In [1]:
```
%config InlineBackend.figure_format = 'png'
```

In [2]:
```
import os
import shutil
import time

# prevent lengthy SPM output
from nipype.utils.logger import logging, logger, fmlogger, iflogger
#logger.setLevel(logging.getLevelName('CRITICAL'))
#fmlogger.setLevel(logging.getLevelName('CRITICAL'))
#iflogger.setLevel(logging.getLevelName('CRITICAL'))

import numpy as np
from scipy.stats.stats import pearsonr, spearmanr
from scipy.stats import wilcoxon
import sklearn as sk
from sklearn.linear_model.base import BaseEstimator, RegressorMixin
import sklearn.metrics as skm
import sklearn.cross_validation as cv
import matplotlib
#matplotlib.use('Agg')
#import matplotlib.pyplot as plt


#from nipype.utils.config import config
#config.enable_debug_mode()

import nipype.pipeline.engine as pe

from spm_2lvl import do_spm          #spm workflow --> give directory + co
from feature_selection import determine_model_all
from cluster_tools import get_clustermeans
from cfutils import get_subjects, get_subject_data
```

```
INFO:interface:stdout 2012-02-
19T12:51:42.237679:/software/matlab_versions/2010b/bin//matlab
```

In [3]:
```
X = get_subjects()
_, pdata = get_subject_data(X)
X = pdata.subject
y = pdata.lsas_pre - pdata.lsas_post
dcsidx = np.nonzero(pdata.classtype==2)[0]
pcbidx = np.nonzero(pdata.classtype==3)[0]
```

In [4]:
```
#wf = do_spm(X, y, analname='all_subjects', run_workflow=False)
#wf.base_dir = os.path.realpath('..')
#wf.run()
```

get cluster coordinates

In [5]:
```
def get_coords(img, affine):
    coords = []
    labels = np.setdiff1d(np.unique(img.ravel()), [0])
```

```
    cs = []
    for label in labels:
        cs.append(np.sum(img==label))
    for label in labels[argsort(cs)[::-1]]:
        coords.append(np.dot(affine,
                             np.hstack((np.mean(np.asarray(np.nonzero(img=
                                                axis = 1),
                                        1)))[:3].tolist())
    return coords
```

In [6]:
```
def print_table(img, affine, vox2vol=1., valimg=None):
    coords = []
    labels = np.setdiff1d(np.unique(img.ravel()), [0])
    cs = []
    for label in labels:
        cs.append(np.sum(img==label)*vox2vol)
    valmax = []
    csidx = argsort(cs)[::-1]
    cs = np.array(cs)[csidx]
    if valimg is not None:
        print "Cluster   \tX\tY\tZ\tk\tT"
    else:
        print "Cluster   \tX\tY\tZ\tk"
    for cidx, label in enumerate(labels[csidx]):
        xyz = np.dot(affine,
                     np.hstack((np.mean(np.asarray(np.nonzero(img==label))
                                        axis = 1),
                                1)))[:3]
        coords.append(xyz.tolist())
        if valimg is not None:
            idx = np.nonzero(img==label)
            maxval = np.max(np.abs(valimg[idx]))*np.sign(np.median(valimg[
            valmax.append(maxval)
            print "Cluster %02d\t%d\t%d\t%d\t%5d\t%.2f" % (cidx, xyz[0], x
        else:
            print "Cluster %02d\t%d\t%d\t%d\t%5d" % (cidx, xyz[0], xyz[1],
```

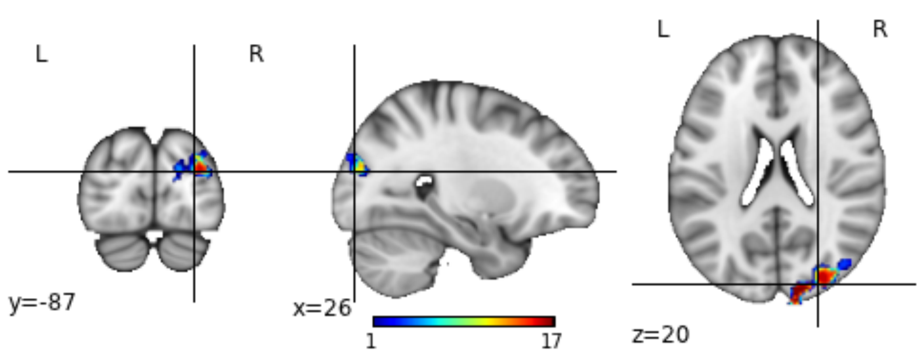
In [7]:
```
from nipy.labs import viz
from nibabel import load
def show_slices(img, coords=None, threshold=0.1, cmap=None, prefix=None,
                show_colorbar=None, formatter='%.2f'):
    if cmap is None:
        cmap = pylab.cm.hot
    data, aff = img.get_data(), img.get_affine()
    anatimg = load('/usr/share/fsl/data/standard/MNI152_T1_1mm_brain.nii.g
    anatdata, anataff = anatimg.get_data(), anatimg.get_affine()
    anatdata = anatdata.astype(np.float)
    anatdata[anatdata<10.] = np.nan
    outfile = 'cluster.svg'
    if prefix:
        outfile = '_'.join((prefix, outfile))
    outfile = os.path.join('figures', outfile)
    if coords is None:
        osl = viz.plot_map(np.asarray(data), aff, threshold=threshold,
                           cmap=cmap, black_bg=False)
```

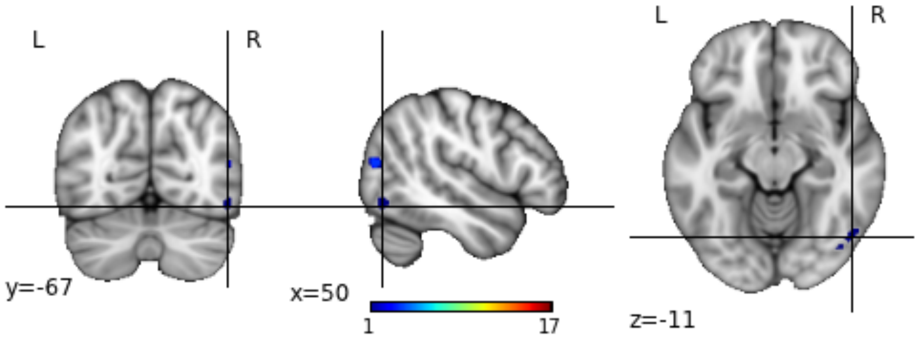
```
            osl.frame_axes.figure.savefig(outfile, transparent=True)
        else:
            for idx,coord in enumerate(coords):
                outfile = 'cluster%02d' % idx
                if prefix:
                    outfile = '_'.join((prefix, outfile))
                outfile = os.path.join('figures', outfile)
                osl = viz.plot_map(np.asarray(data), aff, anat=anatdata, anat_
                                   threshold=threshold, cmap=cmap,
                                   black_bg=False, cut_coords=coord)
                if show_colorbar:
                    cb = colorbar(gca().get_images()[1], cax=axes([0.4, 0.075,
                                  orientation='horizontal', format=formatter)
                    cb.set_ticks([cb._values.min(), cb._values.max()])
                    show()
                osl.frame_axes.figure.savefig(outfile+'.svg', bbox_inches='ti
                osl.frame_axes.figure.savefig(outfile+'.png', dpi=600, bbox_ir
```

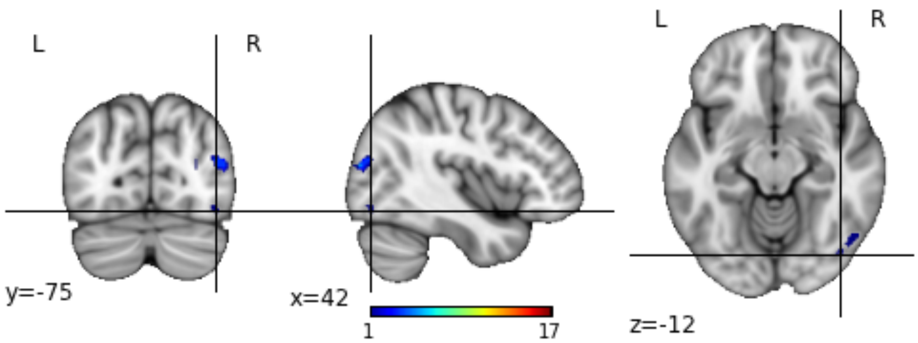
In [8]:
```
def plot_regression_line(x,y, xlim, color='r'):
    model=sk.linear_model.LinearRegression().fit(x[:,None],y)
    xplot = np.arange(xlim[0], xlim[1])[:,None]
    plot(xplot, model.predict(xplot), color=color)
```

In [9]:
```
import os
from scipy.ndimage import label
import scipy.stats as ss
def get_labels(data, min_extent=5):
    labels, nlabels = label(data)
    for idx in range(1, nlabels+1):
        if sum(labels==idx)<min_extent:
            labels[labels==idx] = 0
    return labels, nlabels
```

In [10]:
```
base_dir = '/mindhive/gablab/satra/sad/'
filename = os.path.join(base_dir, 'scripts', 'clustermean.nii.gz')
img=load(filename)
labels, nlabels = label(abs(img.get_data())>0)
coords = get_coords(labels, img.get_affine())
show_slices(img, coords, cmap=pylab.cm.jet, prefix='overlap', show_colorba
            formatter='%d')
```

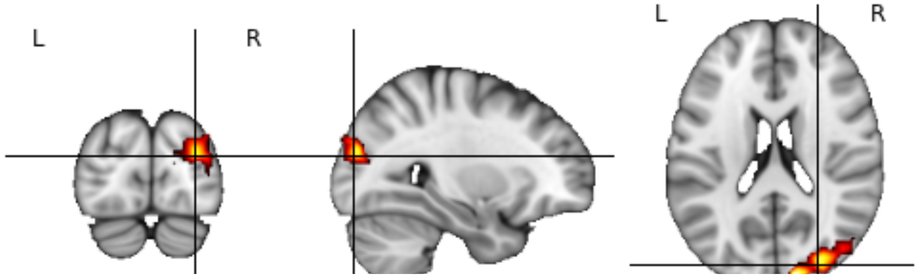




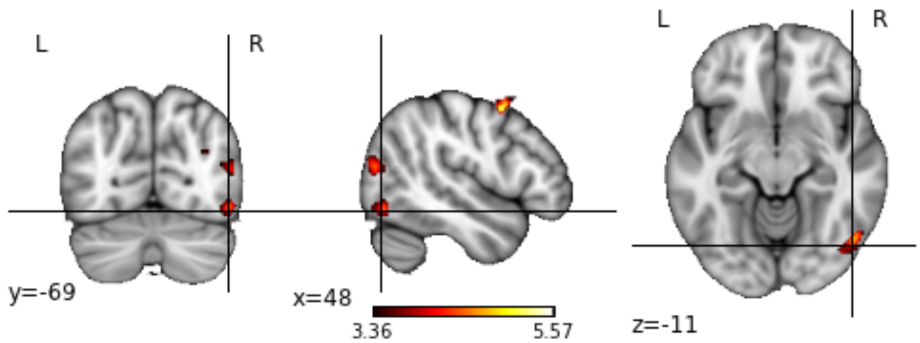


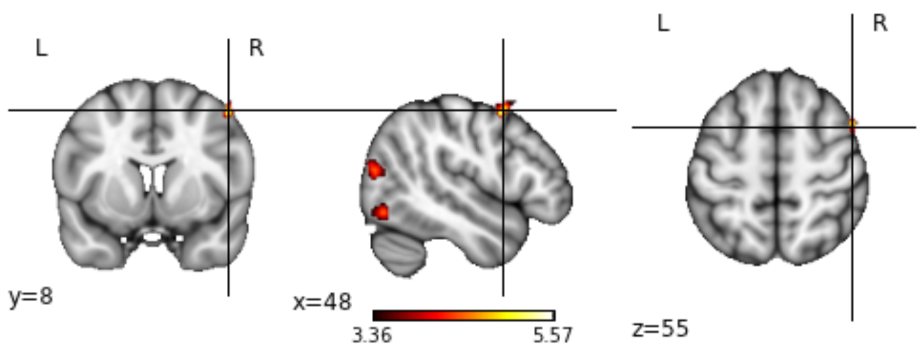
In [11]:
```
base_dir = '/mindhive/gablab/satra/sad/'
filename = os.path.join(base_dir, 'all_subjects', 'conest', 'spmT_0001.im
img=load(filename)
labels, nlabels = get_labels(img.get_data()>ss.t.ppf(1-0.001,33), 20)
data = img.get_data()
data[labels==0] = 0
#cmeans = get_clustermeans(X, labels, nlabels)
coords = get_coords(labels, img.get_affine())
print_table(labels, img.get_affine(),
            np.prod(img.get_header().get_zooms()), img.get_data())
show_slices(img, coords, threshold=0.5, prefix='uncorrected', show_colorba
```

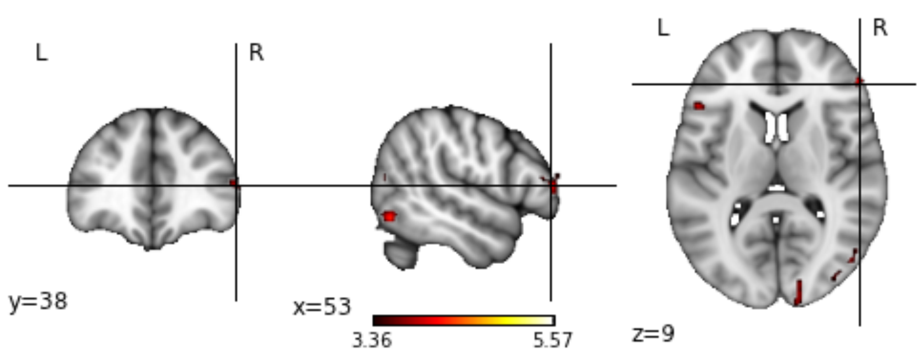
Cluster	X	Y	Z	k	T
Cluster 00	28	-85	19	9760	5.57
Cluster 01	48	-69	-11	1264	4.76
Cluster 02	48	8	55	496	5.21
Cluster 03	53	38	9	472	4.13
Cluster 04	-49	24	10	360	4.02
Cluster 05	-21	-61	-7	184	3.98
Cluster 06	-24	-77	30	160	3.48

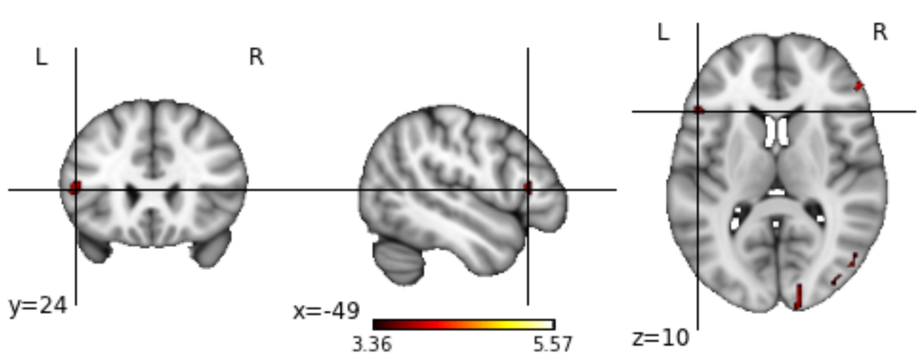


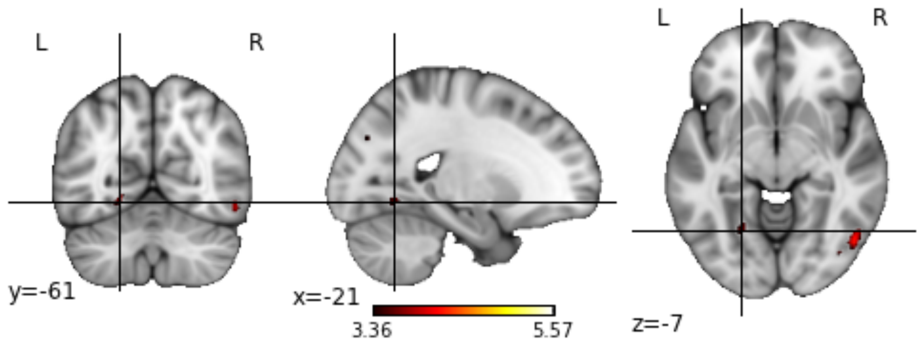

y=-85
x=28
3.36
5.57
z=19

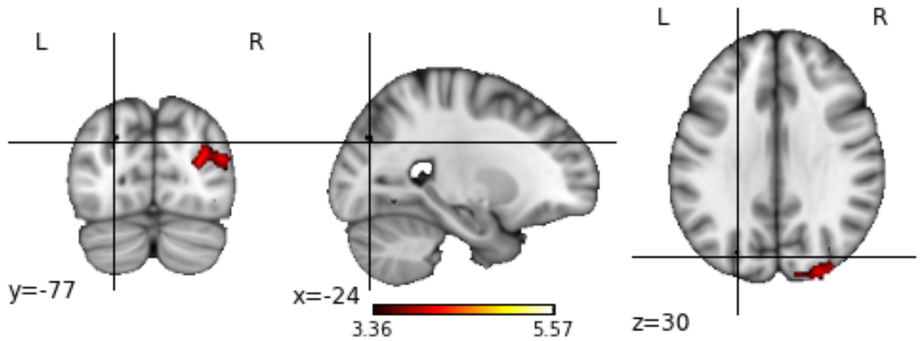

L
R
y=-69
x=48
3.36
5.57
L
R
z=-11


L
R
y=8
x=48
3.36
5.57
L
R
z=55


L
R
y=38
x=53
3.36
5.57
L
R
z=9

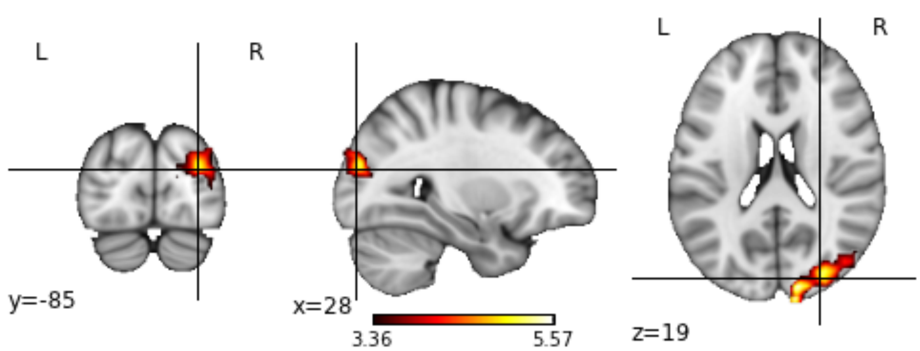

L
R
y=24
x=-49
3.36
5.57
L
R
z=10

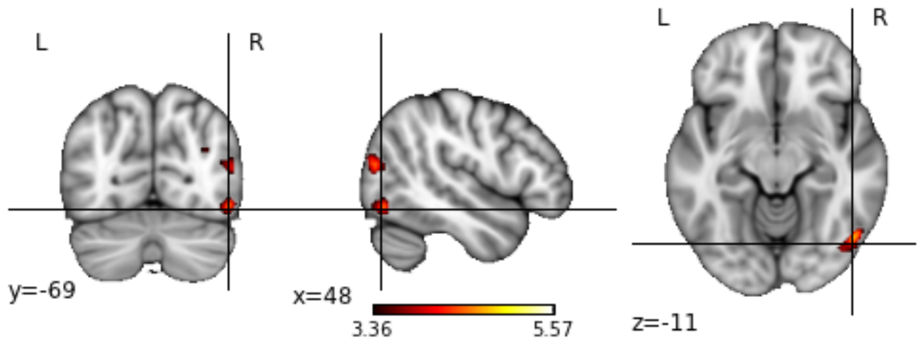




In [12]:
```
import os
from scipy.ndimage import label
base_dir = '/mindhive/gablab/satra/sad/'
filename = os.path.join(base_dir, 'all_subjects', 'thresh', 'spmT_0001_thr
img=load(filename)
labels, nlabels = label(abs(img.get_data())>0)
cmeans = get_clustermeans(X, labels, nlabels)
coords = get_coords(labels, img.get_affine())
print_table(labels, img.get_affine(),
            np.prod(img.get_header().get_zooms()), img.get_data())
show_slices(img, coords, prefix='topocorrect', show_colorbar=True)
```

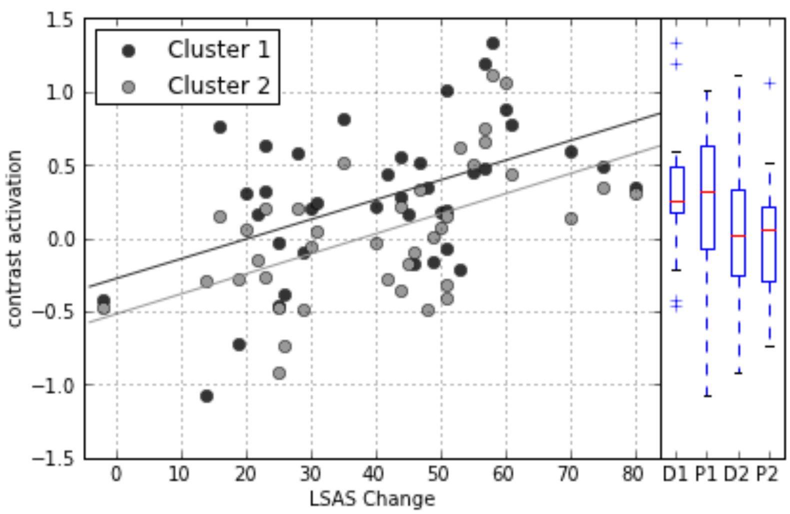
Cluster	X	Y	Z	k	T
Cluster 00	28	-85	19	9760	5.57
Cluster 01	48	-69	-11	1264	4.76






```
In [13]: close('all')
         axes([0.1,0.1,0.7,0.8])
         plot(y, cmeans[:,0], 'o', color=[0.2,0.2,0.2])
         plot(y, cmeans[:,1], 'o', color=[0.6,0.6,0.6])
         xlim([-5, 84])
         xlabel('LSAS Change')
         ylabel('contrast activation')
         legend(('Cluster 1', 'Cluster 2'), 'best', numpoints=1)
         plot_regression_line(y, cmeans[:,0], [-4,85], color=[0.2,0.2,0.2])
         plot_regression_line(y, cmeans[:,1], [-4,85], color=[0.6,0.6,0.6])
         grid()
         axes([0.8,0.1,0.15,0.8])
         boxplot([cmeans[dcsidx,0], cmeans[pcbidx,0], cmeans[dcsidx,1], cmeans[pcb:
         yticks([])
         xticks([1,2,3,4], ['D1', 'P1', 'D2', 'P2'])
         savefig('figures/scatter_means_all.svg',bbox_inches='tight', transparent=T
         savefig('figures/scatter_means_all.png', dpi=600,bbox_inches='tight', trar
         print 'r: C1', pearsonr(cmeans[:,0], y)
         print 'r: C2', pearsonr(cmeans[:,1], y)
```

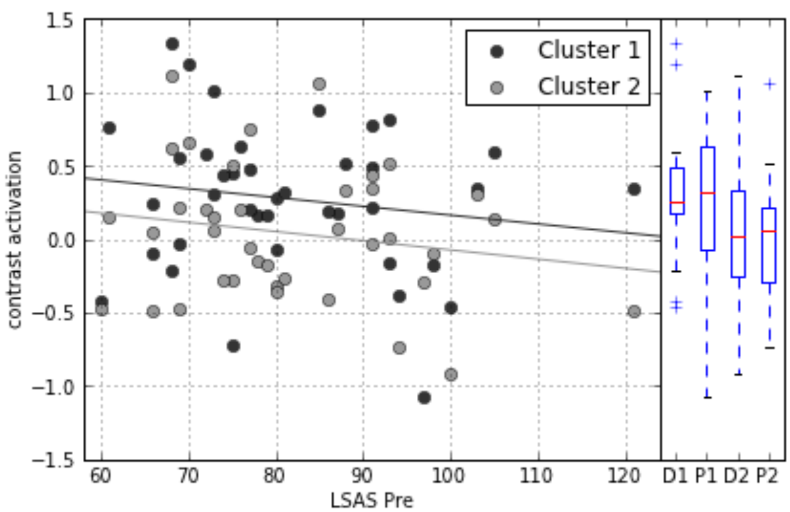
```
r: C1 (0.48514858956652174, 0.0017459420489864509)
r: C2 (0.54136848262878923, 0.00037241303817518978)
```




```
In [14]: close('all')
         axes([0.1,0.1,0.7,0.8])
         plot(pdata.lsas_pre, cmeans[:,0], 'o', color=[0.2,0.2,0.2])
```

```
plot(pdata.lsas_pre, cmeans[:,1], 'o', color=[0.6,0.6,0.6])
xlim([58, 124])
xlabel('LSAS Pre')
ylabel('contrast activation')
legend(('Cluster 1', 'Cluster 2'), 'best', numpoints=1)
plot_regression_line(pdata.lsas_pre, cmeans[:,0], [57, 125], color=[0.2,0.
plot_regression_line(pdata.lsas_pre, cmeans[:,1], [57, 125], color=[0.6,0.
grid()
axes([0.8,0.1,0.15,0.8])
boxplot([cmeans[dcsidx,0], cmeans[pcbidx,0], cmeans[dcsidx,1], cmeans[pcb
yticks([])
xticks([1,2,3,4], ['D1', 'P1', 'D2', 'P2'])
savefig('figures/scatter_means_all_lsaspre.svg',bbox_inches='tight', trans
savefig('figures/scatter_means_all_lsaspre.png', dpi=600,bbox_inches='tigh
print 'r: C1', pearsonr(cmeans[:,0], pdata.lsas_pre)
print 'r: C2', pearsonr(cmeans[:,1], pdata.lsas_pre)
print 'r: C1D', pearsonr(cmeans[dcsidx,0], pdata.lsas_pre[dcsidx])
print 'r: C2D', pearsonr(cmeans[dcsidx,1], pdata.lsas_pre[dcsidx])
print 'r: C1P', pearsonr(cmeans[pcbidx,0], pdata.lsas_pre[pcbidx])
print 'r: C2P', pearsonr(cmeans[pcbidx,1], pdata.lsas_pre[pcbidx])
```

```
r: C1 (-0.16095048699429301, 0.32766201406019846)
r: C2 (-0.18417747913131668, 0.261692057110544)
r: C1D (-0.14358866223194502, 0.56975027637211828)
r: C2D (-0.24952790428404886, 0.31800409852280737)
r: C1P (-0.16990164301777813, 0.46155369765786414)
r: C2P (-0.1158078786499194, 0.61715284762787637)
```

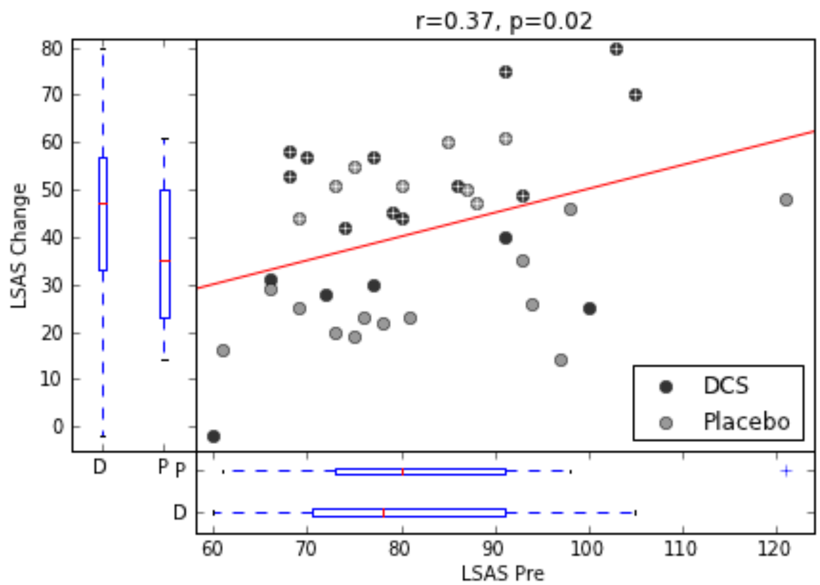


In [15]:
```
def Rmodel(y_true, y_pred):
    robjects.globalenv['y_true'] = robjects.FloatVector(y_true)
    robjects.globalenv['y_pred'] = robjects.FloatVector(y_pred)
    robjects.r("model = lm('y_true~y_pred')")
    print robjects.r("summary(model)")
```

In [16]:
```
close('all')
a1 = axes([0.05, 0.2, 0.15, 0.75])
boxplot([y[dcsidx], y[pcbidx]])
ylim([-5, 82])
ylabel('LSAS Change')
xticks([1,2],('D','P'))
a2 = axes([0.2, 0.05, 0.75, 0.15])
```

```
a2 = axes([0.2, 0.05, 0.75, 0.15])
boxplot([pdata.lsas_pre[dcsidx], pdata.lsas_pre[pcbidx]],
        vert=False)
xlim([58, 124])
xlabel('LSAS Pre')
yticks([1,2],('D','P'))
a3 = axes([0.2, 0.2, 0.75, 0.75]) #, sharex=a2, sharey=a1)
plot(pdata.lsas_pre[dcsidx], y[dcsidx], 'o', color=(0.2, 0.2, 0.2))
plot(pdata.lsas_pre[pcbidx], y[pcbidx], 'o', color=(0.6, 0.6, 0.6))
plot_regression_line(pdata.lsas_pre, y, [57, 125])
a3.set_xticks([])
a3.set_yticks([])
ylim([-5, 82])
xlim([58, 124])
grid()
legend(('DCS', 'Placebo'), 'lower right', numpoints=1)
title('r=%.2f, p=%.2f' % pearsonr(y, pdata.lsas_pre))
savefig('figures/corr_pre_delta.svg', bbox_inches='tight', transparent=Tru
savefig('figures/corr_pre_delta.png', dpi=600, bbox_inches='tight', transp
idx50 = np.nonzero(y>0.5*pdata.lsas_pre)[0]
plot(pdata.lsas_pre[idx50], y[idx50], '+', color='w')
savefig('figures/corr_pre_delta_50.svg', bbox_inches='tight', transparent=
savefig('figures/corr_pre_delta_50.png', dpi=600, bbox_inches='tight', tra
print 'D', mean(pdata.lsas_pre[dcsidx]), '+-', std(pdata.lsas_pre[dcsidx])
print 'P', mean(pdata.lsas_pre[pcbidx]), '+-', std(pdata.lsas_pre[pcbidx])
```

```
D 81.1111111111 +- 13.0847192
P 82.380952381 +- 13.393232799
```



In [17]:
```
from rpy2 import robjects
from rpy2.robjects.packages import importr
stats = importr('stats')
base = importr('base')
```

In [18]:
```
c1 = robjects.FloatVector(cmeans[:,0])
c2 = robjects.FloatVector(cmeans[:,1])
lsasd = robjects.FloatVector(y)
robjects.globalenv['c1'] = c1
```

```
robjects.globalenv['c2'] = c2
robjects.globalenv['lsaspre'] = robjects.FloatVector(pdata.lsas_pre)
robjects.globalenv['group'] = robjects.IntVector(pdata.classtype-2)
robjects.globalenv['lsasd'] = lsasd
m1 = robjects.r("model1 = lm('lsasd~c1 + c2 + lsaspre + lsaspre:group +c1:
m2 = robjects.r("model2 = lm('lsasd~lsaspre + lsaspre:group')")
m3 = robjects.r("model3 = lm('lsasd~lsaspre')")
m4 = robjects.r("model4 = lm('lsasd~group')")
m5 = robjects.r("model5 = lm('lsasd~c1 + c2 + lsaspre')")
```

In [19]:
```
print robjects.r("summary(model1)")
print robjects.r("summary(model2)")
print robjects.r("summary(model3)")
print robjects.r("anova(model3, model2)")
print robjects.r("summary(model4)")
print robjects.r("anova(model1, model5)")
```

```
Call:
lm(formula = "lsasd~c1 + c2 + lsaspre + lsaspre:group +c1:group +
c2:group")

Residuals:
     Min       1Q   Median       3Q      Max
-19.7886  -9.7221   0.7661   8.5806  23.5959

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
(Intercept)   -19.35098   12.25982  -1.578  0.12431
c1              6.99024    8.57336   0.815  0.42090
c2             22.81833    8.19825   2.783  0.00895 **
lsaspre         0.76585    0.15144   5.057 1.68e-05 ***
lsaspre:group  -0.12637    0.05612  -2.252  0.03133 *
c1:group        3.65047   10.92309   0.334  0.74041
c2:group      -11.17317   11.91793  -0.938  0.35552
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 11.81 on 32 degrees of freedom
Multiple R-squared: 0.6417,     Adjusted R-squared: 0.5745
F-statistic: 9.551 on 6 and 32 DF,  p-value: 4.967e-06



Call:
lm(formula = "lsasd~lsaspre + lsaspre:group")

Residuals:
    Min      1Q  Median      3Q     Max
-35.973 -11.985  -0.339  14.132  22.628

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   -1.88058   16.26468  -0.116  0.90859
lsaspre        0.59756    0.20092   2.974  0.00522 **
lsaspre:group -0.13576    0.06312  -2.151  0.03828 *
```

```
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 16.24 on 36 degrees of freedom
Multiple R-squared: 0.2374,     Adjusted R-squared: 0.1951
F-statistic: 5.604 on 2 and 36 DF,  p-value: 0.007604



Call:
lm(formula = "lsasd~lsaspre")

Residuals:
    Min      1Q  Median      3Q     Max
-34.623 -13.605   2.389  14.922  29.395

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -0.1687    17.0226  -0.010   0.9921
lsaspre       0.5030     0.2054   2.449   0.0192 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 17.02 on 37 degrees of freedom
Multiple R-squared: 0.1394,     Adjusted R-squared: 0.1162
F-statistic: 5.995 on 1 and 37 DF,  p-value: 0.01920


Analysis of Variance Table

Model 1: lsasd ~ lsaspre
Model 2: lsasd ~ lsaspre + lsaspre:group
  Res.Df     RSS Df Sum of Sq      F  Pr(>F)
1     37 10718.2
2     36  9497.8  1    1220.4 4.6258 0.03828 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1


Call:
lm(formula = "lsasd~group")

Residuals:
    Min      1Q  Median      3Q     Max
-48.278 -13.929  -1.278  11.647  33.722

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   46.278      4.158  11.130 2.29e-13 ***
group         -9.849      5.666  -1.738   0.0905 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 17.64 on 37 degrees of freedom
Multiple R-squared: 0.07549,    Adjusted R-squared: 0.0505
F-statistic: 3.021 on 1 and 37 DF,  p-value: 0.0905
```

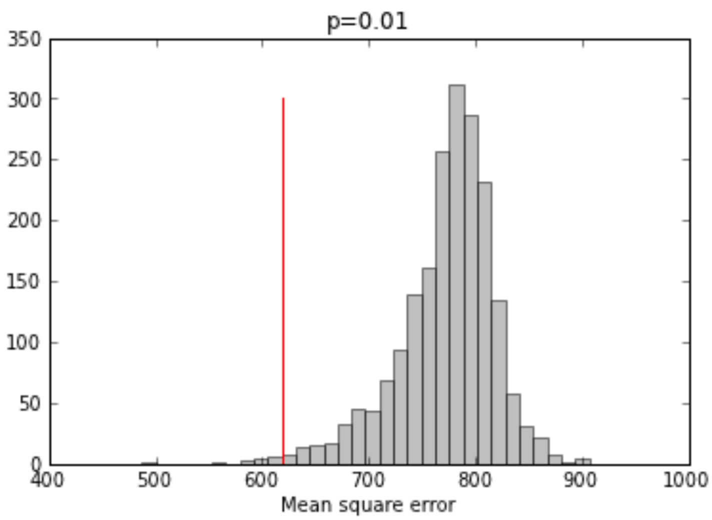
```
Analysis of Variance Table

Model 1: lsasd ~ c1 + c2 + lsaspre + lsaspre:group + c1:group + c2:group
Model 2: lsasd ~ c1 + c2 + lsaspre
  Res.Df    RSS Df Sum of Sq     F  Pr(>F)
1     32 4462.8
2     35 5539.8 -3   -1077.0 2.574 0.07119 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

In [20]:
```
from sklearn.linear_model import LinearRegression
import sklearn.cross_validation as cv
result = []
Xnew = np.vstack((pdata.lsas_pre, pdata.lsas_pre*(pdata.classtype-2))).T
for train, test in cv.StratifiedKFold(pdata.classtype, 18):
    model = LinearRegression()
    model.fit(Xnew[train], y[train])
    result.append([y[test], model.predict(Xnew[test])])
result_lsas = result
y_true = []; y_pred = []
for a,b in result:
    y_true.extend(a.tolist())
    y_pred.extend(b.tolist())
result = np.array(np.vstack((y_true, y_pred))).T
```

In [21]:
```
value, distribution, pvalue = cv.permutation_test_score(LinearRegression()
                                                        score_func=skm.mea
                                                        cv=cv.StratifiedKF
                                                        n_permutations=200
                                                        )
```

In [22]:
```
hist(distribution, 32, alpha=0.5, color='gray')
plot([value, value], [0,300], 'r')
title('p=%.2f' % (1-pvalue))
xlabel('Mean square error')
```

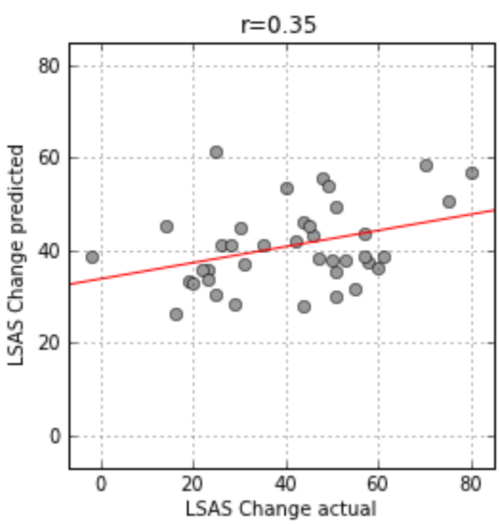
Out[22]:
```
<matplotlib.text.Text at 0x4d27e10>
```



In [23]:
```
print np.corrcoef(result.T)
Rmodel(result.T[0], result.T[1])
```

```
[[ 1.          0.35391993]
 [ 0.35391993  1.        ]]

Call:
lm(formula = "y_true~y_pred")

Residuals:
    Min      1Q  Median      3Q     Max
-41.329 -13.401  -0.674  14.265  27.500

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  11.5355    13.0814   0.882   0.3836
y_pred        0.7192     0.3124   2.302   0.0271 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 17.16 on 37 degrees of freedom
Multiple R-squared: 0.1253,     Adjusted R-squared: 0.1016
F-statistic: 5.298 on 1 and 37 DF,  p-value: 0.02708
```

In [24]:
```
plot(result[:,0], result[:,1], 'o', color=[0.6,0.6,0.6])
minv = np.min(result)-5
maxv = np.max(result)+5
plot_regression_line(result[:,0], result[:,1], [minv-1, maxv+1], color='r'
xlabel('LSAS Change actual')
ylabel('LSAS Change predicted')
axis('scaled')
ylim([minv, maxv])
xlim([minv, maxv])
grid()
title('r=%.2f' % np.corrcoef(result.T)[0,1])
savefig('figures/loo_lsaspre.svg')
savefig('figures/loo_lsaspre.png', dpi=600)
```



```
In [25]: result = []
         Xnew = np.hstack((np.vstack((pdata.lsas_pre, pdata.lsas_pre*(pdata.classty
                                cmeans))
         for train, test in cv.StratifiedKFold(pdata.classtype, 18):
             model = LinearRegression()
             model.fit(Xnew[train], y[train])
             result.append([y[test], model.predict(Xnew[test])])
         y_true = []; y_pred = []
         for a,b in result:
             y_true.extend(a.tolist())
             y_pred.extend(b.tolist())
         result = np.array(np.vstack((y_true, y_pred))).T
```

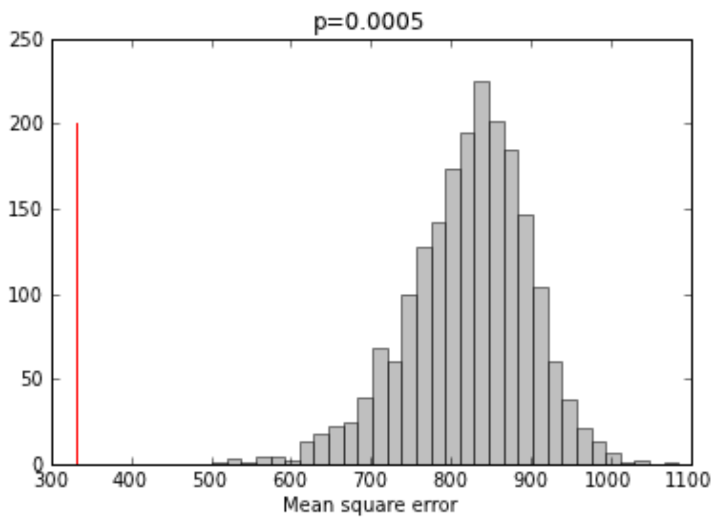
```
In [26]: np.corrcoef(result.T)
```

```
Out[26]: array([[ 1.        ,  0.72534911],
                [ 0.72534911,  1.        ]])
```

```
In [27]: value, distribution, pvalue = cv.permutation_test_score(LinearRegression()
                                                                   score_func=skm.mea
                                                                   cv=cv.StratifiedKF
                                                                   n_permutations=200
                                                                   )
```

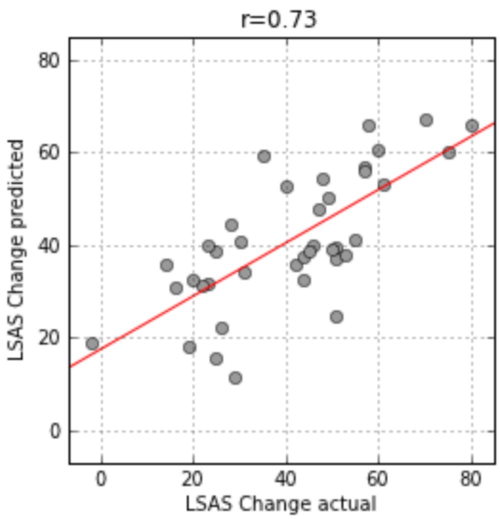
```
In [28]: pvalue = min(pvalue, 1-1./2000)
         hist(distribution, 32, alpha=0.5, color='gray')
         plot([value, value], [0,200], 'r')
         title('p=%.4f' % (1-pvalue))
         xlabel('Mean square error')
```

```
Out[28]: <matplotlib.text.Text at 0x4ed1fd0>
```

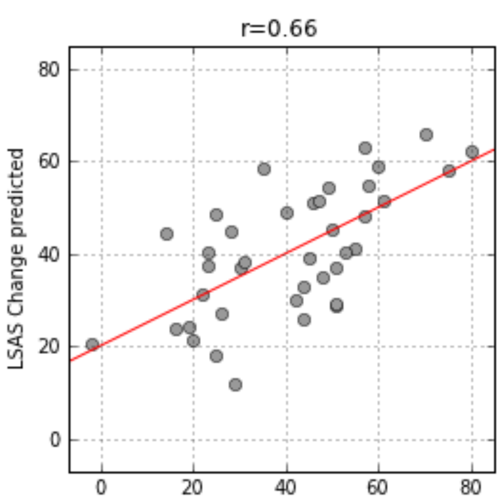



```
In [29]: plot(result[:,0], result[:,1], 'o', color=[0.6,0.6,0.6])
         xlabel('LSAS Change actual')
         ylabel('LSAS Change predicted')
         minv = np.min(result)-5
         maxv = np.max(result)+5
         plot_regression_line(result[:,0], result[:,1], [minv-1, maxv+1], color='r'
```

```
axis('scaled')
ylim([minv, maxv])
xlim([minv, maxv])
grid()
title('r=%.2f' % np.corrcoef(result.T)[0,1])
savefig('figures/loo_group_cluster.svg')
savefig('figures/loo_group_cluster.png', dpi=600)
```



In [30]:
```
cvres = np.load('result_cv.npz')
minv = np.min(cvres['aout'])-5
maxv = np.max(cvres['aout'])+5
plot(cvres['aout'][:,0], cvres['aout'][:,1], 'o', color=[0.6,0.6,0.6])
plot_regression_line(cvres['aout'][:,0], cvres['aout'][:,1], [minv-1, max
xlabel('LSAS Change actual')
ylabel('LSAS Change predicted')
axis('scaled')
ylim([minv, maxv])
xlim([minv, maxv])
grid()
title('r=%.2f' % np.corrcoef(cvres['aout'].T)[0,1])
savefig('figures/fullcv_results.svg')
savefig('figures/fullcv_results.png', dpi=600)
```

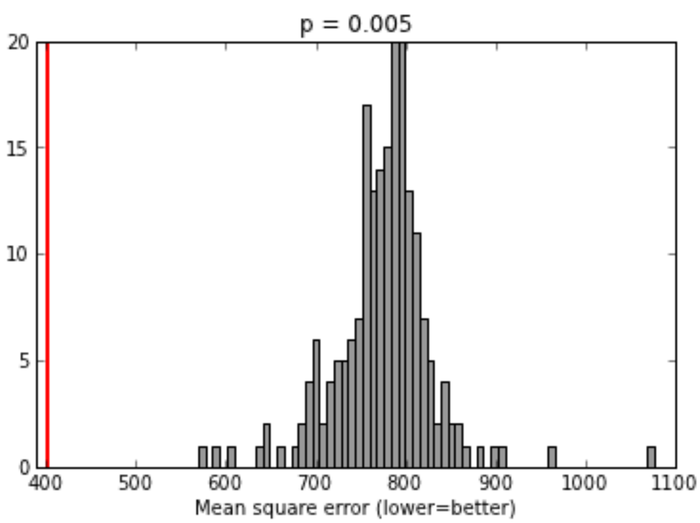


LSAS Change actual

```
In [31]: skm.explained_variance_score(cvres['aout'][:,0], cvres['aout'][:,1])
         Rmodel(cvres['aout'][:,0], cvres['aout'][:,1])
```

```
Call:
lm(formula = "y_true~y_pred")

Residuals:
    Min      1Q  Median      3Q     Max 
-30.217  -8.168   3.147  11.093  20.483 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)   5.9900     6.9811   0.858    0.396    
y_pred        0.8631     0.1633   5.285 5.83e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 13.85 on 37 degrees of freedom
Multiple R-squared: 0.4302,	Adjusted R-squared: 0.4148 
F-statistic: 27.93 on 1 and 37 DF,  p-value: 5.829e-06
```

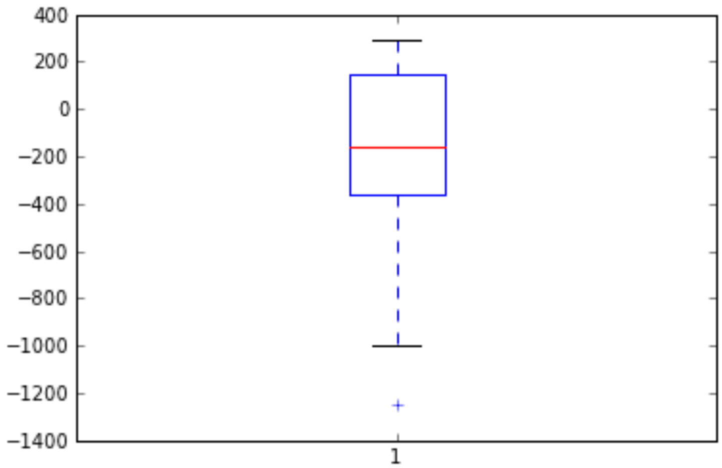
```
In [32]: permdata = np.load('permtest200.npz')
         hist(permdata['distribution'], 64, color=[0.6,0.6,0.6])
         plot([permdata['value'], permdata['value']], [0, 20], color='r', linewidtl
         title('p = %.3f' % max(1./200, (1-permdata['pvalue'])))
         xlim([390, 1100])
         xlabel('Mean square error (lower=better)')
         savefig("figures/permtest_hist.svg")
         savefig("figures/permtest_hist.png", dpi=600)
```




```
In [33]: msedata = []
         for idx, res in enumerate(result_lsas):
             msedata.append((skm.mean_square_error(res[0], res[1]),
                             skm.mean_square_error(cvres['result'][idx][0],
```

```
cvres['result'][idx][1])))
```

```
In [34]: print wilcoxon(np.diff(msedata, axis=1).ravel())
         boxplot(np.diff(msedata, axis=1))

         (44.0, 0.070709320478686236)
Out[34]: {'boxes': [<matplotlib.lines.Line2D at 0x6714290>],
          'caps': [<matplotlib.lines.Line2D at 0x6594950>,
           <matplotlib.lines.Line2D at 0x6714850>],
          'fliers': [<matplotlib.lines.Line2D at 0x6714bd0>,
           <matplotlib.lines.Line2D at 0x6802550>],
          'medians': [<matplotlib.lines.Line2D at 0x6714e50>],
          'whiskers': [<matplotlib.lines.Line2D at 0x6585810>,
           <matplotlib.lines.Line2D at 0x65941d0>]}
```



Other factors: MADRS, Sex, Comorbidity

```
In [35]: otherdata = np.recfromcsv('ControlParameters_Prediction.csv', usecols=[1,2
         otherX = otherdata.view(np.int).reshape(39,3)
         names = otherdata.dtype.names
```

```
In [36]: robjects.globalenv['y_true'] = robjects.FloatVector(y)
         robjects.globalenv['lsaspre'] = robjects.FloatVector(pdata.lsas_pre)
         robjects.globalenv['group'] = robjects.IntVector(pdata.classtype-2)
         for i,name in enumerate(names):
             robjects.globalenv[name] = robjects.FloatVector(otherX[:,i])
         m1str = 'y_true~lsaspre + lsaspre:group + %s + %s' % ('+'.join(names), ':g
         m1 = robjects.r("m1 = lm(%s)" % m1str)
         print robjects.r("summary(m1)")
         m3 = robjects.r("m3 = lm('y_true~lsaspre + lsaspre:group')")
         print robjects.r("anova(m3,m1)")
```

```
Call:
lm(formula = y_true ~ lsaspre + lsaspre:group + sex + madrs_pre +
    comorbid_anxiety_disorder + sex:group + madrs_pre:group +
```

```
    comorbid_anxiety_disorder)

Residuals:
     Min       1Q   Median       3Q      Max 
-29.6333 -11.3696   0.7282  11.1426  24.7014 

Coefficients:
                          Estimate Std. Error t value Pr(>|t|)    
(Intercept)                 4.9135    16.9447   0.290  0.77377    
lsaspre                     0.9711     0.2629   3.693  0.00085 ***
sex                       -20.4286     9.5126  -2.148  0.03968 *  
madrs_pre                  -0.9329     0.7386  -1.263  0.21599    
comorbid_anxiety_disorder  -3.2736     6.1502  -0.532  0.59833    
lsaspre:group              -0.6527     0.2745  -2.378  0.02375 *  
group:sex                  24.9847    12.5279   1.994  0.05497 .  
group:madrs_pre             0.9200     0.9227   0.997  0.32644    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1 

Residual standard error: 16.11 on 31 degrees of freedom
Multiple R-squared: 0.3543,	Adjusted R-squared: 0.2085 
F-statistic:  2.43 on 7 and 31 DF,  p-value: 0.04167 


Analysis of Variance Table

Model 1: y_true ~ lsaspre + lsaspre:group
Model 2: y_true ~ lsaspre + lsaspre:group + sex + madrs_pre + 
comorbid_anxiety_disorder + 
    sex:group + madrs_pre:group + comorbid_anxiety_disorder
  Res.Df    RSS Df Sum of Sq      F Pr(>F)
1     36 9497.8                           
2     31 8042.6  5    1455.2 1.1218 0.3694
```

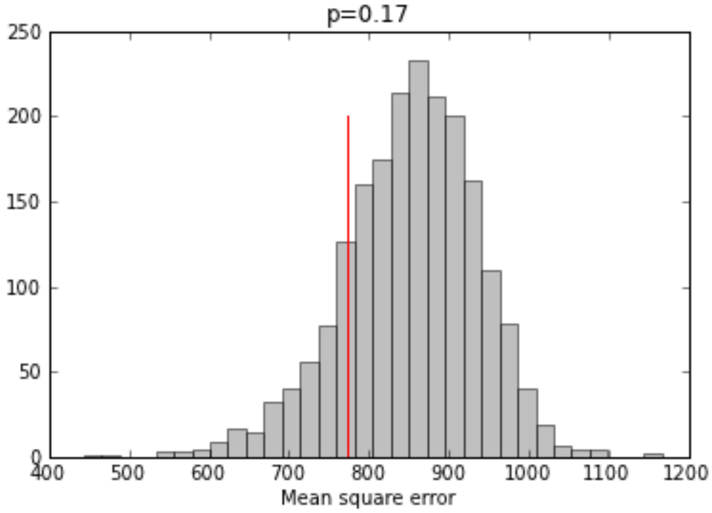
In [37]:
```
result = []
Xnew = np.hstack((np.vstack((pdata.lsas_pre, pdata.lsas_pre*(pdata.classty
                  otherX))
for train, test in cv.StratifiedKFold(pdata.classtype, 18):
    model = LinearRegression()
    model.fit(Xnew[train], y[train])
    result.append([y[test], model.predict(Xnew[test])])
y_true = []; y_pred = []
for a,b in result:
    y_true.extend(a.tolist())
    y_pred.extend(b.tolist())
result = np.array(np.vstack((y_true, y_pred))).T
```

In [38]:
```
value, distribution, pvalue = cv.permutation_test_score(LinearRegression()
                                                        score_func=skm.mea
                                                        cv=cv.StratifiedKF
                                                        n_permutations=200
                                                        )
```

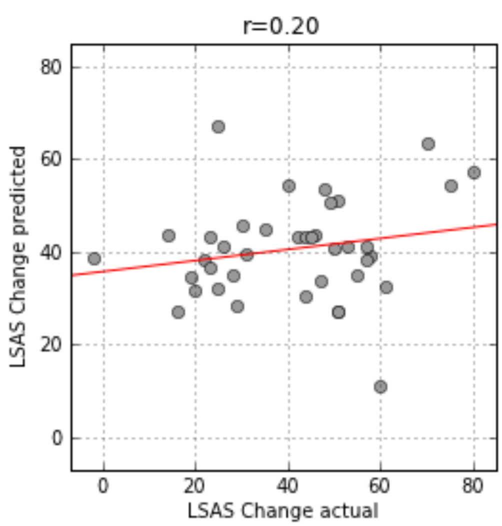
In [39]:
```
hist(distribution, 32, alpha=0.5, color='gray')
```

```
plot([value, value], [0,200], 'r')
title('p=%.2f' % (1-pvalue))
xlabel('Mean square error')
```

Out[39]: <matplotlib.text.Text at 0x6808790>



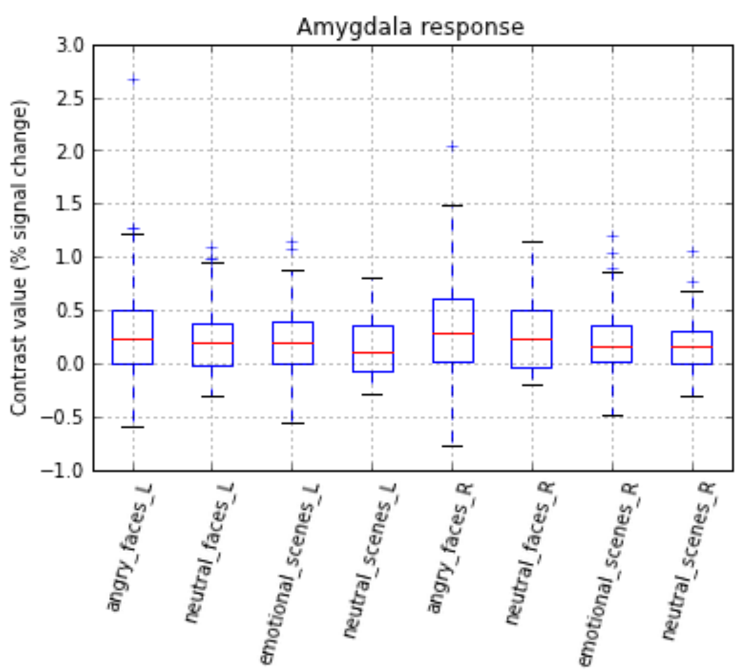
In [40]:
```
plot(result[:,0], result[:,1], 'o', color=[0.6, 0.6, 0.6])
xlabel('LSAS Change actual')
ylabel('LSAS Change predicted')
minv = np.min(result)-5
maxv = np.max(result)+5
plot_regression_line(result[:,0], result[:,1], [minv-1, maxv+1], color='r'
axis('scaled')
ylim([minv, maxv])
xlim([minv, maxv])
grid()
title('r=%.2f' % np.corrcoef(result.T)[0,1])
savefig('figures/loo_depression.svg')
savefig('figures/loo_depression.png', dpi=600)
```



Amygdala responses

In [41]:
```
amygdata = recfromcsv('AmygdalaResponses.csv', names=True)
amygX = amygdata.view(np.float64).reshape(39,8)
names = []
for name in amygdata.dtype.names:
    if '_1' in name:
        names.append(name.replace('_1','_R'))
    else:
        names.append(name+'_L')
```

In [42]:
```
bp = boxplot(amygX)
xticks(arange(1,9), names, rotation=75)
ylabel("Contrast value (% signal change)")
grid()
title('Amygdala response')
savefig('figures/amygdala_response.svg', bbox_inches='tight')
savefig('figures/amygdala_response.png', dpi=600, bbox_inches='tight')
```



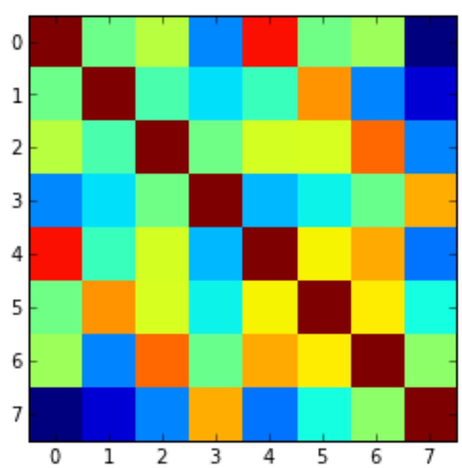
In [43]:
```
robjects.globalenv['y_true'] = robjects.FloatVector(y)
robjects.globalenv['lsaspre'] = robjects.FloatVector(pdata.lsas_pre)
robjects.globalenv['group'] = robjects.IntVector(pdata.classtype-2)
for i,name in enumerate(names):
    robjects.globalenv[name] = robjects.FloatVector(amygX[:,i])
m1str = 'y_true~lsaspre + lsaspre:group + %s + %s' % ('+'.join(names), ':g
m1 = robjects.r("m1 = lm(%s)" % m1str)
print robjects.r("summary(m1)")
m2 = robjects.r("m2 = lm('y_true~lsaspre + lsaspre:group + angry_faces_R +
print robjects.r("summary(m2)")
m3 = robjects.r("m3 = lm('y_true~lsaspre + lsaspre:group')")
print robjects.r("anova(m3,m1)")
```

```
Call:
lm(formula = y_true ~ lsaspre + lsaspre:group + angry_faces_L +
    neutral_faces_L + emotional_scenes_L + neutral_scenes_L +
    angry_faces_R + neutral_faces_R + emotional_scenes_R +
neutral_scenes_R +
    angry_faces_L:group + neutral_faces_L:group + emotional_scenes_L:group
+
    neutral_scenes_L:group + angry_faces_R:group + neutral_faces_R:group +
    emotional_scenes_R:group + neutral_scenes_R)

Residuals:
    Min      1Q  Median      3Q     Max
-37.896  -5.868  -1.560   9.892  26.182

Coefficients:
                          Estimate Std. Error t value Pr(>|t|)
(Intercept)               10.41843   24.43309   0.426   0.6742
lsaspre                    0.41167    0.31164   1.321   0.2007
angry_faces_L             -7.22086   42.04757  -0.172   0.8653
neutral_faces_L          -65.96567   59.87470  -1.102   0.2830
emotional_scenes_L        29.57672   27.21910   1.087   0.2895
neutral_scenes_L          24.18155   63.63084   0.380   0.7077
angry_faces_R             24.37981   58.96838   0.413   0.6835
neutral_faces_R           60.95377   63.69678   0.957   0.3495
emotional_scenes_R       -75.37781   40.66474  -1.854   0.0779 .
neutral_scenes_R           6.08784   49.26220   0.124   0.9028
lsaspre:group             -0.08846    0.10475  -0.845   0.4079
group:angry_faces_L       -5.13978   51.69948  -0.099   0.9218
group:neutral_faces_L     80.38850   68.62127   1.171   0.2545
group:emotional_scenes_L -38.48076   44.63915  -0.862   0.3984
group:neutral_scenes_L   -11.52532   50.43130  -0.229   0.8214
group:angry_faces_R       -8.67095   70.62746  -0.123   0.9035
group:neutral_faces_R    -95.08737   77.81603  -1.222   0.2353
group:emotional_scenes_R  89.63743   66.07628   1.357   0.1893
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 18.3 on 21 degrees of freedom
Multiple R-squared: 0.4354,     Adjusted R-squared: -0.02169
F-statistic: 0.9525 on 17 and 21 DF,  p-value: 0.5349



Call:
lm(formula = "y_true~lsaspre + lsaspre:group + angry_faces_R +
angry_faces_R:group + neutral_faces_R + neutral_faces_R:group")

Residuals:
    Min      1Q  Median      3Q     Max
-35.533 -11.448  -1.034  13.955  24.262

Coefficients:
                      Estimate Std. Error t value Pr(>|t|)
(Intercept)           -5.13664   19.33876  -0.266   0.7922
lsaspre                0.61734    0.22797   2.708   0.0108 *
```

```
angry_faces_R          -1.41219   10.76728  -0.131   0.8965
neutral_faces_R         6.38532   19.22618   0.332   0.7420
lsaspre:group          -0.11859    0.08683  -1.366   0.1815
group:angry_faces_R     6.38895   16.81749   0.380   0.7065
group:neutral_faces_R -11.46601   25.48513  -0.450   0.6558
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 17.15 on 32 degrees of freedom
Multiple R-squared: 0.2442,     Adjusted R-squared: 0.1025
F-statistic: 1.723 on 6 and 32 DF,  p-value: 0.1476


Analysis of Variance Table

Model 1: y_true ~ lsaspre + lsaspre:group
Model 2: y_true ~ lsaspre + lsaspre:group + angry_faces_L +
neutral_faces_L +
    emotional_scenes_L + neutral_scenes_L + angry_faces_R +
neutral_faces_R +
    emotional_scenes_R + neutral_scenes_R + angry_faces_L:group +
    neutral_faces_L:group + emotional_scenes_L:group +
neutral_scenes_L:group +
    angry_faces_R:group + neutral_faces_R:group + emotional_scenes_R:group
+
    neutral_scenes_R
  Res.Df    RSS Df Sum of Sq      F Pr(>F)
1     36 9497.8
2     21 7032.3 15    2465.5 0.4908 0.9191
```

In [44]:
```
imshow(corrcoef(amygX.T), interpolation='nearest')
```

Out[44]: <matplotlib.image.AxesImage at 0x70b1550>



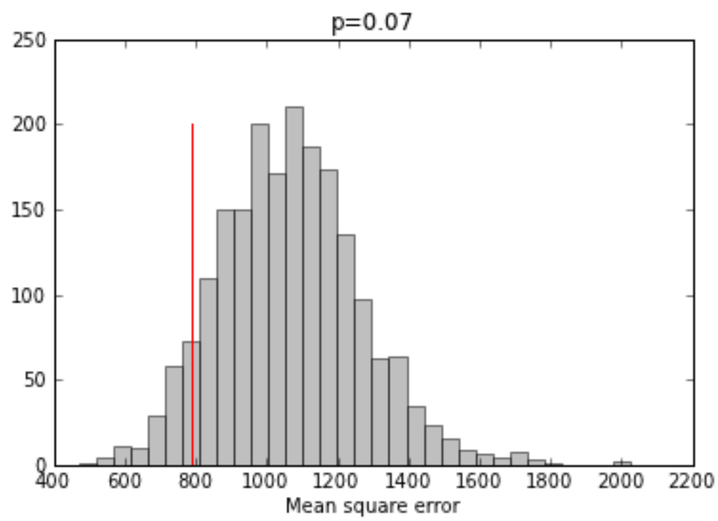
In [45]:
```
result = []
Xnew = np.hstack((np.vstack((pdata.lsas_pre, pdata.lsas_pre*(pdata.classty
                     amygX))
for train, test in cv.StratifiedKFold(pdata.classtype, 18):
    model = LinearRegression()
    model.fit(Xnew[train], y[train])
    result.append([y[test], model.predict(Xnew[test])])
y_true = []; y_pred = []
```

```
for a,b in result:
    y_true.extend(a.tolist())
    y_pred.extend(b.tolist())
result = np.array(np.vstack((y_true, y_pred))).T
```

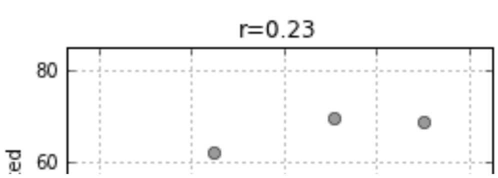
In [46]:
```
value, distribution, pvalue = cv.permutation_test_score(LinearRegression()
                                                        score_func=skm.mea
                                                        cv=cv.StratifiedKF
                                                        n_permutations=200
                                                        )
```

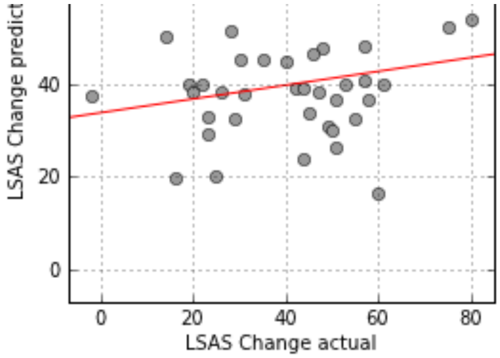
In [47]:
```
hist(distribution, 32, alpha=0.5, color='gray')
plot([value, value], [0,200], 'r')
title('p=%.2f' % (1-pvalue))
xlabel('Mean square error')
```

Out[47]:
```
<matplotlib.text.Text at 0x6e315d0>
```



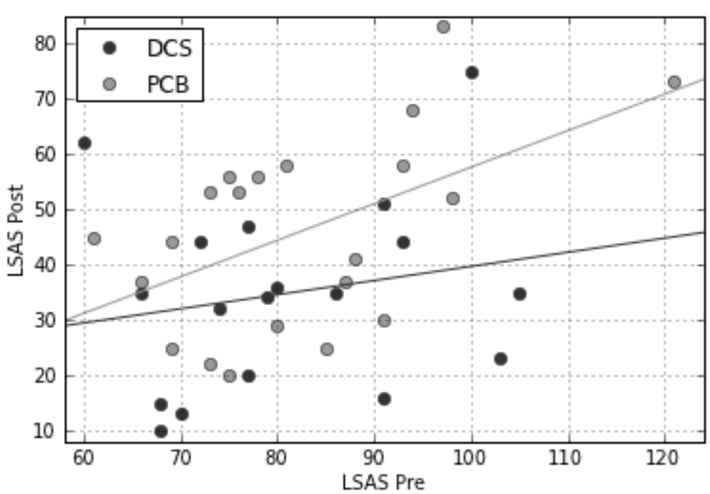
In [48]:
```
plot(result[:,0], result[:,1], 'o', color=[0.6, 0.6, 0.6])
xlabel('LSAS Change actual')
ylabel('LSAS Change predicted')
minv = np.min(result)-5
maxv = np.max(result)+5
plot_regression_line(result[:,0], result[:,1], [minv-1, maxv+1], color='r'
axis('scaled')
ylim([minv, maxv])
xlim([minv, maxv])
grid()
title('r=%.2f' % np.corrcoef(result.T)[0,1])
savefig('figures/loo_amygdala.svg')
savefig('figures/loo_amygdala.png', dpi=600)
```






```
In [49]: print pearsonr(pdata.lsas_pre,pdata.lsas_post)
         print pearsonr(pdata.lsas_pre,pdata.lsas_pre-pdata.lsas_post)
         print 'DP:', pearsonr(pdata.lsas_pre[dcsidx],pdata.lsas_post[dcsidx])
         print 'PP:', pearsonr(pdata.lsas_pre[pcbidx],pdata.lsas_post[pcbidx])
         print 'DD:',pearsonr(pdata.lsas_pre[dcsidx],pdata.lsas_pre[dcsidx]-pdata.l
         print 'PD:',pearsonr(pdata.lsas_pre[pcbidx],pdata.lsas_pre[pcbidx]-pdata.l
         print spearmanr(pdata.lsas_pre,pdata.lsas_post)
         print spearmanr(pdata.lsas_pre,pdata.lsas_pre-pdata.lsas_post)
         plot(pdata.lsas_pre[dcsidx], pdata.lsas_post[dcsidx], 'o', color=(0.2,0.2,
         plot(pdata.lsas_pre[pcbidx], pdata.lsas_post[pcbidx], 'o', color=(0.6,0.6,
         legend(['DCS', 'PCB'], 'best', numpoints=1)
         plot_regression_line(pdata.lsas_pre[dcsidx], pdata.lsas_post[dcsidx], [55,
                              color=[0.2,0.2,0.2])
         plot_regression_line(pdata.lsas_pre[pcbidx], pdata.lsas_post[pcbidx], [55,
                              color=[0.6,0.6,0.6])
         xlabel('LSAS Pre')
         ylabel('LSAS Post')
         xlim([58,124])
         ylim([8,85])
         grid()
```

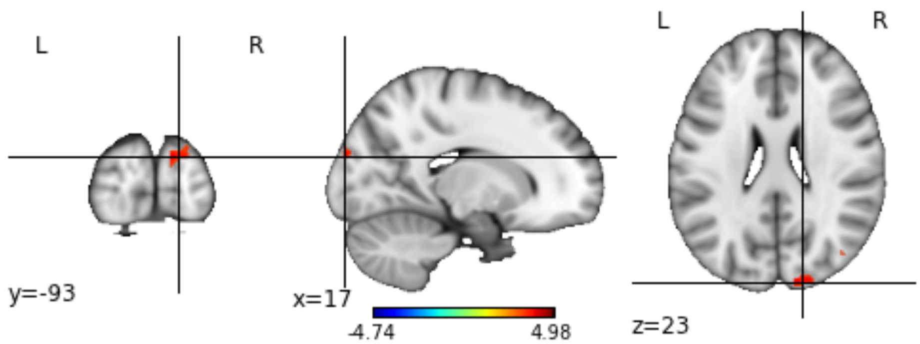
```
(0.36957463542651603, 0.020582499321071621)
(0.37342153991691346, 0.019203555707493748)
DP: (0.19781321306137134, 0.4313860574718511)
PP: (0.51996702206134149, 0.015686563949718763)
DD: (0.50692543723149752, 0.031787597769797171)
PD: (0.29883566124575828, 0.18821010684663192)
(0.30034458449575868, 0.0632007389267807)
(0.27659575035945, 0.088273733122193665)
```

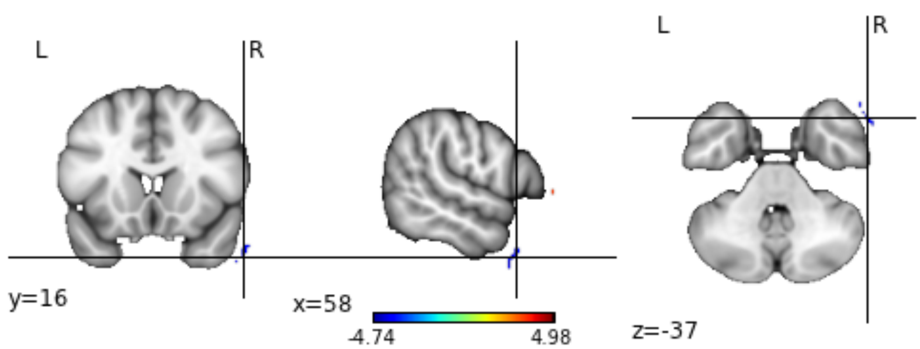


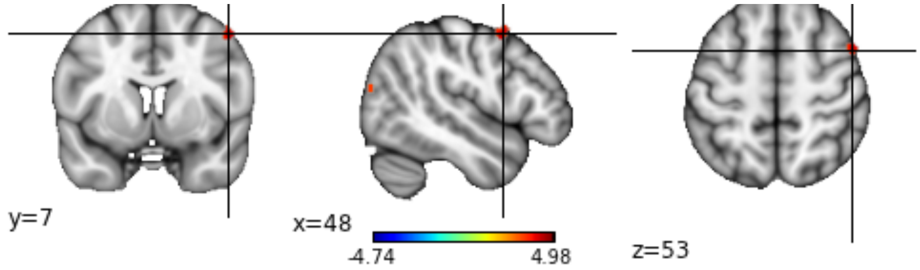
LSAS delta

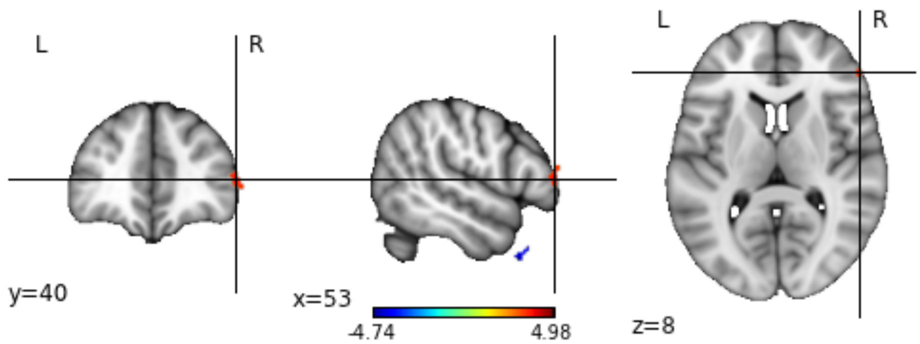
In [50]:
```
filename = os.path.join(base_dir, 'lsasdelta_all', 'conest', 'spmT_0001.in
img=load(filename)
print img.get_header()['descrip']
labels, nlabels = get_labels(abs(img.get_data())>ss.t.ppf(1-0.001,35), 20)
data = img.get_data()
data[labels==0] = 0
#cmeans = get_clustermeans(X, labels, nlabels)
coords = get_coords(labels, img.get_affine())
print_table(labels, img.get_affine(),
            np.prod(img.get_header().get_zooms()), img.get_data())
show_slices(img, coords, threshold=0.5, prefix='uncorrected_lsasdelta', sh
            cmap=cm.jet)
```

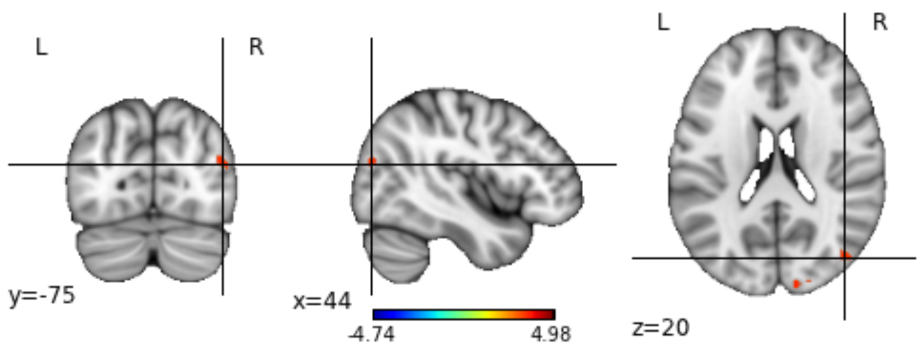
```
SPM{T_[35.0]} - contrast 1: LSAS Delta Response
Cluster         X       Y       Z        k      T
Cluster 00      17      -93     23       672    4.14
Cluster 01      58      16      -37      376    -4.76
Cluster 02      48      7       53       360    4.99
Cluster 03      53      40      8        296    3.93
Cluster 04      44      -75     20       256    3.70
```









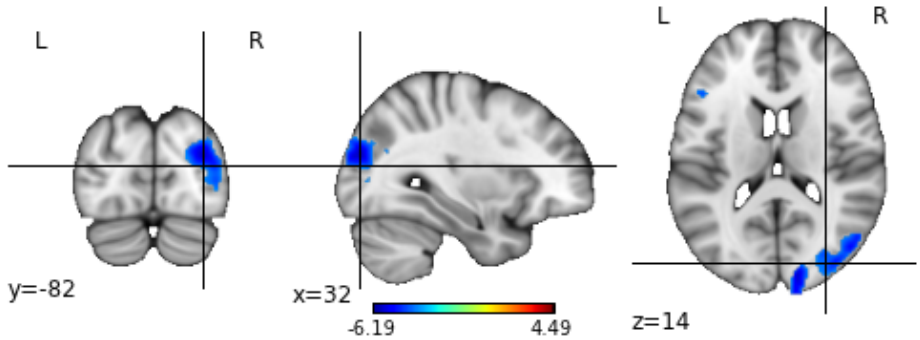


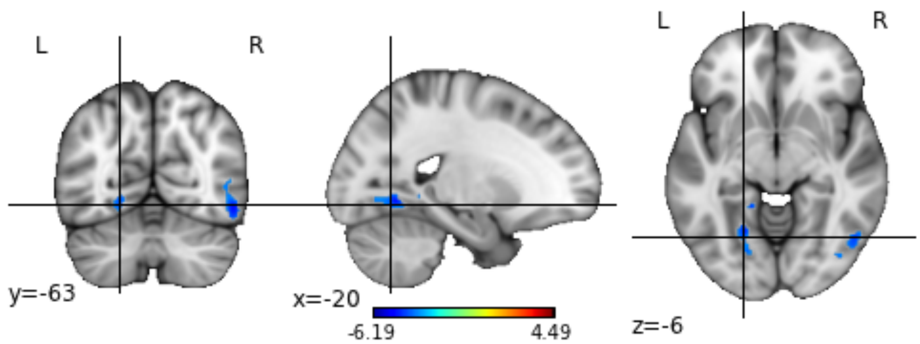
LSAS Post

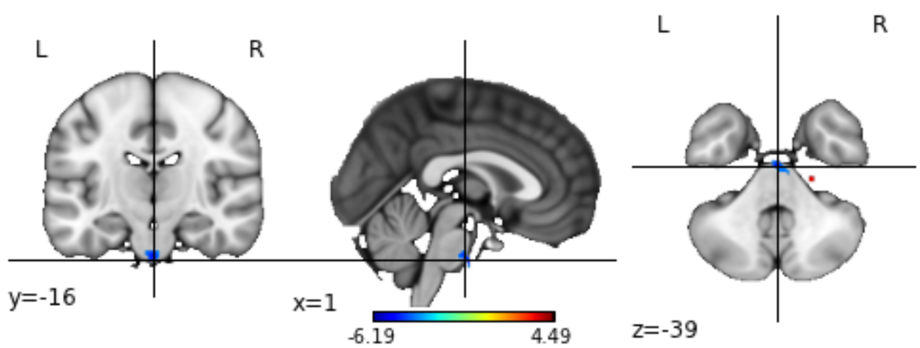
In [51]:
```
filename = os.path.join(base_dir, 'lsaspost_all', 'conest', 'spmT_0001.img
img=load(filename)
print img.get_header()['descrip']
labels, nlabels = get_labels(abs(img.get_data())>ss.t.ppf(1-0.001,35), 20)
data = img.get_data()
data[labels==0] = 0
#cmeans = get_clustermeans(X, labels, nlabels)
coords = get_coords(labels, img.get_affine())
print_table(labels, img.get_affine(),
            np.prod(img.get_header().get_zooms()), img.get_data())
show_slices(img, coords, threshold=0.5, prefix='uncorrected_lsaspost', sho
            cmap=cm.jet)
```

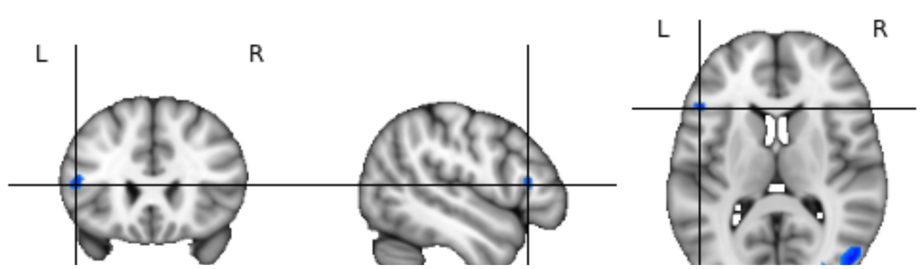
```
SPM{T_[35.0]} - contrast 1: LSAS Delta Response
Cluster          X       Y       Z       k       T
```

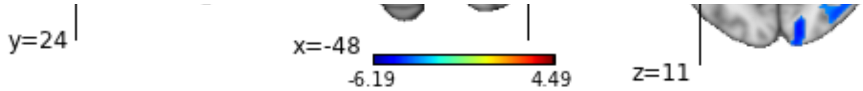
Cluster 00	32	-82	14	13864	-6.22
Cluster 01	-20	-63	-6	672	-4.67
Cluster 02	1	-16	-39	432	-4.68
Cluster 03	-48	24	11	312	-4.23
Cluster 04	23	-23	-44	296	4.51
Cluster 05	-16	-45	-3	232	-4.84
Cluster 06	48	7	55	216	-4.49

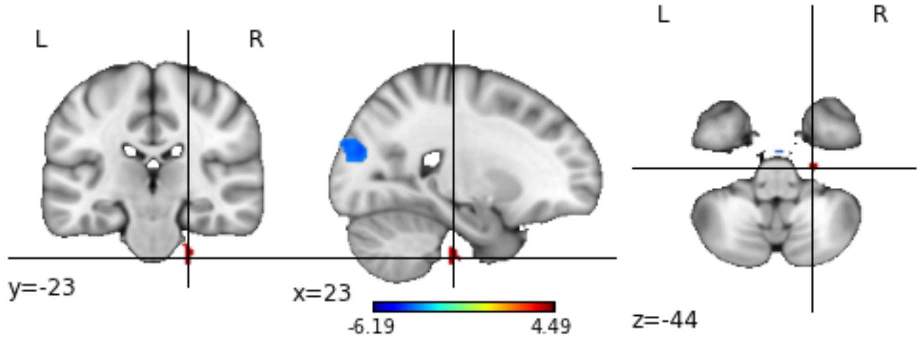


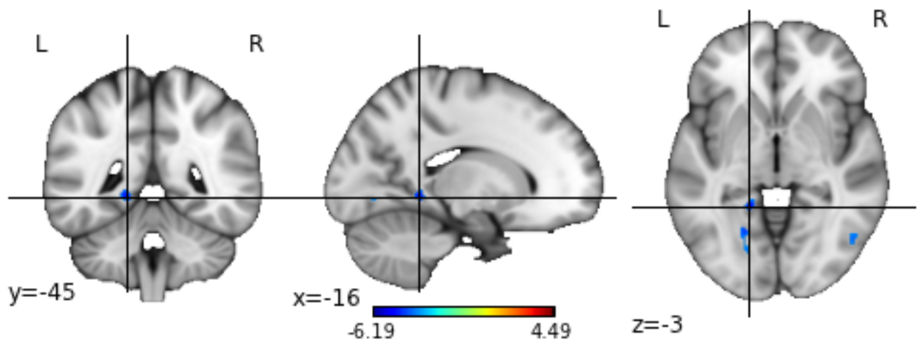


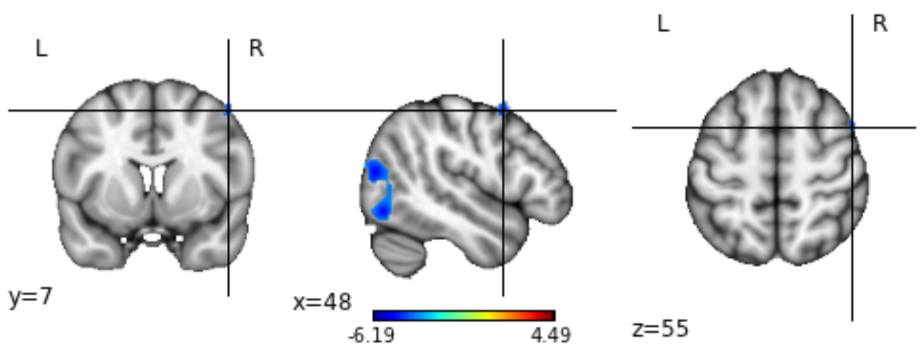












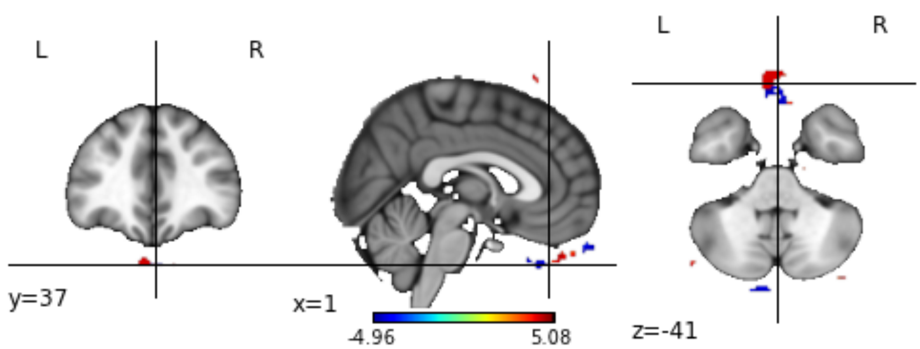
LSAS pre

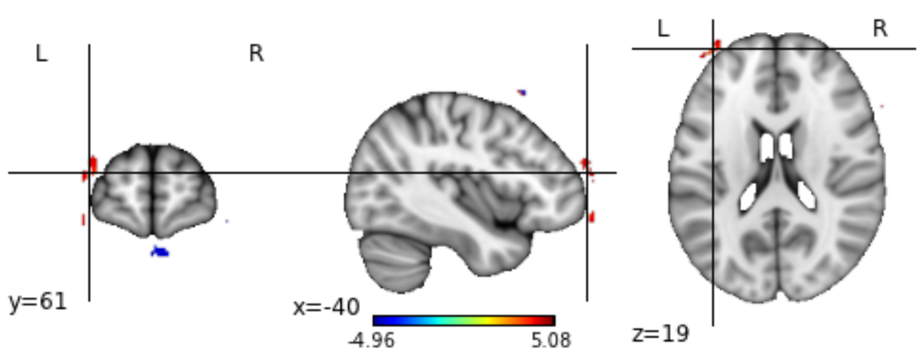
In [52]:
```
filename = os.path.join(base_dir, 'lsaspre_all', 'conest', 'spmT_0001.img'
img=load(filename)
print img.get_header()['descrip']
labels, nlabels = get_labels(abs(img.get_data())>ss.t.ppf(1-0.001,35), 20)
data = img.get_data()
data[labels==0] = 0
#cmeans = get_clustermeans(X, labels, nlabels)
```

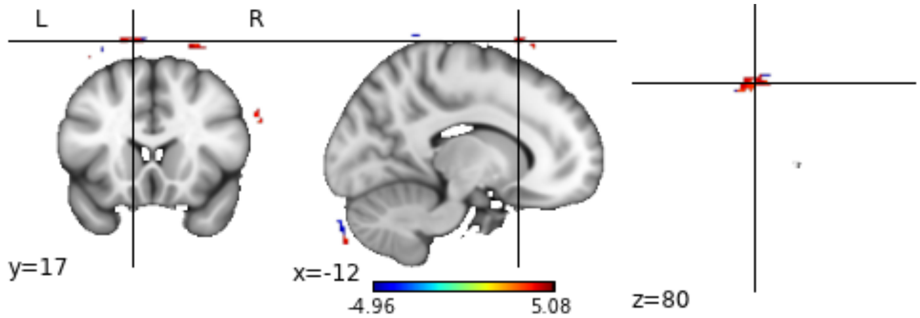
```
coords = get_coords(labels, img.get_affine())
print_table(labels, img.get_affine(),
            np.prod(img.get_header().get_zooms()), img.get_data())
show_slices(img, coords, threshold=0.5, prefix='uncorrected_lsaspre', show
            cmap=cm.jet)
```

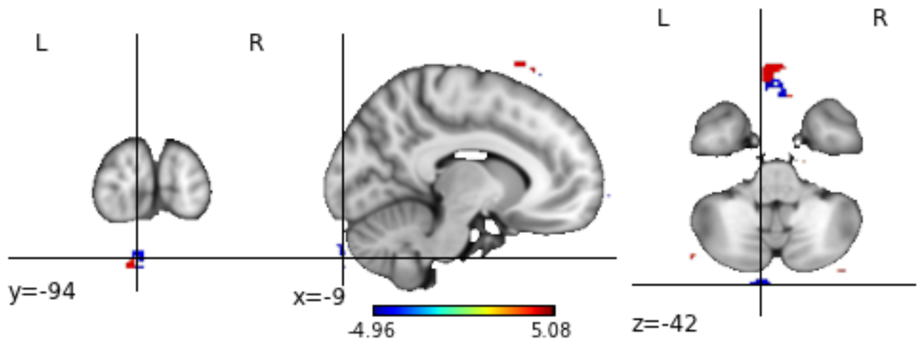
SPM{T_[35.0]} - contrast 1: LSAS Delta Response

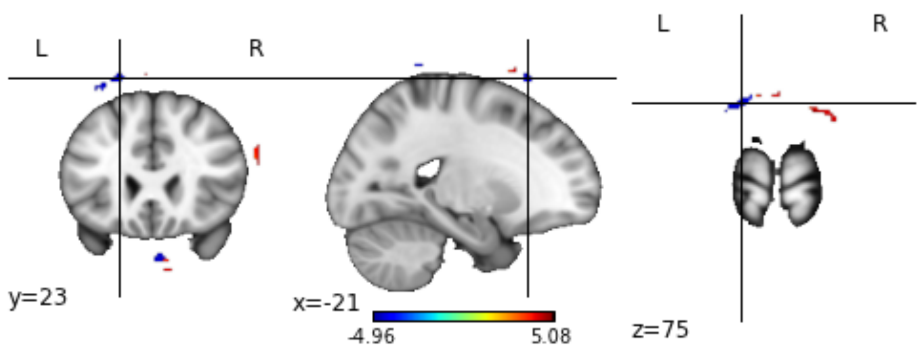
Cluster	X	Y	Z	k	T
Cluster 00	1	37	-41	2200	5.10
Cluster 01	-40	61	19	824	4.53
Cluster 02	-12	17	80	504	4.46
Cluster 03	-9	-94	-42	480	-4.68
Cluster 04	-21	23	75	424	-4.43
Cluster 05	20	-19	-47	408	4.62
Cluster 06	67	21	28	384	4.54
Cluster 07	-38	65	-11	360	4.51
Cluster 08	4	62	-33	344	-4.45
Cluster 09	-61	-50	-49	272	4.98
Cluster 10	-32	24	70	248	-4.39
Cluster 11	30	17	76	224	4.84
Cluster 12	41	-88	-38	224	5.02
Cluster 13	-57	30	39	208	-4.74
Cluster 14	-2	29	75	200	4.35
Cluster 15	-3	76	-6	200	-4.34
Cluster 16	65	-38	-34	184	4.49
Cluster 17	-17	-47	84	176	-4.71
Cluster 18	-51	-78	-46	168	4.36
Cluster 19	42	66	-9	168	-4.34

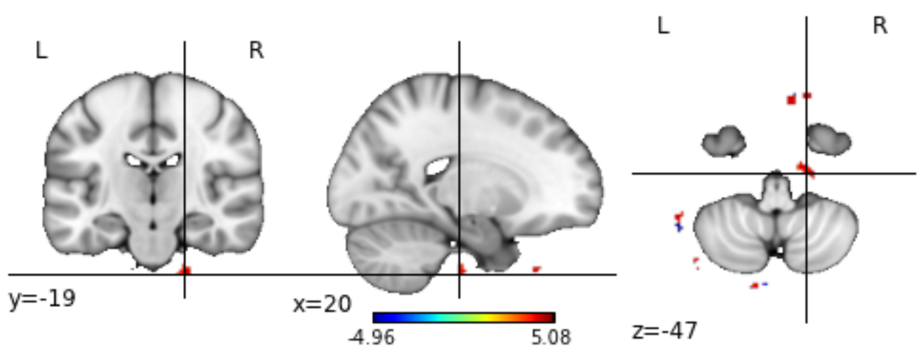


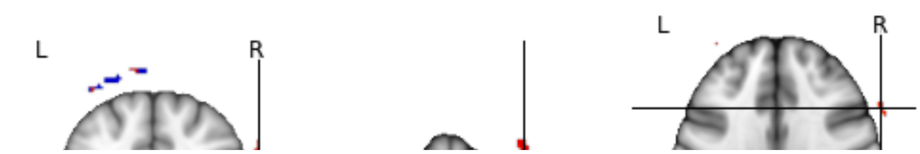


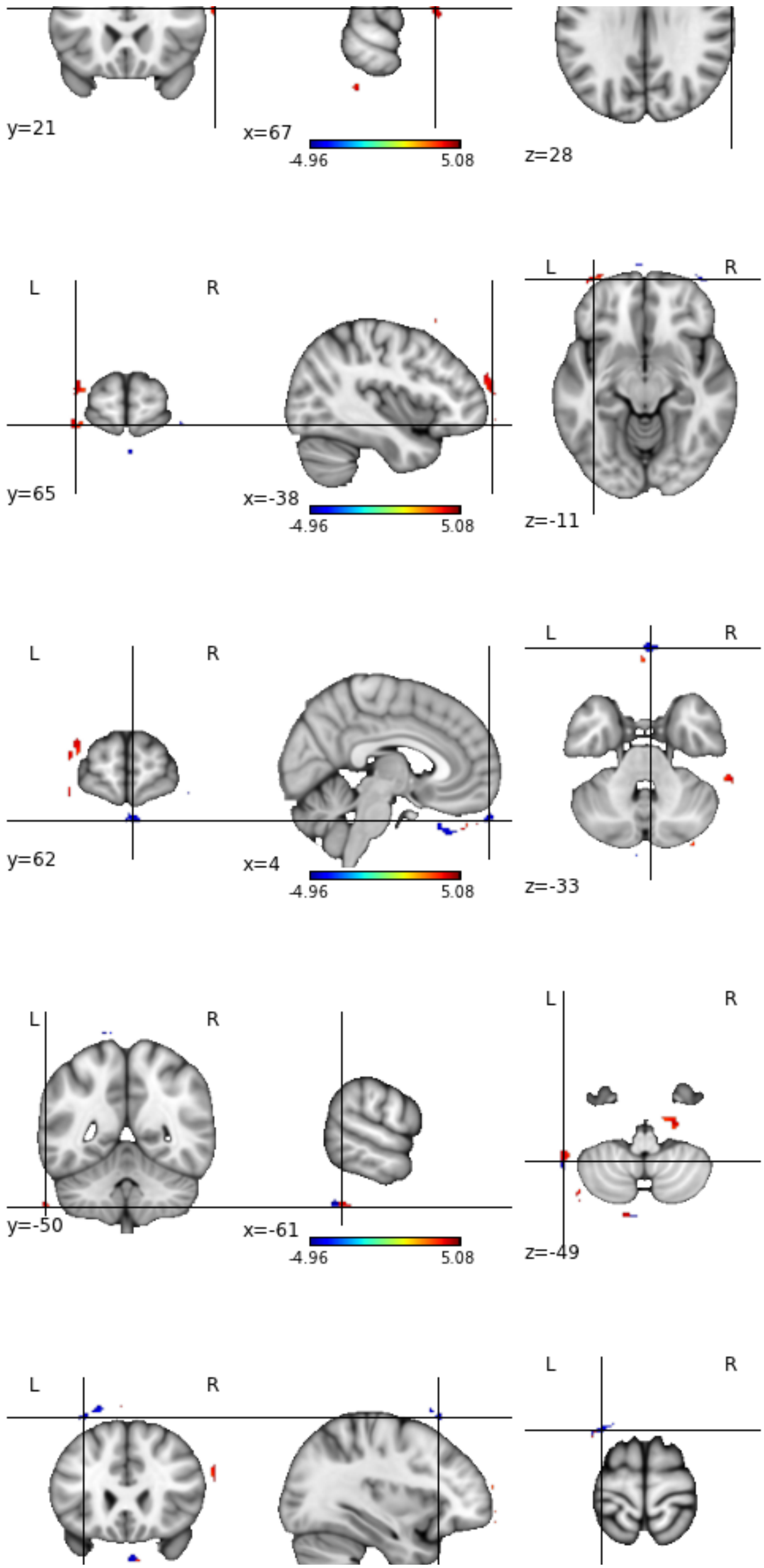

L
R
y=17
x=-12
-4.96
5.08
z=80

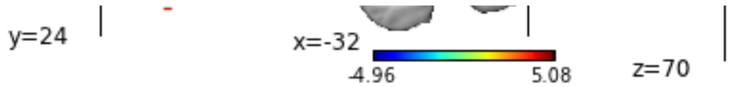

L
R
y=-94
x=-9
-4.96
5.08
L
R
z=-42

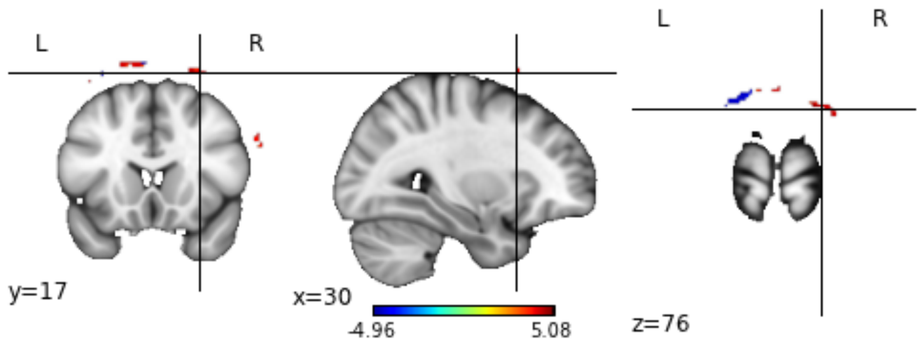

L
R
y=23
x=-21
-4.96
5.08
L
R
z=75

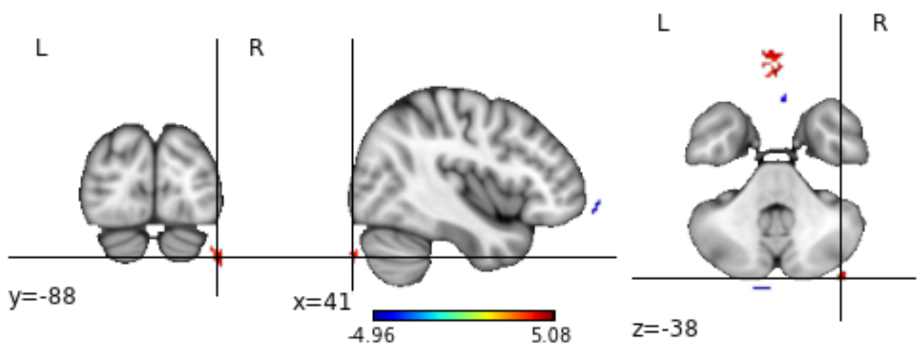

L
R
y=-19
x=20
-4.96
5.08
L
R
z=-47

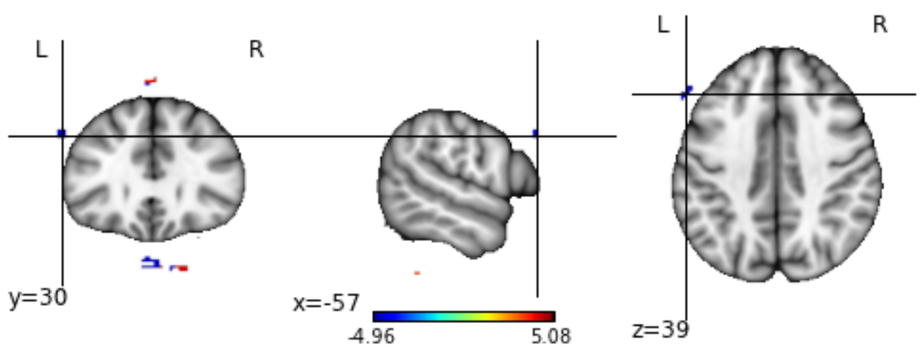

L
R
L
R

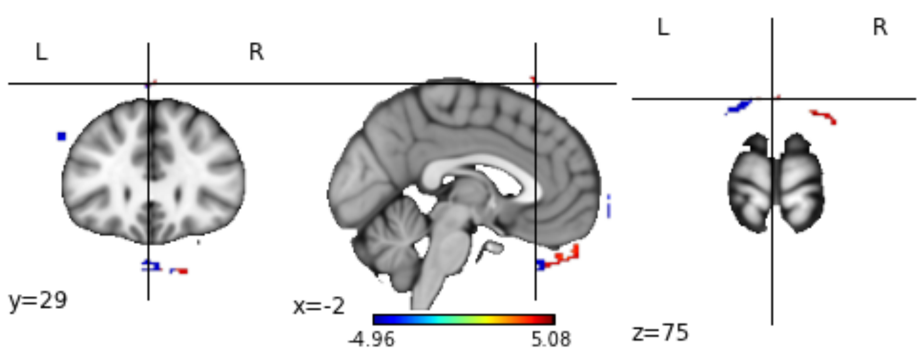

y=21
x=67
-4.96
5.08
z=28
L
R
y=65
x=-38
-4.96
5.08
L
R
z=-11
L
R
y=62
x=4
-4.96
5.08
L
R
z=-33
L
R
y=-50
x=-61
-4.96
5.08
L
R
z=-49
L
R
L
R


y=24
x=-32
-4.96
5.08
z=70

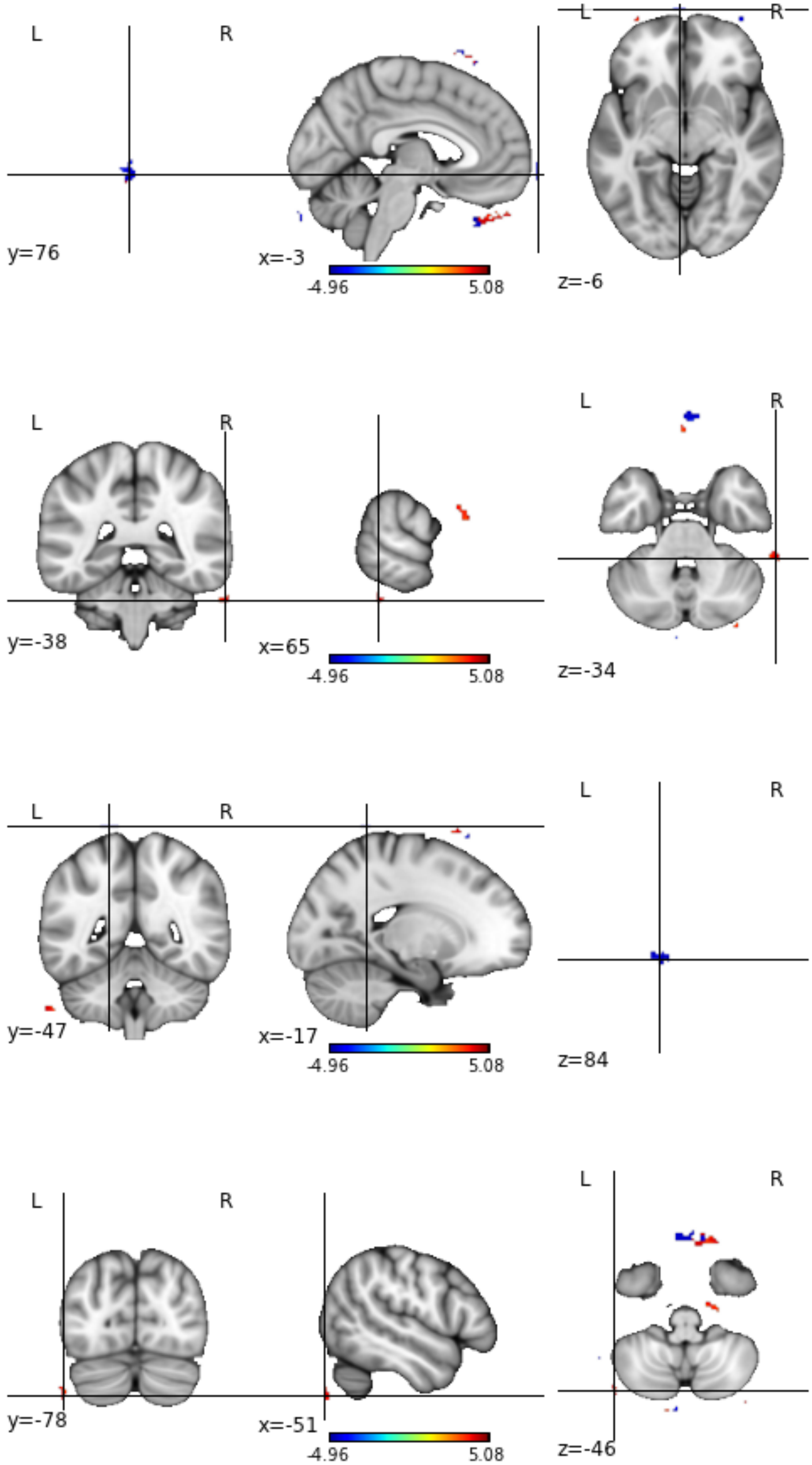

L
R
L
R
y=17
x=30
-4.96
5.08
z=76

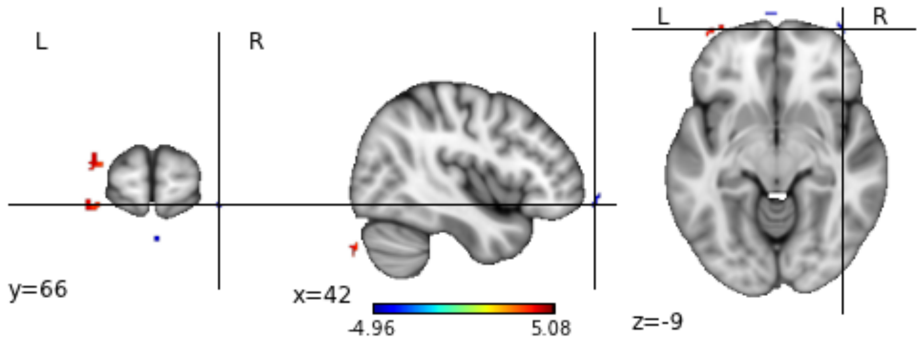

L
R
L
R
y=-88
x=41
-4.96
5.08
z=-38


L
R
L
R
y=30
x=-57
-4.96
5.08
z=39


L
R
L
R
y=29
x=-2
-4.96
5.08
z=75

```
/software/python/EPD/7.2/lib/python2.7/site-packages/numpy/ma/core.py:3785: UserWarning: Warning: converting a masked element to nan.
  warnings.warn("Warning: converting a masked element to nan.")
```





In [52]: