ELSEVIER

NeuroImage

www.elsevier.com/locate/ynimg
NeuroImage 21 (2004) 965–972

Voxel-based homogeneity probability maps of gray matter in groups: assessing the reliability of functional effects

Reza Momenan,[a,*] Robert Rawlings,[a] Grace Fong,[a] Brian Knutson,[b] and Daniel Hommer[a]

[a] *Section for Brain Electrophysiology and Imaging, LCS, National Institute on Alcohol Abuse and Alcoholism, National Institutes of Health, Bethesda, MD 20892-1256, USA*
[b] *Department of Psychology, Stanford University, Stanford, CA, USA*

Received 8 July 2003; revised 23 October 2003; accepted 29 October 2003

A subject of increasing importance in magnetic resonance imaging (MRI) is the analysis of intersubject structural differences, particularly when comparing groups of subjects with different conditions or diagnoses. On the other hand, determining structural homogeneity across subjects using voxel-based morphological (VBM) methods has become even more important to investigators who test for group brain activation using functional magnetic resonance images (fMRI) or positron emission tomography (PET). In the absence of methods that evaluate structural differences, one does not know how much reliability to assign to the functional differences. Here, we describe a voxel-based method for quantitatively assessing the homogeneity of tissues from structural magnetic resonance images of groups. Specifically, this method determines the homogeneity of gray matter for a group of subjects. Homogeneity probability maps (HPMs) of a given tissue type (e.g., gray matter) are generated by using a confidence interval based on binomial distribution. These maps indicate for each voxel the probability that the tissue type is gray for the population being studied. Therefore, HPMs can accompany functional analyses to indicate the confidence one can assign to functional difference at any given voxel. In this paper, examples of HPMs generated for a group of control subjects are shown and discussed. The application of this method to functional analysis is demonstrated.

Keywords: Voxel-based; Group analysis; Gray matter; Structural; Functional; Segmentation; Registration; Homogeneity probability map; Confidence interval

Introduction

In analyzing functional and structural differences among groups, two factors may contribute to misleading conclusions: (1) inherent constraints and limitations of existing algorithms to determine the optimal transformation matrix for realigning or registering images; and (2) intersubject variability in size and shape of brain structures and location of activations. Aside from the resolution of image acquisition systems and the magnetic field inhomogeneity, two components contribute to misregistration: (1) inherent limitations of, commonly utilized, rigid body transformations; and (2) underestimated activation due to partial voluming effects on smaller areas of activation or overestimated activation due to blurring of larger areas. As a result, the analyses of functional data fail to indicate whether observed differences in activations are solely due to functional variability or a combination of structural and functional differences. Although random error is of additional concern as a source of misinterpreting the functional data, it will not be discussed in this paper. In the latter case, the correlated noise results in a displacement of the peak activation so that the observed location only lies in some confidence region of the "true location" (Ma et al., 1999).

Nonlinear registration

One remedy for the misregistration is to realign and register images by nonlinear warping to a standard volume, for example, atlas or reference image. However, this approach may in fact compound the above-mentioned problem by making the structural and functional attributes inseparable. This happens because structural warping forces images into a reference volume and then the functional data are accordingly resampled and interpolated to fit the warped cortex resulting in an adverse effect that further complicates the analysis of functional data by making separation of structural and functional differences impossible. It is also unlikely that any existing warping method can actually align cytoarchitectural regions (Roland et al., 1997). Therefore, there is a possibility of substantial inaccuracy in aligned voxels.

Partial voluming

Another problem, partial voluming, is caused by limitations on magnetic resonance imaging (MRI) resolution and may be more pronounced in studies where brain atrophy is present. Therefore, it might commonly occur in studies focusing on aging, alcoholism, and degenerative disease such as Alzheimer's. The effect of partial voluming, resulting from brain atrophy in subjects with Alzheimer's disease, on brain glucose metabolic rates measured by positron emission tomography (PET) has

* Corresponding author. Section for Brain Electrophysiology and Imaging, LCS, National Institute on Alcohol Abuse and Alcoholism, National Institutes of Health, Room 3C114, Building 10, 10 Center Drive, MSC 1256, Bethesda, MD 20892-1256. Fax: +1-301-402-4354.
E-mail address: rezam@nih.gov (R. Momenan).
Available online on ScienceDirect (www.sciencedirect.com.)

1053-8119/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2003.10.038

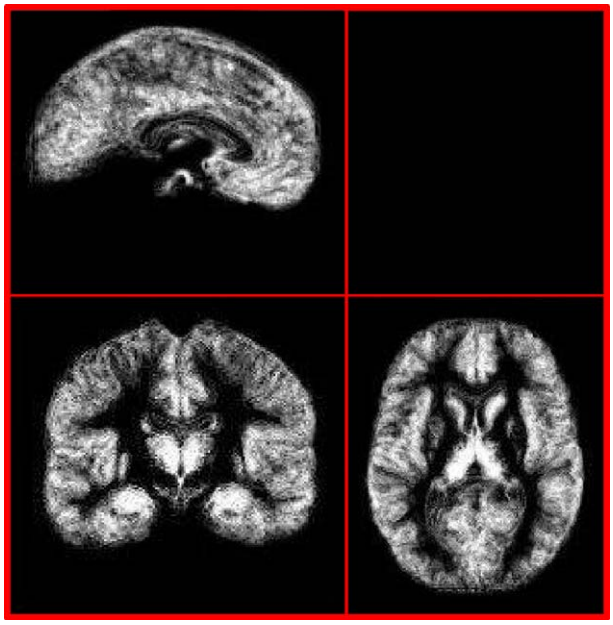

Fig. 1. Orthogonal views of the summation image for 11 subjects of a control group.

been examined (Ibanez et al., 1998), and various methods have been suggested for correcting these effects to obtain a better estimate of regional CBF and metabolism (Iida et al., 2000; Muller-Gartner et al., 1992). However, these corrections do not distinguish the effects due to functional differences from those due to structural differences.

It is a common practice that blurring (or spatial smoothing) is applied to the images as a method of compensating for activation localization inaccuracy. This method—in case of gray matter nonhomogeneity and in regions of partial voluming—can make the matter worse because larger numbers of nongray matter voxels may contribute to the observed "activation" value at each voxel.

Among the various methods for examining the relationship between structural and functional images, voxel-based morphometry (VBM) methods provide the most suitable option because they do not rely upon human inspection of difference images. Several investigators have proposed and developed VBM methods for the voxel-wise comparison of gray matter images from different groups of subjects (Ashburner and Friston, 2000). These methods utilize nonlinear warping of images and the Gaussian random field approach for group comparison. It should be noted that various nonlinear registration methods and voxel-based methods will have different degrees of regional, anatomical, and tissue compositional accuracy depending on the landmarks that they use for optimization of registration or comparison of voxels (Salmond et al., 2002).

Intersubject variability

Intersubject variability in the location of activation poses a more complex problem. This problem arises because even in subjects with optimally matched structures, the actual site of activation for a given stimulus or behavior may differ. Therefore, even nonlinear warping cannot guarantee separation of functional and structural differences in the case of variability in activation sites between subjects in a group. Other investigators (Salmond et al., 2002) have shared this concern by stating that even if anatomical differences were removed by exact registration or normalization it may artificially alter the tissue composition of the images and hence fail to prevent matching dissimilar anatomies based on the tissue intensity.

Another possible approach to addressing localization while avoiding nonlinear transformations involves utilization of shape descriptors—curvature(s), length, height, thickness, etc. followed by comparison of corresponding regions or structures. However, this approach encounters a major problem: for each region or structure, there must be enough descriptors to allow unique identification of such region or structure. This is not only at times computationally costly but may also not be possible to identify these descriptors (particularly for small structures). Because some of these descriptors are not currently automated, they therefore require tedious manual processing of the images vulnerable to operator variability.

In this paper, we propose a voxel-based method to generate homogeneity probability maps (HPMs) within groups. This method allows structural comparison of the cortical surface and gray matter,

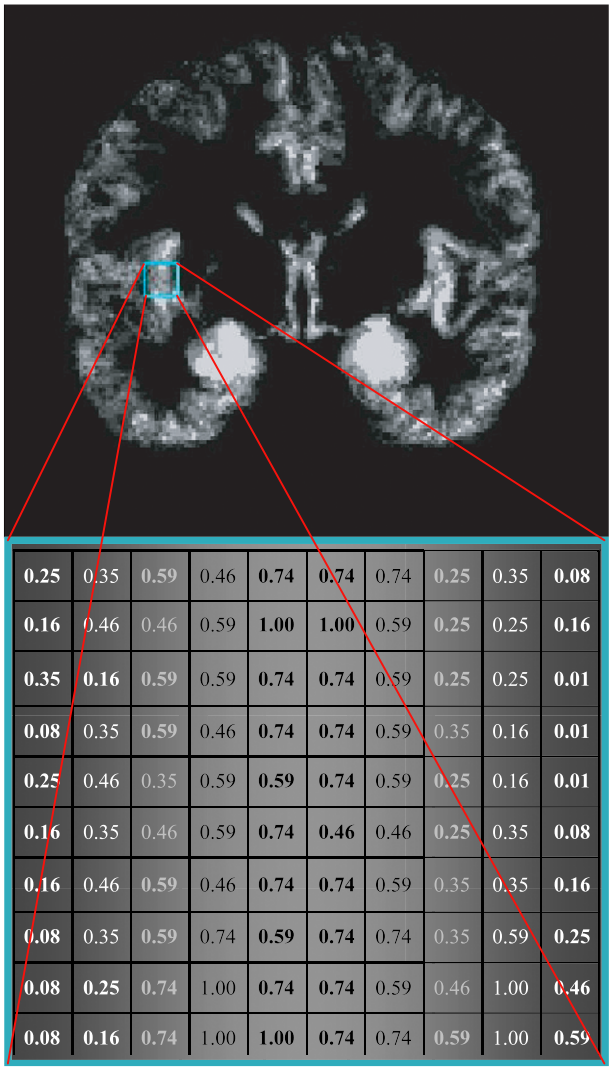

Fig. 2. A sample of the map of estimated probability $\hat{p}$ of similar tissue type (gray matter).

in general, regardless of the type of registration and segmentation methods used. We utilize this method not as a correction mechanism for variation in the activation sites but rather as a complementary tool for evaluating the structural reliability of functional differences.

Method

The objective of our method is to address registration limitations and partial voluming effects by localizing regional or structural homogeneity (and by contrast, heterogeneity) for any given tissue type (e.g., gray matter or white matter). In the case of functional images, we provide homogeneity maps of gray matter over the entire brain as an index of reliability that the observed functional differences are actually due to functional rather than structural differences. In the following sections we describe this method, which includes image registration, image segmentation, generation of a density volume, and finally generation of homogeneity probability map (HPM) for a group of subjects.

Image realignment or registration and segmentation

All images are first coregistered to a standard (e.g., Talairach) volume. The individual registered images are then segmented into various tissue types (e.g., CSF, gray matter, and white matter). The segmentation allows investigators to avoid influence of intensity variations in the images due to differences in acquisition parameters. Clearly, both registration and segmentation algorithms affect the outcome of this method as they do in all analyses. In this paper, we utilize analysis of functional neural images (AFNI) (Cox, 1996) software for realignment or registration along with an intensity based (Momenan et al., 1997) algorithm for segmentation.

Density volume

A binary mask B_m of gray matter is then generated from the segmented volume of each brain:

$$B_m(i,j,k) = \begin{cases} 1 & \text{pixel} \subset \text{gray matter} \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

where $B_m(i,j,k)$ is a binary value in each volume, m, in the group at voxel location i,j,k. In this fashion, a voxel which has any amount of gray matter will be assigned the value 1. Otherwise, the voxel will be assigned the value 0. As a result, we create a binary image of the gray matter in the brain and treat the resulting data in binomial mode.

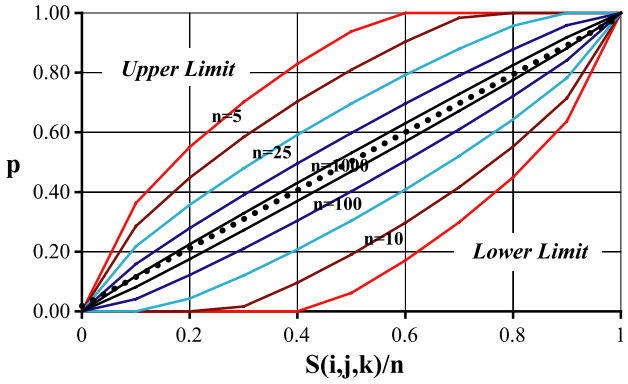

Fig. 4. Plot providing a 95% confidence interval for p in binomial sampling as a function of $S(i,j,k)$ and number of subjects n.

In the proposed method, we use addition instead of subtraction to compare cortical regions. To do this, registered binary images of gray matter are summed together for all corresponding voxels. This avoids effects that are encountered with subtraction methods (Fox and Freeborough, 1997) (i.e., having to deal with both positive and negative values). As a result, a single density volume S is generated (Fig. 1) in which the voxel where the gray matter in each volume overlaps with the others has a higher intensity than those voxels with fewer overlapping gray matter. Therefore, the density volume is computed and generated from:

$$S(i,j,k) = \sum_{m=1}^{n} W_{i,j,k} \times B_m(i,j,k) \quad (2)$$

where $S(i,j,k)$ is the corresponding voxel in the density volume at voxel i,j,k and $w_{i,j,k}$ is the weighting factor between 0 and 1 at each voxel. The weighting factor is determined by the segmentation algorithm. $w_{i,j,k}$ is zero for voxels with no gray matter and equal to 1 if the voxel is determined to be 100% gray matter. In a deterministic segmentation method (Momenan et al., 1997), $w_{i,j,k}$ is always equal to 1 because the binary image $B_m(i,j,k)$ only represents the gray matter. In a probabilistic segmentation method, $w_{i,j,k}$ is equal to the

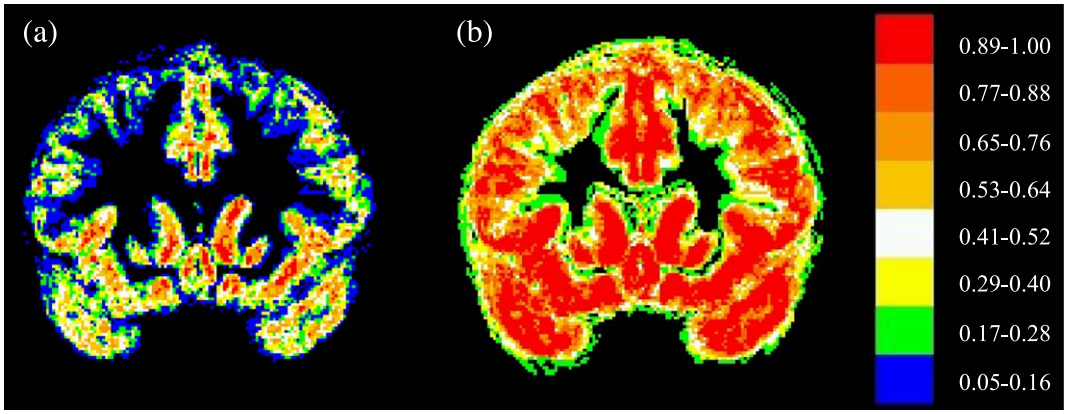

Fig. 3. The lower (a) and upper (b) confidence limits for the group based on a 95% confidence interval.

value of the proportion of the voxel that is determined to be gray matter. Therefore, the value of each density voxel ranges from zero (no gray matter in all subjects) to n (100% gray matter in all subjects) for a given voxel.

Homogeneity probability map (HPM)

The next step is to generate a homogeneity probability map (HPM) of the group. The HPM indicates the probability that each voxel in every subject of the group belongs to a given tissue type (in this case gray matter; see Fig. 2).

To this end, the estimated probability of a given voxel in the group being gray matter is computed by:

$$\hat{p}(i,j,k) = \frac{S(i,j,k)}{n} \tag{3}$$

where n is the number of subjects in the group.

We can then compute the approximate confidence interval for the probability of any given voxel being gray matter for the entire group from (Goodall, 1995):

$$\hat{p} - Z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} < p < \hat{p} + Z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \tag{4}$$

where α is the chosen significance level, $Z_{\alpha/2}$ denotes the $N(0,1)$ at that level, and p is the true probability of a given voxel being composed of the desired tissue type. Note that this is only one of several methods of approximating a confidence interval. Other methods can provide more accurate confidence intervals for small sample sizes (Johnson and Kotz, 1969).

The actual probability of any given voxel being gray matter lies somewhere between the lower and upper limits of this interval (Fig. 3). We chose the 95% confidence interval for our analysis.

Fig. 3a shows the lower (or conservative) confidence limit. As a result, fewer voxels have high probability of being gray matter. However, the upper (or liberal) confidence limit shown in Fig. 3b indicates more homogenous voxels. Note that the lower and upper limits approach convergence as: (1) the structural (and tissue type) variability among subjects decreases ($S(i,j,k) \rightarrow n$); and more importantly as (2) n, the number of subjects, increases (Fig. 4).

Subjects and scans

Twelve physically and psychiatrically healthy normal volunteers (six women and six men; right handed; mean age: 31) participated in a previously described (Knutson et al., 2003) parametric monetary incentive delay (MID) task in which they pressed a button to either gain or avoid losing monetary incentives of varying amounts (US$5.00, US$1.00, US$0.20, US$0.00).

During the task, 432 T_2^*-weighted echo-planar MR volumes (TR = 2 s, TE = 40 s, flip angle = 90°) were acquired with 22 sagittal slices (3.752 × 3.8 mm) centered on the intrahemispheric fissure using a 1.5-T magnet.

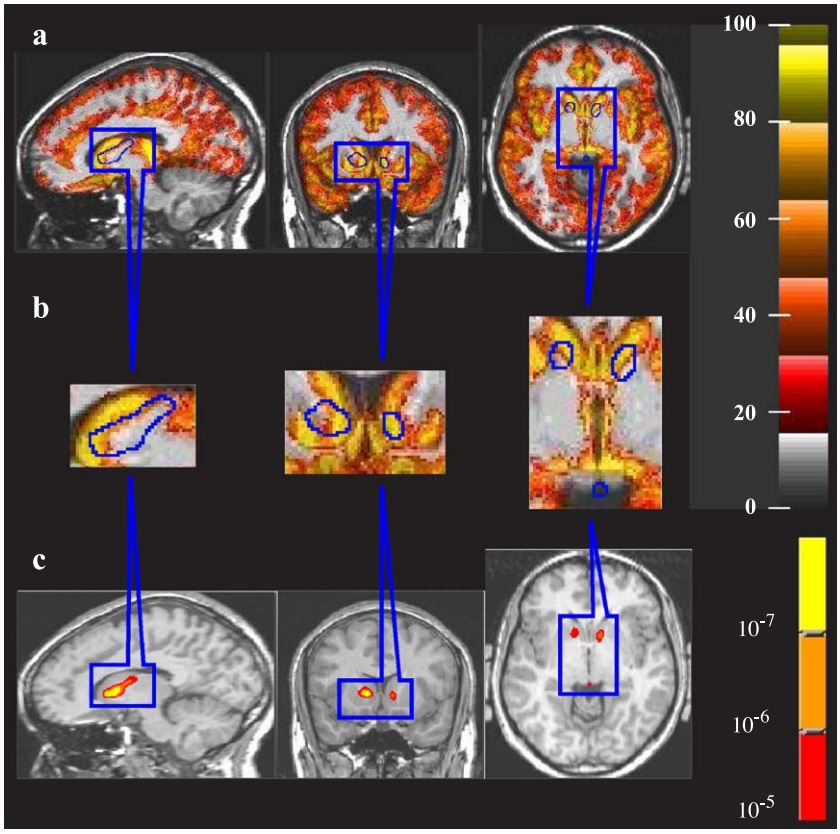

Fig. 5. The lower homogeneity probability map for (a) 11 control subjects with the color bar showing percent probability with the boundary of functional activation overlaid on the HPMs; (b) the magnified region of activation overlaid on HPM; (c) the corresponding z-scores from the functional analysis for incentive anticipation tasks of the control subjects.

The full-brain high resolution coronal structural volume was collected using T_1-weighted spoiled grass sequence (TR = 100 ms, TE = 7 ms, flip angle = 90°) of 256 × 256 × 124 matrix with voxel size of 0.9375 × 0.9375 × 2.0 mm^3. All volumes were Talairach registered using AFNI (Cox, 1996) resulting in isotropic volumes with 161 × 151 × 191 matrix of 1 × 1 × 1 mm^3 voxels. Only 11 of the above-mentioned normal subjects (six females) were used for this voxel-based analysis to match 11 segmented images available at the time of this publication.

Results

Structural analysis

In the present analyses we selected the lower limit of the confidence interval to be conservative. As shown in coronal views of Fig. 3, in the cortex, where variability is most expected, the homogeneity probability is smaller. On the other hand, subcortical nuclei show greater homogeneity. The homogeneities along Anterior Commissure (AC), Posterior Commissure (PC), midsagittal cortical surface, striatum, and thalamus are also large (Figs. 5a and 6a and HPMs of control subjects, respectively).

Accordingly, as indicated in Figs. 5a and 6a, the areas of high confidence (lower confidence limit above 70%) in these groups include both subcortical and cortical regions. Cortical regions included some of functionally relevant structures such as mesial prefrontal cortex, as well as other cortical areas spanning from the temporal lobe to the genu of the corpus callosum. Subcortical regions included the insula (left: −39, 7, 9; right: 39, 7, 9), hippocampus (left: −30, 24, −9; right: 30, 24, −9), caudate (left: −11, −7, 9; right: 11, −7, 9), thalamus (left: −12, 19, 8; right: 12, 19, 8), amygdale (left: −20, −4, −19; right: 20, −4, 19), parahippocampal gyrus (left: 26, −20, −10; right: −26, −20, −10), and fusiform gyri (left: −40, 48, −16; right: 40, 48, −16).

In contrast, the cortical regions in the parietal and occipital lobes (with the exception of V1 in the visual cortex) did not warrant high levels of confidence. Neither did the dorsolateral and orbitofrontal cortices in the frontal lobe. Subcortical structures such as the putamen (left: −24, 0, 3; right: 24, 0, 3) and globus pallidus (left: −17, 4, −2; right: 17, 4, −2) also elicited lower confidence. Also note that in both conservative and optimistic limits of the images shown in Fig. 3—as expected—the homogeneity near the intersection of cortical surfaces with cerebrospinal fluid and white matter is low.

Functional analysis reliability check

One motivation for developing this approach is to determine which functional differences may derive from structural inhomogeneities. Therefore, we propose the utilization of

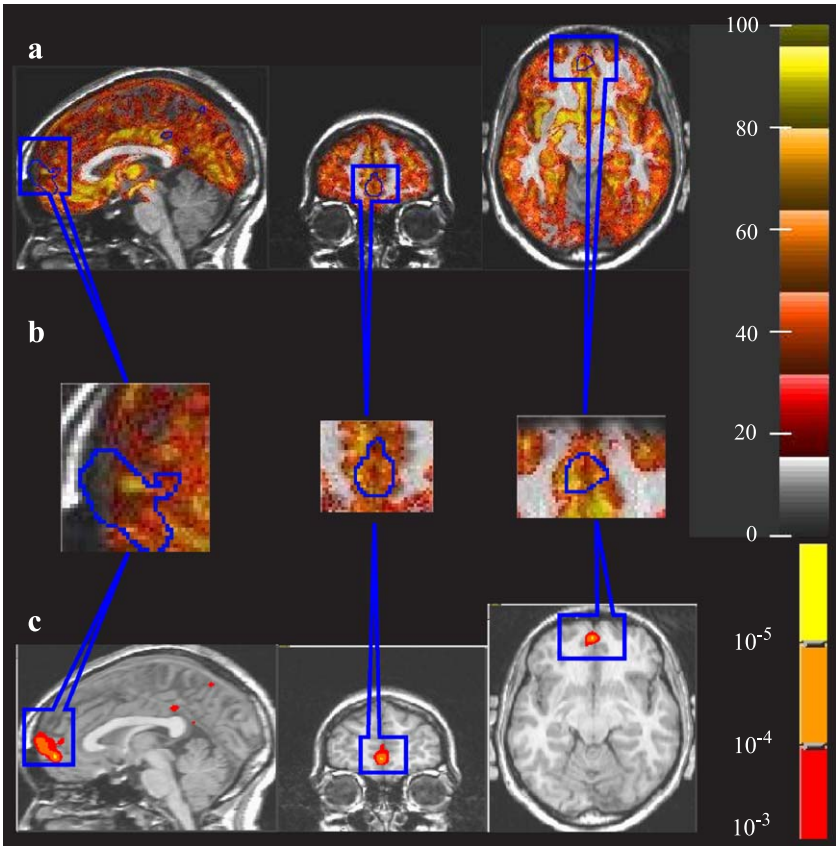

Fig. 6. The lower homogeneity probability map for (a) 11 control subjects with the color bar showing percent probability with the boundary of functional activation overlaid on the HPMs; (b) the magnified region of activation overlaid on HPM; (c) the corresponding z-scores from the functional analysis for incentive feedback vs. baseline of the control subjects.

homogeneity probability maps (HPMs) as a supplementary index of the reliability of functional differences. It is important to note that HPMs do not correct or evaluate the method selected for functional analysis, but rather, evaluate how reliably the functional differences reflect functional rather than structural differences.

Functional analysis

Functional analyses focused on changes in blood oxygen level-dependent (BOLD) contrast that occurred as subjects anticipated gaining or losing money, and received feedback indicating that they had gained or lost money. The experiment was set up and analyses were performed as described by Knutson et al. (2003) as follows.

All analyses were conducted using AFNI (Cox, 1996). For preprocessing, voxel time series were interpolated to correct for nonsimultaneous slice acquisition within each volume (using sinc interpolation and the rightmost slice as a reference), concatenated across both task sessions, and then corrected for three-dimensional motion (using the third volume of the first session as a reference). Motion correction estimates were then visually inspected to confirm that no participant's head moved more than 1.5 mm in any dimension from one volume acquisition to the next.

Preprocessed time series data for each individual were first analyzed with multiple regression (Neter et al., 1996) to allow covariance of "nuisance" variables related to head motion and scanning session in modeling trials of interest. Both anticipation and outcome stages were modeled for each trial type and convolved with a gamma-variate function to model a prototypical hemodynamic response before inclusion in the regression model (Cohen, 1997). Resulting functional maps were coregistered with structural maps, spatially normalized by rigid body transformation to Talairach space, and slightly spatially smoothed (FWHM = 4 mm) to account for anatomical variability. Using a random-effects type analysis, orthogonal general linear tests were then conducted to compare conditions of interest for all subjects as a group. T-statistic maps were then converted to Z-scores, spatially blurred (FWHM = 4 mm), and normalized. Active voxels were defined as those showing a significant effect using a threshold of $P < 0.00001$, uncorrected, or $P < 0.05$, corrected. Comparisons of interest included: (1) anticipation of monetary gain vs. anticipation of no monetary outcome; (2) anticipation of monetary loss vs. anticipation no monetary outcome; (3) "hit" vs. "miss" outcomes on potential gain trials; and (4) "hit" vs. "miss" outcomes on potential loss trials. Table 1 shows the group maximum Z scores and Talairach coordinates of activation foci as described by Knutson et al. (2003). Note that two (rightmost) columns are added that indicate the minimum and maximum probabilities, respectively, of the center of each focus being gray matter. Hence, the table provides an estimate of how reliably the Z scores results from functional activation rather than structural differences.

Figs. 5a and 6a show the homogeneity probability maps (HPMs) for healthy volunteer functional maps shown in two separate functional magnetic resonance imaging (fMRI) analyses of Figs. 5c and 6c (Knutson et al., 2003), respectively. For clarity of comparison, the blue outlines overlaid in Figs. 5a and 6a indicate the boundary of activation regions and Figs. 5b and 6b are their magnified copies, respectively. As shown in Figs. 5b and 6b, the activation regions largely overlap with the high HPMs, but may have subregions of activations that have low HPMs. Table 1 shows the HPM confidence limits for the focal or peak point of the above-mentioned analyses. In Table 1, the focal point of 12 out of 20 regions has low HPMs. However, if the HPM method had been available then the 12 low HPM points would not have been selected as the focal point of

Table 1
Group maximum Z-scores and Talairach coordinates of activation foci ($P < 0.0001$, uncorrected; $n = 12$)

Area (Brodmann's area)	Anticipation: potential gain vs. no outcome		Outcome: gain vs. no outcome		Confidence limits for 0.95 confidence interval	
	Max Z	TC (R,A,S)	Max Z	TC (R,A,S)	Lower limit	Upper limit
R. Ant. Insula (13)	4.30	36, 14, 1			0.59	1.00
L. Ant. Insula (13)					0.46	1.00
R. NAcc	**4.89**	**11, 12, 0**			0.35	0.92
L. NAcc	5.11	−9, 10, 0			0.59	1.00
R. Caudate	5.76	10, 9, 4			0.16	0.75
L. Caudate	6.42	−8, 1, 7			0.00	0.65
R. Putamen					0.35	0.92
L. Putamen					0.00	0.26
Thalamus	4.86	−3,−13, 13			1.00	1.00
R. Amygdala					0.74	1.00
L. Amygdala					1.00	1.00
Mes. Prefrontal Ctx (10/32)			**4.45**	**1, 53−6**	0.59	1.00
Frontal Pole (10)			4.33	2, 65, 7	0.00	0.41
Ant. Cingulate (24)					0.00	0.54
Post. Cingulate (26/30)			4.03	5, −51, 22	0.25	0.84
Parietal Ctx (7)			3.91	4, −62, 53	0.25	0.84
Mes. Prefrontal Ctx (32)	4.51	3, 27, 35			0.59	1.00
Sup. Motor Area (6)	3.96	0,−4, 49			0.35	0.92
L. Motor Ctx. (4)	4.94	−28,−55, 42			0.46	0.99
Cerebellar Vermis[a]	4.67	0,−70,−26			–	–

Boldface indicates foci that were used to construct volumes-of-interest. TC = Talairach coordinates, R = right, A = anterior, S = superior.
[a] The gray matter and white matter segmentations for cerebellar areas were not available.

activation. From Figs. 5b and 6b, it is clear that there are alternative points in these regions of activation with high HPMs. Therefore, the maximum activations among these high HPMs should be chosen. In this sense, these results confirm the functional analysis of anticipation (Fig. 5c) and feedback (Fig. 6c) as demonstrating functional activation at the foci, but caution against considering the activation foci on the edges of gray matter structures. These results demonstrate the usefulness of homogeneity probability map of gray matter by identifying the heterogeneity of activation areas (e.g., edges, shell vs. core of structure) and, therefore, helping investigators avoid confusing structural variations with functional differences.

Conclusion

We have described a voxel-based method for generating homogeneity probability maps (HPMs) to evaluate the homogeneity of tissues and for indexing structural differences within groups. We have demonstrated that this method can also be used to supplement functional data analysis, indicating the structural or tissue homogeneity of the areas of significant functional activations.

In this method, a confidence interval—as opposed to a specific measure—of homogeneity of each voxel is provided. To be conservative, we have chosen the lower limit as the index of reliability. This measure should be used as an index of how much the normalization and segmentation steps as well as the partial voluming and intersubject variability cumulatively affect the homogeneity of a given voxel within a group.

We emphasize that this complementary method can be used regardless of registration, segmentation, and normalization methods that are applied before functional data analysis. HPMs are not intended as a method for evaluating the performance of the aforementioned methods, but are an indicator of tissue homogeneity.

Unlike other similar methods, this technique does not introduce any inherent shortcomings of its own to within and between groups comparisons. The method does, however, inherit two confounding parameters. First, the rigid body registration must adequately match group images into the reference volume. Second, the segmentation program must perform adequately. However, any problems with registration and segmentation arise before implementation of the suggested HPM method. Thus, the proposed method describes the analytical reliability, but not adequacy. Therefore, as the registration and segmentation performance in terms of accuracy and performance are increased, the HPM's would similarly be more precise.

As for greater homogeneity along the AC–PC line and the midsagittal surface, most programs such as AFNI require that these landmarks be identified. This may contribute to better registration of images along the AC–PC line and the midsagittal slice. Therefore, notwithstanding inherent variability among subjects, there is still a good intersubject structural or tissue homogeneity in the striatum and thalamus. Another contributing factor may be the increased homogeneity of the center of the magnetic field.

Using the HPM method, one can also use the areas of high concentration of gray matter as a mask for functional analysis, hence increasing the power of analysis by reducing the number of statistical tests (as opposed to testing the whole brain or an entire gray matter volume). This will also result in computationally less expensive processes because fewer voxels will be tested.

This method also confirms inherent advantages of single-subject analyses. Specifically, the findings suggest that in single group analyses, investigators should confine their analyses to overlapping regions of similar tissue type. This could result in smaller regions of detected activation. Therefore, whole brain may only be possible in single-subject analyses.

Two caveats are in order regarding the HPM method. First, although it is likely that in the regions of greater gray matter homogeneity the statistical differences derive from functional activations in the group, the reverse is not necessarily true. Specifically, in regions of low gray matter homogeneity, one cannot determine that the source of a statistical difference is due to structural differences, functional differences, or both. Second, investigators must choose the minimum acceptable probability threshold for the HPMs of gray matter to accept or reject a functional map of the corresponding regions.

It should be noted that like other existing methods, intersubject variability of the location of functional activation cannot be addressed by the HPM method proposed in this paper.

As a natural extension of group HPMs, a work in progress involves generating between-group HPMs for assessing the between-group functional differences. In this case, there can be additional homogeneity problems beyond the within-group analysis that should be tested to generate between-group HPMs.

Future work will also include improvement of the methods for calculating the confidence intervals as well as investigating the expansion of the HPM method so that a ''correction factor'' or ''correction covariate'' can be determined and directly applied to functional data analysis.

Acknowledgments

The authors would like to thank Shannon Bennett for tremendous assistance in scanning the subjects. We also thank Shannon Bennett and Shaun Steuer for some of the postprocessing of images. We also thank Michael Kerich for his computer support and assistance. Data collection was supported by NIAAA-LCS intramural program, and Brian Knutson was supported by grants from NIMH and NARSAD.

References

Ashburner, J., Friston, K., 2000. Voxel-based morphometry—The method. NeuroImage 11, 805–821.

Cohen, M., 1997. Parametric analysis of fMRI data using linear systems methods. NeuroImage 6, 93–103.

Cox, R., 1996. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. Comput. Biomed. Res. 29, 162–173.

Fox, N., Freeborough, P., 1997. Brain atrophy progression measured from registered serial MRI: validation and application to Alzheimer's disease. J. Magn. Reson. Imaging 7 (6), 1069–1075.

Goodall, G., 1995. Don't get t out of proportion. Teach. Stat. 17 (2), 50–51.

Ibanez, V., Pietrini, P., et al., 1998. Regional glucose metabolic abnormalities are not the result of atrophy in Alzheimer's disease. Neurology 50 (6), 1585–1593.

Iida, H., Law, I., et al., 2000. Quantitation of regional cerebral blood flow corrected for partial volume effect using O-15 water and PET: I. Theory, error analysis, and stereologic comparison. J. Cereb. Blood Flow Metab. 20 (8), 1237–1251.

Johnson, N., Kotz, S., 1969. Discrete Distributions. Wiley, New York.

Knutson, B., Fong, G., et al., 2003. A region of mesial prefrontal cortex

tracks monetary rewarding outcomes: characterization with rapid event-related fMRI. NeuroImage 18, 263–272.
Ma, L., Worsley, K., et al., 1999. Variability of spatial location of activation in fMRI and PET CBF images. NeuroImage 9 (6), 2786 (Poster 178).
Momenan, R., Hommer, D., et al., 1997. Intensity-adaptive segmentation of single-echo T_1-weighted magnetic resonance images. Hum. Brain Mapp. 5, 194–205.
Muller-Gartner, W., Links, J., et al., 1992. Measurement of radiotracer concentration in brain gray matter using positron emission tomography: MRI-based correction for partial volume effects. J. Cereb. Blood Flow Metab. 12, 571–583.
Neter, J., Kutner, M., et al., 1996. Applied Linear Statistical Models. Irwine, Chicago.
Roland, P., Geyer, S., et al., 1997. Cytoarchitectural maps of the human brain in standard anatomical space. Hum. Brain Mapp. 5, 222–227.
Salmond, C., Ashburner, J., et al., 2002. The precision of anatomical normalization in the medial temporal lobe using spatial basis functions. NeuroImage 17 (1), 507–512.