospectively analyzed. Functional imaging was performed with a 1.5-T commercial imager equipped with high-speed gradients for echo-planar (EP) imaging. High-resolution anatomic images were obtained with multisection spin-echo sequences. For fMR imaging, a series of images was acquired in the coronal, axial, or sagittal plane, as the radiologist (V.M.H.) prescribed. Technical parameters for these acquisitions included the following: 18 sections, 64 × 64 matrix, 90° flip angle, 2000/50 (TR/TE), 24-cm field of view, 7-mm section thickness, and 1-mm intersection gap. Each patient performed three tasks: passive text listening, intermittent finger tapping, and silent word generation, according to a standard on-off block-type paradigm. For each paradigm, four or five epochs of tasks were interspersed with epochs of rest. These tasks have been described elsewhere in detail (6–10).

Several filtering processes were applied to the data. Prior to reconstruction, the raw EP data were filtered by using a Hamming filter in the spatial-frequency domain to increase the signal-to-noise ratio (SNR) (11, 12). Voxels outside the brain were excluded from the analysis by eliminating those voxels on the EP image that had signal intensities below a prescribed threshold. In the remaining voxels, the temporal offset due to the order of section acquisition was corrected by means of linear interpolation. The Analysis and Visualization of Functional Neuroimages (AFNI) 3D registration motion correction (13) was applied to the data to attempt to control for random and systematic patient motion.

For conventional processing of image data, the voxel time courses were analyzed with a linear regression model. The model parameters were estimated with a locally developed least-squares-fit algorithm. The observed data were compared voxel by voxel with constant (baseline signal intensity), ramp (signal drift), and boxcar (idealized expected response to the task or stimulus) functions. The boxcar function was convolved with a Poisson function and had a unit amplitude and a period matching each of the on-off cycles of the task or stimulus. A 6-s lag was incorporated into the reference function to account for the expected hemodynamic delay. The statistical significance of the estimated parameter for the boxcar reference function was assessed for each voxel with the Student t test. This t score was then converted to a z score for comparison with ICA results. The level of significance on the z maps was assessed with a test of their corresponding time courses. The null distribution for the test was estimated by randomizing of the boxcar reference function with a nonparametric statistical method (14, 15). This corresponding P value was determined to be less than 0.0001 for a z score greater than 4. Voxels with a z statistic exceeding a threshold value of four were merged with coregistered anatomic images by using the AFNI display program (13).

For ICA, the code of Bell and Sejnowski (1) was used to separate the data into independent components. The resulting data matrix had approximate dimensions of 148 time frames × 14,000 voxels. The spatial maps of each component were overlaid on coregistered 3D volume anatomic datasets. Standard parcellation methods were used to identify the structures in which activation was observed. The spatial ICA component maps were ranked according to z score. The maps were then inspected to select those that best conformed to the expected spatial patterns of activation. The ranking scheme, according to z score, limited the number of components that needed visual inspection. The temporal characteristics of the selected ICA components were then correlated to the reference function to verify that it was temporally related to task performance. The components that had temporal characteristics unrelated to the reference function were classified as artifacts, presumably due to head motion or pulsatile flow artifacts.

Concurrence between the maps processed with the hypothesis-driven method and the blind source separation method was measured by creating an intersect map that showed only those voxels with a significant z score with both methods. A concurrence ratio (CR) was then calculated; it was equal to the number of these concurrent voxels in the intersect map divided by the average of the number of voxels that were independently activated with each of the two methods. CRs were then expressed as percentages.

For each task in each patient, CRs were calculated for each section, and the section with the maximum CR was identified. Also, for each task in each patient, the average CR for all sections with activation in the eloquent cortex was tabulated. In some comparisons, averages were computed for as many as nine or 10 sections with activation in the auditory or motor cortices. In other comparisons, such as that involving the language task, as few as one or two sections were available for averaging. Additionally, overall average CRs were calculated for each patient and all patients.

Results

fMR imaging activation maps of good technical quality were acquired with the conventional hypothesis-driven method; all but two image sets had evidence of patient motion or other corrupting artifacts. These conventional maps showed activation in the superior temporal lobes (text listening); sensori-motor cortices, supplementary motor areas, and cerebellum (finger tapping); and left inferior frontal gyrus (word generation). Examples of each of these eloquent areas are depicted in Figures 1A, 2A, and 3A. Other regions with activation were identified in some subjects.

In each of the 12 patients, one or more of the ICA components produced an activation map corresponding to the expected activation pattern for auditory and motor tasks. In four patients, ICA could not be used to identify a component with the appropriate spatial and temporal characteristics of an activation pattern for language. Examples of selected ICA components for auditory, motor, and language tasks are presented in Figures 1B, 2B, and 3B. Multiple components were mapped and inspected in each task. In most cases, the component with the task activation was among the 20 with the highest z score. Of the multiple components generated with the ICA algorithm, only a few of those identified had maps of eloquent cortical regions.

The maps produced with the conventional and ICA processing methods were similar in most cases. The similarity was demonstrated on the intersect maps (Figs 1C, 2C, and 3C), which show the concurrent pixels activated with both the conventional and ICA maps. The time courses in voxels showing activation with the conventional processing method closely resembled the temporal characteristics of an independent component, which mapped to the appropriate eloquent cortex. The time courses of representative voxels in eloquent cortices (Figs 1A, 2A, and 3A) are shown in Figures 1D, 2D, and 3D. The temporal pattern of the selected ICA components (Figs 1B, 2B,
and 3B) are depicted in Figures 1E, 2E, and 3E. The ICA components that temporally resembled the hemodynamic response correlated with the reference function, with correlation coefficients of .70, .64, and .38 in the cases illustrated. The use of a reference function is not necessary as part of the ICA. We included a correlation to a reference function simply as a method of verifying that the ICA component chosen on the basis of spatially expected patterns also was similar in the temporal domain.

In two patients, the data acquired during the alternating finger-tapping task had evidence of corruption. In one case, the conventional map showed activation in both left and right sensorimotor cortices (Fig 4A), although, with this task and reference function, activation typically is predominant in one hemisphere. Inspection of the time courses of voxels in the right and left motor cortices revealed temporal patterns of signal that suggested incorrect performance of the finger-tapping task. The patient moved the right hand not only when instructed but also when instructed to move the left hand. ICA, however, revealed robust activation in only the contralateral sensorimotor cortex (Fig 4B), as expected.
In the other case, the data acquired during finger tapping was corrupted by task-correlated head motion (Fig 5A and C). Patient movement during the finger-tapping task was demonstrated on the motion-versus-time graph (Fig 5E), which revealed significant inferior-superior motion that was temporally related to the task periods. This motion led to notable false-positive findings in the conventional $z$ maps (Fig 5A and C). However, in the ICA analysis, task activation and head motion were separated into different independent components; therefore, the maps of the component related to finger tapping (Fig 5B and D) were not corrupted.

CRs in all patients in the single section with the highest CR for auditory, motor, and language cortices are listed in Table 1. In patient 1, both the left and the right auditory cortices had a 79% concurrence between the ICA and conventional $z$ score activation map; 67% and 79% concurrence was observed in the left and right motor cortices, respectively, and a 78% concurrence was observed in the Broca area. The overall concurrence for this patient was 76%. In other patients, the CRs ranged from a low of 25% in the Broca area in patient 3 to a high of 98% in the left and right auditory cortices in patient 2. CRs ranged from 55% to 98% for auditory cortices, 35% to 84% for motor cortices, and 25% to 80% for the Broca area, with combined averages of 79%, 68%, and 57%, respectively. Average ratios with auditory, motor, and language tasks ranged from 55% to 85% for each patient, while the overall CR for all tasks in all patients was 70%.

In Figures 6 and 7, multisection comparisons of auditory cortices are shown, and in Figures 8 and 9, the multisection comparisons of motor cortices are shown. The conventional, ICA, and intersect maps are shown for four representative cases. In most patients in this study, one component was identified in each auditory and language task that was spatially related to the appropriate eloquent cortex. In the motor task data for this group, however, components were sometimes split between two ICA components, with the left cortex activated in one component and the right cortex in another. One such case is illustrated in Figure 8. In this case, a component that mapped to only the left sensorimotor cortex and another that mapped to only the right sensorimotor cortex were found. The activation of the left (dominant) hemisphere in this left-hand finger-tapping task
differed from that of the right hemisphere (contralateral to the finger that was active). Because the dominant hemisphere had some activation with both contralateral and ipsilateral finger movements, different ICA components were found in the left and right hemispheres.

Table 2 lists the CRs determined when all the sections covering the relevant activated cortex were used in the calculation. With the auditory and motor data in all patients and with the language data in some patients, this calculation involved multiple sections. Ratios were calculated, on average, in auditory cortices (typically six contiguous coronal sections), motor cortices (typically three sections), and the Broca area (typically one section). CRs for this comparison ranged from a low of 25% in the Broca area in patient 3 to a high of 87% in the left and right auditory cortices in patient 6. Concordance ratios ranged from 34% to 87% in the auditory cortices, 30% to 75% in the motor cortices, and 25% to 80% in the Broca area, with combined averages of 61%, 56%, and 55%, respectively. Average ratios with auditory, motor, and language tasks ranged from 32% to 67% in each patient, while the overall CR for all tasks in all patients was 58%.

**Discussion**

These findings show concurrence between the activation maps produced with a conventional hypothesis-driven method and that produced with a blind source separation method. This overall concurrence is comparable to that of a first and second iteration of a task (ie, test-retest precision) (16). The findings further show that, in the case of task-related motion or improper performance of the task, ICA produced a map that was less severely confounded by artifact than those produced with the conventional method.

Major weaknesses of the ICA method at present are the lack of criteria for determining the physiologic importance of each component and the incomplete development and optimization of the method for clin-
ical use. In the ICA algorithm used in this study, components identified with ICA were ranked according to their $z$ scores, that is, according to the probability that the component did not merely represent a random series of values. We ranked independent components for mapping on the basis of their $z$ scores and then selected them on the basis of their spatial patterns. Thus, we examined the components in order of their $z$ score. The components related to the activation usually had the top $z$ scores. Other components (eg, related to motion of the head) also often had the top $z$ scores. We classified components as sensorimotor-, auditory-, or language-related on the basis of the spatial mapping pattern of the component. In addition, we calculated the correlation with the standard reference function as a means to verify that the component was likely related to a task-induced hemodynamic response and not to some other effect, such as motion. Notably, while this correlation to the task timing was effective in identifying task-related components, a reference function was not incorporated in the ICA algorithm. The components were produced on the basis of an assumption of statistical independence.

Selection of the components based on their spatial patterns introduced a bias in the study, because the investigator chose the components for mapping. Methods based on ICA, which would eliminate this source of bias, have been developed (17). They permit the selection of components based on both temporal and spatial patterns by using a hybrid application of data-driven analysis and an a priori reference function. Such programs may overcome the problem of observer bias in assigning physiologic importance to components identified with blind source separation methods. However, this reduction in bias reduces the data-driven sensitivity of pure ICA, because the hybrid method requires a reference function.

**TABLE 1: CRs in the comparison of conventional z maps and ICA maps determined with single sections**

<table>
<thead>
<tr>
<th>Patient No.</th>
<th>Auditory Cortex CR (%)</th>
<th>Motor Cortex CR (%)</th>
<th>Left Language Cortex CR (%)</th>
<th>Average CR (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Left</td>
<td>Right</td>
<td>Left</td>
<td>Right</td>
</tr>
<tr>
<td>1</td>
<td>79</td>
<td>79</td>
<td>67</td>
<td>79</td>
</tr>
<tr>
<td>2</td>
<td>98</td>
<td>98</td>
<td>81</td>
<td>81</td>
</tr>
<tr>
<td>3</td>
<td>92</td>
<td>92</td>
<td>83</td>
<td>83</td>
</tr>
<tr>
<td>4</td>
<td>84</td>
<td>84</td>
<td>76</td>
<td>76</td>
</tr>
<tr>
<td>5</td>
<td>86</td>
<td>86</td>
<td>67</td>
<td>82</td>
</tr>
<tr>
<td>6</td>
<td>91</td>
<td>91</td>
<td>35</td>
<td>52</td>
</tr>
<tr>
<td>7</td>
<td>76</td>
<td>76</td>
<td>84</td>
<td>71</td>
</tr>
<tr>
<td>8</td>
<td>74</td>
<td>74</td>
<td>57</td>
<td>57</td>
</tr>
<tr>
<td>9</td>
<td>90</td>
<td>90</td>
<td>67</td>
<td>67</td>
</tr>
<tr>
<td>10</td>
<td>61</td>
<td>61</td>
<td>66</td>
<td>66</td>
</tr>
<tr>
<td>11</td>
<td>62</td>
<td>62</td>
<td>52</td>
<td>37</td>
</tr>
<tr>
<td>12</td>
<td>55</td>
<td>55</td>
<td>63</td>
<td>63</td>
</tr>
<tr>
<td>Average</td>
<td>79</td>
<td>79</td>
<td>67</td>
<td>68</td>
</tr>
</tbody>
</table>

Note.—CRs are the highest CR determined from a section selected from those that showed activation in the relevant cortex. Combined averages in auditory, motor, and left language cortex were 79%, 68%, and 57%, respectively.

**Fig 6.** fMR images obtained with the text-listening task in a patient with a cortical dysplasia involving the left occipital lobe. (Images in the leftmost column are those in Figure 1, with the addition of the next three consecutive sections.)

*Top row.*, The $z$ score maps obtained with conventional hypothesis-driven analysis.

*Middle row.*, The $z$ score maps obtained with ICA show good agreement with those in the top row.

*Bottom row.*, The intersect map reflects an 87% CR for these four consecutive sections.
At times, ICA results in the identification of multiple components for activation with a task. For example, different components were identified in the left and the right hemispheres with the finger-tapping task in one patient, and often, more than one component for activation is identified in each hemisphere with finger tapping. For instance, when a right-handed person performs a left-hand task, the left sensorimotor cortex activation is less than that of the right sensorimotor cortex. The participation of the dominant hemisphere in ipsilateral motor tasks has been described previously (5). ICA can be used to...
find other components related to finger tapping in the basal ganglia and cerebellum, because the time courses of activation in the basal ganglia differ from those in the sensorimotor cortices (18). Ipsilateral cerebellum and supplementary motor areas usually are included in the same component as sensorimotor activation, because the duration of activation in these areas is comparable.

### TABLE 2: CRs in the comparison of conventional z maps and ICA maps determined with all sections

<table>
<thead>
<tr>
<th>Patient No.</th>
<th>Auditory Cortex CR (%)</th>
<th>Motor Cortex CR (%)</th>
<th>Left Language Cortex CR (%)</th>
<th>Average CR (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Left</td>
<td>Right</td>
<td>Left</td>
<td>Right</td>
</tr>
<tr>
<td>1</td>
<td>75 (7)</td>
<td>75 (7)</td>
<td>49 (2)</td>
<td>59 (3)</td>
</tr>
<tr>
<td>2</td>
<td>59 (8)</td>
<td>59 (8)</td>
<td>73 (4)</td>
<td>73 (4)</td>
</tr>
<tr>
<td>3</td>
<td>81 (6)</td>
<td>81 (6)</td>
<td>71 (5)</td>
<td>71 (5)</td>
</tr>
<tr>
<td>4</td>
<td>63 (6)</td>
<td>63 (6)</td>
<td>75 (2)</td>
<td>75 (2)</td>
</tr>
<tr>
<td>5</td>
<td>68 (10)</td>
<td>68 (10)</td>
<td>47 (3)</td>
<td>51 (3)</td>
</tr>
<tr>
<td>6</td>
<td>87 (4)</td>
<td>87 (4)</td>
<td>35 (1)</td>
<td>48 (2)</td>
</tr>
<tr>
<td>7</td>
<td>58 (5)</td>
<td>58 (5)</td>
<td>63 (2)</td>
<td>67 (2)</td>
</tr>
<tr>
<td>8</td>
<td>64 (5)</td>
<td>64 (5)</td>
<td>49 (2)</td>
<td>49 (2)</td>
</tr>
<tr>
<td>9</td>
<td>61 (8)</td>
<td>61 (8)</td>
<td>55 (4)</td>
<td>55 (4)</td>
</tr>
<tr>
<td>10</td>
<td>46 (7)</td>
<td>46 (7)</td>
<td>57 (4)</td>
<td>57 (4)</td>
</tr>
<tr>
<td>11</td>
<td>38 (5)</td>
<td>38 (5)</td>
<td>52 (1)</td>
<td>31 (2)</td>
</tr>
<tr>
<td>12</td>
<td>34 (6)</td>
<td>34 (6)</td>
<td>30 (4)</td>
<td>30 (4)</td>
</tr>
<tr>
<td>Average</td>
<td>61</td>
<td>61</td>
<td>55</td>
<td>56</td>
</tr>
</tbody>
</table>

Note.—CRs are the averages of all sections that showed activation in the relevant cortex. Data in parentheses are the number of sections used in the calculation. Combined averages in auditory, motor, and left language cortex were 61%, 56%, and 55%, respectively.
Compared with model-dependent methods, such as conventional correlation analysis, ICA can be used to distinguish functions on the basis of patterns in the data not a correlation of the data with an expected response or reference function. Therefore, it can be used to detect activation when the hemodynamic response is atypical or unexpected. It may demonstrate activation that is missed with model-dependent techniques when the actual hemodynamic response differs from the expected response. As these findings illustrate, conventional methods are more sensitive to corruption of the data caused by the improper performance of the task or by motion during the task than ICA.

ICA does not replace conventional methods. It requires more computation than do conventional methods. It fails to identify activation secondary to word generation more often; the reason for this is not known. Possibly, the hemodynamic responses in voxels within the expressive language regions differ sufficiently so that multiple independent components are generated, and no one component provides a map of the entire region. The conditions under which components are separated are not well known. Additional studies, possibly with synthetic datasets, are required to understand the effects of ICA on variations in the data.

Another reason for the inability of ICA to depict activation consistently in language areas may be the lack of uniformity of the blood oxygen level—dependent (BOLD) SNR across the brain. ICA, as applied in this study, is a spatial technique. Thus, spatial variations in the noise characteristics from one spatial location to another create a spatial bias for ICA. This bias is less significant for time-domain methods such as linear regression, because the BOLD response in each voxel is tested independently. We observed in our data a spatial variation in the SNR across the brain. Specifically, we found that the SNR was higher in the auditory and motor cortices than that in the frontal lobe regions. This observation may explain why the ICA method mapped language areas less robustly than did the regression analysis in this study. The difficulty in detecting language activation with ICA also may reflect the fact that simpler tasks (such as auditory or motor responses) are controlled and confirmed more easily than are more complex cognitive tasks (such as language production).

Additional studies are needed to determine if ICA has substantial advantages in clinical practice or research. While these findings suggest that ICA and conventional methods agree in most cases, more work is needed to determine if systematic differences exist. For ideal simulated fMR imaging responses, ICA was shown not to detect activation as well as traditional techniques such as regression models and Kolmogorov-Smirnov tests (19). However, the variety of abnormal fMR imaging responses is difficult to simulate in patients. ICA may perform better with the abnormal responses than with common responses at clinical fMR imaging. To address this issue, further clinical case studies are needed. By definition, the two methods are used to measure different characteristics of BOLD signal. Therefore, the two methods should not be expected to yield identical results. We would not necessarily expect CRs of 100%. Moreover, the CRs themselves, as calculated herein, were limited as a means of evaluating comparisons. First, a bias was introduced into the CR calculation, because the investigator selected a single threshold for both the ICA and conventional fMR imaging maps. Second, CR values could be ambiguous; for instance, it is possible for two vastly different comparison scenarios to lead to the same CR value. To address these and other potential concerns, a more complete comparison of the two methods may include an exhaustive average CR on the basis of every threshold value pair that is possible between the two methods. Such an exhaustive comparison was beyond the scope of this study. The interest here was to determine if the established techniques for mapping functional activation with each of the two methods produce similar results.

The possibility that ICA has greater accuracy than that of conventional methods when the actual hemodynamic response differs from the expected response needs additional testing. ICA has not been optimized yet for fMR image processing. Optimal matrix size, task paradigms, and technical parameters for ICA have not been determined yet. Additional studies should be performed, with an increased number and variety of patient data under varying conditions and with fMR imaging paradigms that allow further comparison of conventional regression analysis with ICA results.

Conclusion
In most cases, ICA maps of task activation, without the assumption of a specific hemodynamic response, were comparable to maps prepared with conventional methods in which the hemodynamic response is modeled a priori. In cases in which the patient performed the task incorrectly or moved during data acquisition, ICA seemed to provide images of better quality. Further investigation of ICA for clinical fMR imaging is warranted.

References
6. Hammeke TA, Yetkin FZ, Mueller WM, Morris GL III, and
Haughton VM. Functional magnetic resonance imaging of somatosensory stimulation. Neurosurgery 1994;35:677–681