

# Quality Checking an fMRI Group Result (art\_groupcheck)

Paul Mazaika, Feb. 24, 2009

A statistical parameter map of fMRI group analyses relies on the assumptions of the General Linear Model (GLM). The assumptions at the single subject level are that the noise is stationary and unbiased for every voxel of every subject. In the mixed level model, the group level GLM assumes that all subjects are drawn randomly from a population, and that the estimates brought up from the single subject level have equal variances.

The validity of the statistical maps depends on how well those assumptions are satisfied. For a high motion subject, the assumptions may be violated because of biases from task-correlated motions and non-stationary noise caused by artifacts. These errors may lead to biased and high variance estimates for subjects that could not be adequately modeled. The art\_groupcheck program contains quality checks that allow a user to check whether these types of errors are present in the data.

The SPM interface allows a user to examine voxels at the peak locations in the group activation map for outliers. Ideally, all the estimates at a voxel should be similar.

- In SPM, open up Results for the group SPM.mat file.

- Bring up the clusters at the group level, and move the red cursor to one of them.

- In the lower left window, select the "Plot" button, then "Fitted Responses".

- Select the contrast of interest, "adjusted" and plot against scan or time.

- Look for consistent values for all subjects.

Some outliers may be very obvious from this check. If so, then the activation map at this cluster location would be more valid if the subject that produced the outlier were excluded from the group study.

However, this method requires a user to test the peak voxels of all the clusters, and it cannot find voxels that failed to be significant because of a negative outlier. Nor does this method review whether the subjects have zero bias or equal variance in their estimates.

The art\_groupcheck program tries to address these concerns.

## **RUNNING THE ART\_GROUPCHECK PROGRAM**

Start matlab with ml7spm5 to load the SPM paths. From matlab, start the program with:

```
>> art_groupccheck
```

Select a group level SPM.mat file. The program is compatible with SPM5 and SPM2.

The program uses units of percent signal change in order to more easily detect large values that are likely due to artifacts rather than cognitive BOLD effects. The scale factor into percent signal change is  $(\text{peak}/\text{contrast\_sum}) * 100 / \text{mean}$  (see the document on FMRI

Percent Signal Change). The program computes this scale factor for the contrast images of the first two subjects in the SPM.mat file, and prints the result as a suggested value X for the user.

Enter  $(\text{peak}/\text{contrast\_sum}) * 100 / \text{bmean}$ : X

The user can override the suggestion at this point. To compute a new value, check that the design matrix has the same peak for every column, and then use the `art_percentscale` program for each individual subject. Take the average of the scale factor results for all the subjects. Generally, the suggested value is accurate enough.

The next menu shows the three available options to review the group results:

View by:

1. Con and ResMS values for all subjects at one voxel location
2. Image of contrast result for every subject
3. Global Quality scores and suggested subject outliers

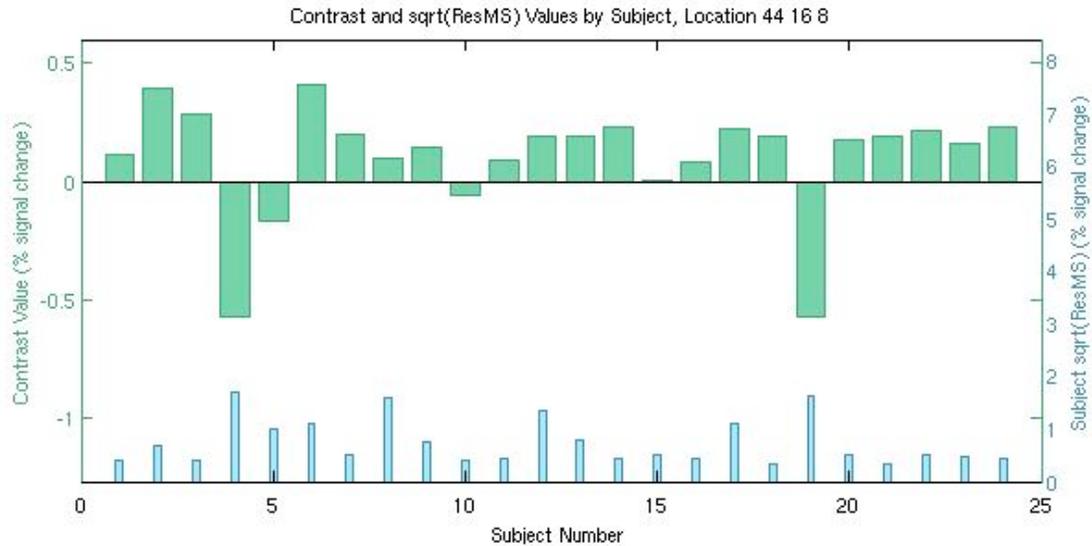
## **INTERPRETING THE OUTPUT**

### **1. Con and ResMS values for all subjects at one voxel location**

This option is similar to reviewing the subject values at a peak location with SPM, but with the capabilities to review ResMS values, quantitative sizes, and any voxel in the brain regardless of whether it is in a cluster.

The user chooses a point (x, y, z) in mm for which to view the results. It is best to have a list of locations to review, or have an image display open at the same time, in order to choose interesting locations. The program displays the contrast value and the  $\sqrt{\text{ResMS}}$  value at a voxel location for each subject (Figure 1). Values are in percent signal change. After one location has been checked, another (x,y,z) location can be viewed by entering the coordinates in the “Select New Coordinates” box.

The group analysis level in SPM uses only the contrast values. The figure shows two subjects (numbers 4 and 19) with extreme negative values on this voxel located at (44, 16, 8) in the right inferior frontal gyrus. While most subjects have values in the range [0 to 0.3], these subjects are 0.5 less than the low end of that range. These extreme values are outliers which might be detected by a statistical outlier criterion such as a Tukey test.



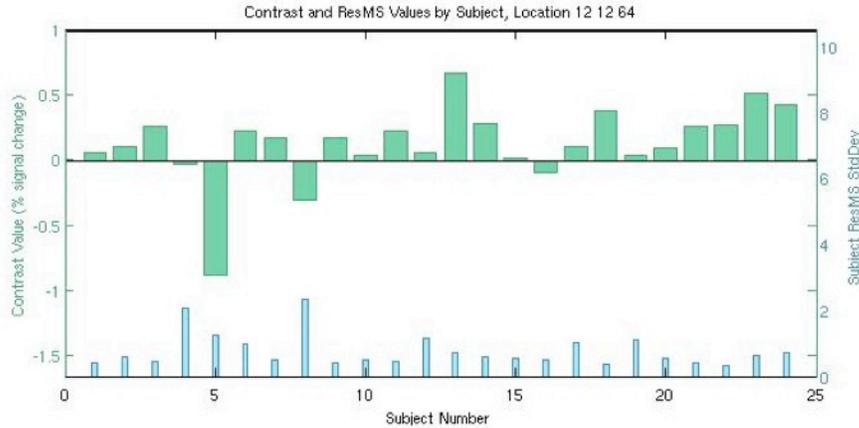
**Figure 1: Contrast and sqrt(ResMS) values at a single voxel in the inferior frontal gyrus. Units are percent signal change. Subjects 4 and 19 have outlier contrast values, while subjects 4, 8, 12, and 19 have abnormally large variance.**

For fMRI, we might also judge that a contrast as large in magnitude as 0.5% is unreasonable for the NoGo-Go contrast. Thus, we might also consider these points to be quantitative outliers relative to *a priori* expectations of the size of single subject results.

At the same time, there is also less confidence in the accuracy of these values because the ResMS values are high for these subjects. ResMS is the residual mean square error. The lower plot shows sqrt(ResMS) rather than ResMS itself, and is in units of percent signal change. The outlier subjects have sqrt(ResMS) values larger than 1%, whereas most of the subjects have errors ~0.2%.

Putting this evidence together at this voxel, subjects 4 and 19 appear to be distributional outliers from the Tukey test, quantitative outliers from the absolute magnitude test, and variance outliers from the ResMS test. Since all these results are only for a single voxel, it will be helpful to check to see if the same subjects are outliers on other voxels before these subjects would be excluded from a group analysis. Figure 2 shows the results for the same set of subjects at a different voxel located in the preSMA area (12, 12, 64).

Comparing Figures 1 and 2, it can be seen that different subjects may be indicated as outliers on different voxels. Thus, to choose outlier subjects it may be necessary to review the results on many voxels. The other menu options of the art\_groupcheck program allow a user to survey the results of numerous voxels in a faster way.



**Figure 2: Contrast and  $\sqrt{\text{ResMS}}$  values at a single voxel in the pre-SMA region of the superior frontal gyrus. Units are percent signal change. Subjects 5 and 13 have outlier contrast values, while subjects 4 and 8 have abnormally large variance.**

### ***The effect of non-equal ResMS values on statistical significance***

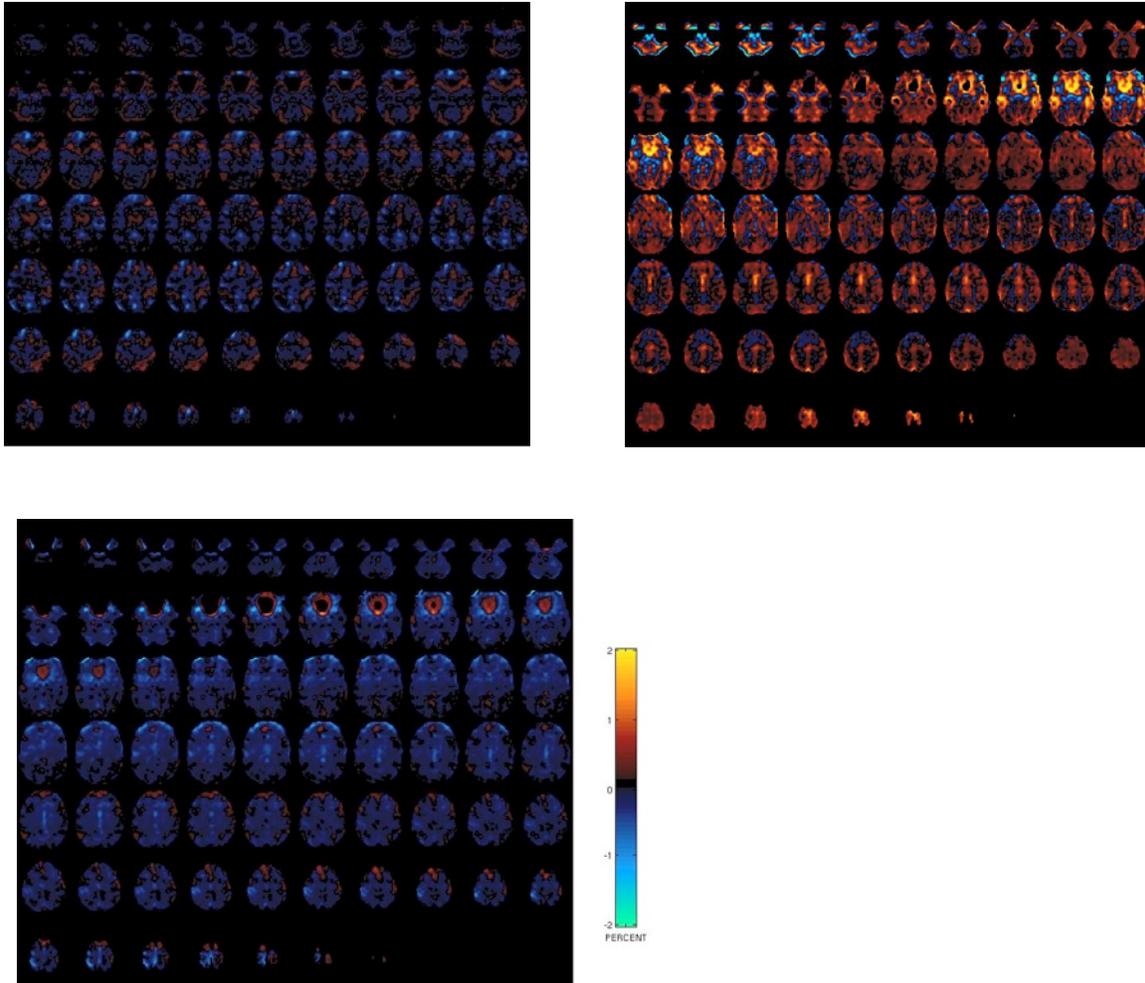
*Ideally, the contrast values will appear to be drawn from a normal distribution, and the  $\sqrt{\text{ResMS}}$  values will be identical across subjects. In practice, the  $\sqrt{\text{ResMS}}$  values will vary somewhat across subjects. If the variances vary widely, then the GLM may overestimate the statistical significance of the result, because the GLM assumes equal variances to count the number of degrees of freedom, and the number of degrees of freedom influences the statistical significance calculation. SPM assumes the variances are equal, and does not use the variances in the group level calculation. Other statistical analysis programs, such as FSL and fMRIstat, can use different single subjects variances in the group level calculation to improve the group results. Reference [1] is an introductory overview of these capabilities.*

*However, the ResMS value does not account for biases that may be in the estimates, so unequal variance is not a complete characterization of potential errors. In order to detect potential biases, it is helpful to review the contrast maps themselves, which is Option 2 of the art\_groupcheck program.*

*[1] Mumford, J.A. and Nichol, T. 2006. Modeling and Inference of Multisubject fMRI Data, IEEE Engr. in Medicine and Biology (25)#2, pp.42-51.*

## 2. Image of contrast result for every subject

This option allows a quick review of the contrast results on all voxels for each subject, and helps identify whether biased estimates may have occurred in certain brain regions. The user can choose to view the images in the axial, sagittal, or coronal planes. The contrast results for each subject are displayed as a montage image, and a user can choose a different subject by moving a slider at the base of the “movie” window. The contrast maps for three subjects are shown in Figure 3.



**Figure 3:** Contrast images of the NoGo-Go for three subjects. These images are *not* statistical activation maps. The contrasts are scaled into percent signal change, and all maps use the same scale ranging from (-2% to +2%). The upper left subject shows contrasts of the expected size (less than 1%), while the upper right subject has poor contrasts over the whole brain, and the lower subject has poor contrasts in the inferior frontal regions only.

The contrasts are scaled into percent signal change, shown with a scale from (-2% to +2%) shown at the right of each picture. Dark colors (deep red and deep blue) are the expected size of contrasts, while bright colors are likely to due task-correlated motion of artifacts in the data. These images allow a user to view the regional spatial context around a voxel to determine if some artifactual effect may have affected many voxels. Many subjects will have good estimates over the entire brain, but some may be poor over the whole brain, or in one region of the brain.

The contrast images provide a visual check that estimates from a subject are not corrupted in a particular region of the brain. (One note of caution: left and right are not preserved in these images). Generally, it will be advantageous to exclude subjects from a group analysis if they have bright artifactual responses over large regions in these displays. This choice of excluded subjects will tend to remove subjects with outlier values on many voxels, or equivalently, subjects whose estimates were heavily biased by non-cognitive effects. Thus, a GLM on the remaining subjects after excluding the outlier subjects will better reflect cognitive activations rather than artifacts in the population.

### **3. Global Quality scores and suggested subject outliers**

The third menu option will start the `art_groupoutlier` program, for which the details are described in the Outlier Protocol document. Whereas the contrast images are a pictorial review of the estimated contrasts on all the voxels, the Global Quality option makes summary statistics over the whole brain of the contrast estimates. Bright contrast images will show up as outlier subjects in the Global Quality scores from the `art_groupoutlier` program.

Note that all the single subjects involved in the group contrast will be analyzed with the same group mask when `art_groupoutlier` is started by this menu. When `art_groupoutlier` is started by itself, each subject will be analyzed with its own head mask. There is usually little difference in the results.