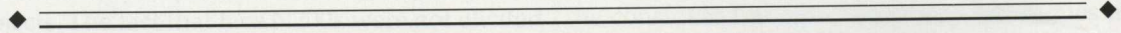


Functional Volumes Modeling: Theory and Preliminary Assessment

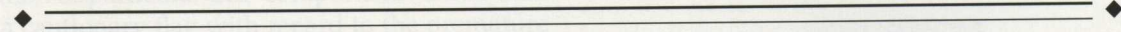
Peter T. Fox,* Jack L. Lancaster, Lawrence M. Parsons, Jin-Hu Xiong, and Frank Zamarripa

Research Imaging Center, University of Texas Health Science Center at San Antonio, San Antonio, Texas 78284-6240



Abstract: A construct for metanalytic modeling of the functional organization of the human brain, termed functional volumes modeling (FVM), is presented and preliminarily tested. FVM uses the published literature to model brain functional areas as spatial probability distributions. The FVM statistical model estimates population variance (i.e., among individuals) from the variance observed among group-mean studies, these being the most prevalent type of study in the functional imaging literature. The FVM modeling strategy is tested by: (1) constructing an FVM of the mouth region of primary motor cortex using published, group-mean, functional imaging reports as input, and (2) comparing the confidence bounds predicted by that FVM with those observed in 10 normal subjects performing overt-speech tasks. The FVM model correctly predicted the mean location and spatial distribution of per-subject functional responses. FVM has a wide range of applications, including hypothesis testing for statistical parametric images. *Hum. Brain Mapping* 5:306–311, 1997. © 1997 Wiley-Liss, Inc.

Key words: FVM; functional brain model; statistical model; Talairach; spatial hypothesis testing



INTRODUCTION

Mapping the functional organization of the human brain is a highly productive, rapidly growing field. Functional imaging studies are localizing the neural populations performing specific mental operations in the domains of perception, action, cognition, and emotion. A considerable portion of the functional-imaging literature has been reported as response coordinates (loci^{x-y-z}) referenced to the Talairach Atlas [Talairach and Tournoux, 1988]. Standardized anatomical referencing makes this literature uniquely well-

sued for metaanalysis. On the other hand, the majority of studies reporting on Talairach space do so for the purpose of creating group-mean, statistical parametric images (SPI[n]), pooling n subjects within the standardized space. Intersubject averaging typically precludes quantifying intersubject variability in functional anatomy. Functional volumes modeling (FVM) addresses the latter shortcoming through exploiting the former strength. Specifically, FVM estimates individual (per-subject) variability in the brain locations of specific mental operations through an analysis of variability among group-mean studies in the reported literature, and allows modeling of groups of various sizes based on this estimate.

THEORY AND STATISTICAL MODEL

FVMs model brain functional areas as bounded volumes. The bounds of an FVM express confidence

Contract grant sponsors: NIMH and NIDA; Contract grant number: Human Brain Project P20 MH/DA52176.

*Correspondence to: Peter T. Fox, Research Imaging Center, University of Texas Health Science Center, 7703 Floyd Carl Drive, San Antonio, TX 78240-6240.

Received for publication 9 May 1997; accepted 12 May 1997

Model size

limits for the spatial distribution of the functional area. The most general form for expressing confidence bounds is based on population parameters and relates to the probabilities for individual samples from the population. For present purposes, the individual sample is a single subject. Insofar as the functional imaging literature reports group-mean studies (SPI[n]), single-subject probabilities are not reported. They can, however, be estimated from the variance among some number of group-mean studies, by means of a statistical model (i.e., the FVM statistical model). A basic premise of the FVM construct is that the variance among group-mean images (SPI[n]) will be less than that of single-subject images (SPI[1]), with the reduction in variance being predicted by the central limit theorem.

An FVM model has two forms: a general and a specific form. A general-form FVM has six parameters: three defining the means (μ_x, μ_y, μ_z), and three defining standard deviations ($\sigma_x, \sigma_y, \sigma_z$) of the population. These six parameters are derived from the literature and computed with Equation (1) (means) and Equation (2) (standard deviations). All six general-form parameters are expressed as millimeters within the Talairach space. An FVM is applied using two additional parameters, which specify the size of the group to which the FVM will be applied (n) and the confidence limits chosen by the user (m; expressed as the number of standard deviations).

Model location

An unbiased estimate of the population mean is calculated using Equation (1),

$$\langle Q \rangle = \frac{n_1}{N} \langle q_1 \rangle + \frac{n_2}{N} \langle q_2 \rangle + \dots + \frac{n_K}{N} \langle q_K \rangle \quad (1)$$

where $\langle Q \rangle$ is the weighted mean of the centroid coordinates $\langle q_i \rangle$ from K experiments. The $\langle q_i \rangle$ represent the mean x, y, or z coordinates as reported in a published SPI[n]. For the weighted mean, $\langle Q \rangle$, to estimate the population mean, the weight factors must be calculated as $\langle n_i/N \rangle$, where $\langle n_i \rangle$ is the number of subjects in a group and $\langle N \rangle$ is the total number of subjects in all groups. Equation (1) is used to estimate μ_x, μ_y , and μ_z for an FVM by successively substituting $\langle x_i \rangle, \langle y_i \rangle$, and $\langle z_i \rangle$ for the $\langle q_i \rangle$'s. Values for $\langle x_i \rangle, \langle y_i \rangle, \langle z_i \rangle$, and $\langle n_i \rangle$ are generally reported in tables accompanying each reported SPI. Thus, the necessary data are readily available in a usable format and rely only on the hypotheses of well-behaved averaging and the comparability of the coordinates across centers and across modalities.

An unbiased estimate of population variance is calculated using Equation (2a),

$$\sigma_q^2 = \frac{N}{N-1} \sigma_{(q)}^2 + \frac{N-K}{N-1} \sigma_{q_i}^2 \quad (2a)$$

which simplifies to Equation (2b),

$$\sigma_q^2 = \sigma_1^2 + \sigma_2^2. \quad (2b)$$

In Equation (2a), the population variance for coordinate q is σ_q^2 . This term is calculated using the raw variance among the K groups in the input data ($\sigma_{(q)}^2$) and the average variance estimated across experiments $\sigma_{q_i}^2$. The first variance term in Eq. (2b), (σ_1^2), is mostly attributed to interlaboratory sources, while the second term, (σ_2^2), is attributed to intralaboratory sources of variance. Interlaboratory variance is largely methodological; intralaboratory variance is largely physiological. This distinction is used in model application.

Model application

For specific applications of an FVM model, the following modification of Equation (2b) is used to estimate variance for group size n from the estimated population variance:

$$\sigma_q^2(n) = \sigma_1^2 + \frac{\sigma_2^2}{n}. \quad (3)$$

Here n is the size of the experimental group to be modeled. Note that the interlaboratory variance component (σ_1^2) does not change with group size. The model is then formulated with a mean location of $\langle Q \rangle$ (Equation 1) and bounding limits at $\pm m$ times $\sigma_q(n)$ (Equation 3). Here m is the number of standard deviations from the mean (i.e., the confidence limits). There are two general approaches to using this FVM model. In the first case, when using the model to evaluate an experiment in a single laboratory, the variance is calculated for group size n, but the interlaboratory component (σ_1^2) is set equal to zero. In the second case, when using the model to search the published literature, the variance is calculated with group size n = 1 as indicated in Equation (3). We are assuming that the mean location determined for the general form of the model with data from many laboratories is a good estimate of the population mean and is suitable for the specific form without modification.

METHODS

The FVM construct was tested as follows. First, an FVM of a single functional area, the mouth area of the primary motor cortex, was constructed from the published literature. Second, the model was used to predict the spatial distribution of per-subject (SPI[1]) and group-mean (SPI[10]) data. Third, the M1-mouth region was identified in 10 individual subjects. Fourth, the assumption of well-behaved image averaging was tested, using the newly acquired data. Fifth, the spatial distribution of M1-mouth loci^{x-y-z} observed in the newly acquired PET data were compared to those predicted from the model.

M1-mouth model

An FVM model of the mouth region of the primary motor cortex in the left cerebral hemisphere was constructed, as follows. Input data from the reported literature were retrieved using BrainMap[™] software tools and database [Fox and Lancaster, 1995; Lancaster et al., 1997]. Model-input data were limited to: 1) group-mean studies, 2) normal subjects, 3) using tasks including overt speech, 4) reporting an activation identified as primary motor cortex, and 5) studies not performed at the Research Imaging Center. Data retrieval was initiated by a BrainMap[™] query on behavior:response:speech. The retrieved studies were filtered for "contrasted" speech, by which is meant that the task state included overt speech while the control state did not. This search and filter strategy identified six SPI[n], ranging in group-size (n) from 8–35, and totaling 102 subjects [Petersen et al., 1988; Petrides et al., 1993; Paus et al., 1993; Andreasen et al., 1995; Bookheimer et al., 1995; Buckner et al., 1995]. From these input data, the FVM model of the mouth region of the primary motor cortex in the left cerebral hemisphere was constructed, using Equations (1) and (3). The resulting mean and standard deviations for the general model are given in Table I. For each axis (x, y, z), confidence bounds were separately computed for the 50th (z = 0.68), 68th (z = 1.0), and 95th (z = 1.6) percentiles. For the volume (all axes), confidence bounds were computed for the 50th (z = 1.26), 68th (z = 1.55), and 95th (z = 2.39) percentiles.

M1-mouth mapping

Primary motor cortex for mouth was mapped in 10 right-handed normal men, for the purpose of testing the above-described FVM model. Mapping was performed with positron-emission tomography (PET)

TABLE I. Comparison of M1-mouth FVM model and M1-mouth mapping*

	Xmean	Xsd	Ymean	Ysd	Zmean	Zsd
Model	-45	4.9	-10	5.0	36	6.4
SPI[1]	-46	4.0	-10	6.5	39	4.8
SPI[10]	-47		-10		36	

* The model is derived from a metaanalysis of published, group-mean studies. The mapping data are from single-subject analysis of an overt speech task. All values are in millimeters, referable to Talairach and Tournoux [1988].

blood-flow imaging, using H₂ ¹⁵O as the blood-flow tracer. Each volunteer underwent a series of nine PET scans: three scans per condition in each of three conditions. The rest condition was visual fixation. Both task states involved overt reading of visually presented paragraphs. In one task condition, the subject heard the paragraph as he read it (Chorus); in the other task condition no auditory input was provided (Solo). Group-mean analyses of these data were reported by Fox et al. [1996].

Using intrasubject averaging, a z-score SPI was made for each subject, yielding 10 SPI[1]. In addition, a single SPI[10] was created, averaging all 10 subjects. Both the SPI[10] and the 10 SPI[1] were created within the Talairach space, using SN for spatial normalization [Lancaster et al., 1995]. Within each of the 10 SPI[1] and the single SPI[10], the most intense response above a z-score of 2.0 ($P < 0.01$) on the lateral brain surface in the region of the central sulcus was designated as the putative M1-mouth area. Responses (loci^{x-y-z}) were specified by the Talairach coordinates of the center-of-mass of the activated area. The loci^{x-y-z} for the 10 SPI[1] were compared to the locus^{x-y-z} for the SPI[10] and to the M1-mouth FVM.

RESULTS

Loci^{x-y-z} for left M1-mouth were identified in each of the 10 SPI[1] and in the single SPI[10] (Table I). For each axis (x, y, z), the mean of the 10, left M1-mouth loci^{x-y-z} (from the 10 SPI[1]) was within 3 mm (mean 1.3 mm) of the M1-mouth locus^{x-y-z} in the group-mean image (SPI[10]). This is an initial confirmation that response locations observed in group-mean images closely approximate the numeric average of response locations in individual subjects.

For each axis (x, y, z), the standard deviations of the 10 single-subject M1-mouth loci^{x-y-z} were 6.5 mm or less (mean, 5.0 mm; Table I). Modeled means and standard deviations (derived solely from group-mean

TABLE II. Test of MI-mouth FVM model on 27 coordinates, from 9 normal individuals*

Modeled probability	Percent of data bounded by model			vol
	x	y	z	
50%	60%	40%	60%	30%
68%	80%	40%	90%	70%
95%	100%	90%	100%	100%

* Percentages indicate the fraction of data accounted for by the model for each predicted probability bound in x, y, and z and for the entire volume (vol).

data) were very close to those observed in the single-subject test data (Table I). Modeled confidence bounds for the 50th, 69th, and 95th percentiles were also close to those observed in single-subject test data (Table II). The spatial distribution of the data is displayed relative to the 95% x-y confidence bounds in Figure 1.

DISCUSSION

This study demonstrates the feasibility of metanalytic modeling of the per-subject spatial distribution of brain functional areas using response coordinates ($loci^{x-y-z}$) of group-mean statistical parametric images (SPI[n]) as reported in the brain mapping literature for input data. The FVM modeling strategy makes several operational assumptions, which are preliminarily confirmed by the present study. Specifically, the FVM mathematical model assumes that image averaging (i.e., the creation of SPI[n] for $n > 1$) is well-behaved with respect to response locations. By well-behaved, we mean: 1) that the response location ($locus^{x-y-z}$) in an SPI[n] is effectively identical to the mean of the individual (SPI[1]) response $loci^{x-y-z}$ of n subjects (Table I); and 2) that the variance among m group-mean (SPI[n]) response $loci^{x-y-z}$ is less than the variance among the individual (SPI[1]) response loci, by a factor of $1/n$, as with the standard error of the mean for numeric averaging. That image averaging is well-behaved with respect to response locations was suggested by the data of Fox et al. [1988] and is again confirmed here. This is not, however, a mathematical necessity.

Image averaging is necessarily well-behaved with respect to intensity values, as these are simple numeric averages [Fox et al., 1988]. However, *with respect to location coordinates ($loci^{x-y-z}$)*, averaging of images is not simple numeric averaging. The local maximum in a multisubject SPI (SPI[n]) is determined by the three-

dimensional (3-D) intensity contour created by averaging the images of n subjects. This intensity contour is a function of the actual intensity values in each individual image and the degree of spatial overlap among the response loci. Individual images with more intense activations weight the averaged image more heavily. Images with responses lying far from the group-response centroid weight the averaged response locus little or not at all. Thus, *with respect to location coordinates*, image averaging creates an intensity-weighted mode. If the spatial distribution of per-subject $loci^{x-y-z}$ is unimodal and if the mode is the mean (as with a Gaussian distribution), a group-mean locus will closely approximate the mean of the per-subject loci. If the spatial distribution is less normal, image averaging will not be well-behaved with respect to location coordinates. Thus, the data reported here provide confirmation that image averaging is well-behaved with respect to location coordinates and suggest that the spatial distribution of per-subject $loci^{x-y-z}$ is Gaussian. If this is the case, the thorny problem of individual variability may prove far more manageable than previously supposed.

Functional volumes modeling has a variety of applications, including interpretation of group-mean and single-subject SPIs, hypothesis testing for SPIs, automated labeling of local maxima in SPI, database queries, neuroscience education, and therapeutic/diagnostic targeting, as follows. Group mean SPIs accurately detect mean locations of activation, but provide no estimate of error about that mean. FVM derives an estimate of population variance, from which spatial probabilities can be computed for studies of any sample size. From this, the likelihood that any observed $locus^{x-y-z}$ falls within the confidence bounds of a modeled $locus^{x-y-z}$ can be determined. Clinical applications of this concept would include determining whether an activation in a patient fell outside the normal spatial-probability distribution, i.e., was evidence of disease-induced functional plasticity.

A recurring criticism of voxel-based (SPI) analyses is the lack of formal hypothesis testing. FVMs have potential applications in this domain. The simplest application would be to use FVMs to specify volumes within which activations are predicted, based upon prior task decomposition and literature metanalysis. Activations within the specified volumes confirm the hypothesis; activations outside the specified volume disconfirm the hypothesis. The fraction of the total activation accounted for by the model can be computed and tested with a chi-square statistic for goodness-of-fit of the data to the model. A second application [Friston, 1997] is to adjust the probability

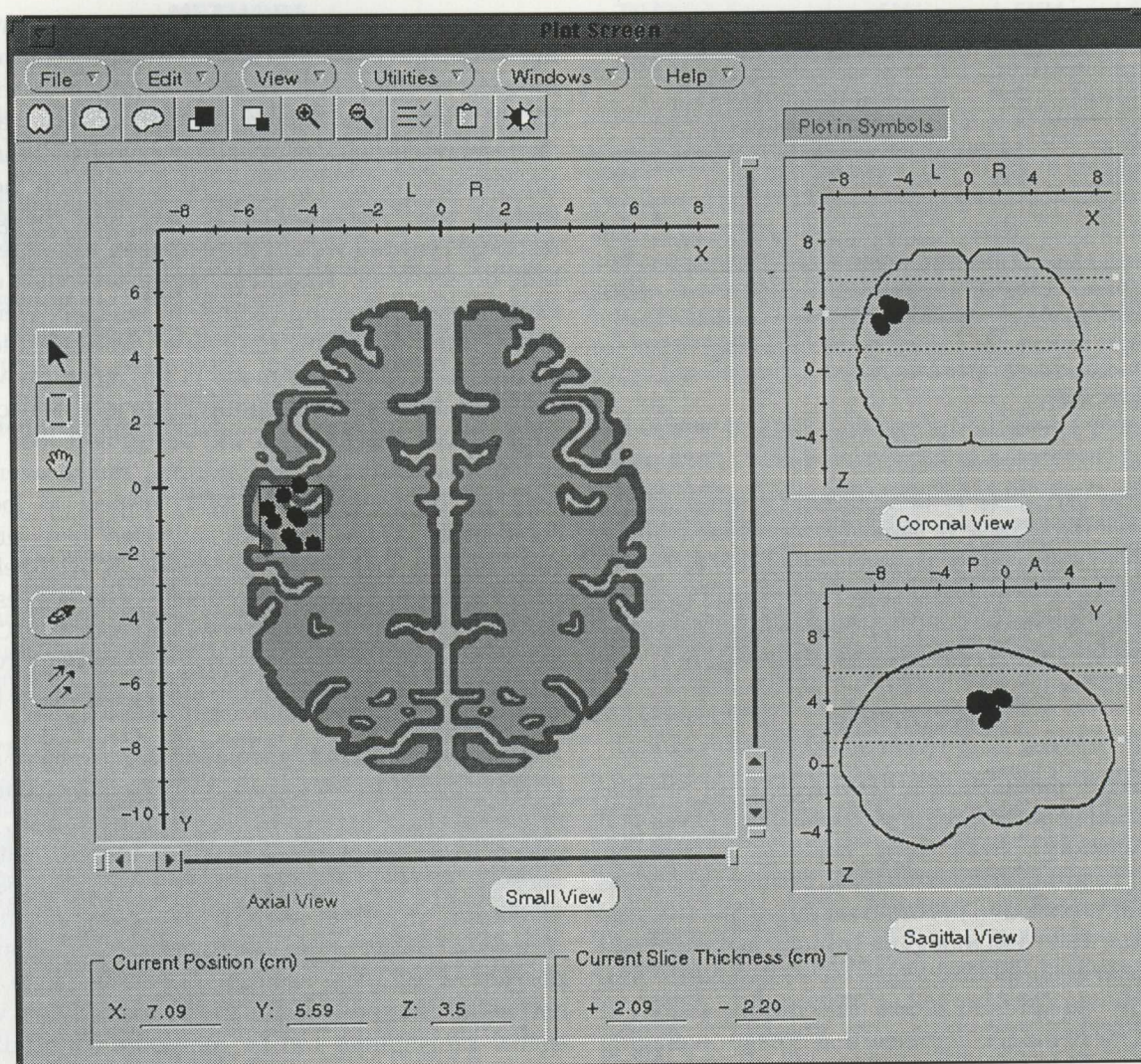


Figure 1.

Spatial distribution of $\text{loci}^{x-y-z}[1]$ for M1-mouth in 10 normal volunteers. Rectangle in the axial view illustrates probability FVM model-computed 95% confidence bounds for per-subject data. Plotted in BrainMap™ [Lancaster et al., 1997].

computation of an SPI to correct only for the number of resolution elements within the FVM and, thereby, to increase the regional sensitivity of SPI analysis.

Identifying local maxima as specific functional areas is time-consuming and subjective. FVM can be used to automate this task, as follows. A response falling within an FVM receives the functional label of the FVM. In fact, the probability that any observed response belongs to a described function area can be described as a probability distribution, with the attached label being that with the greatest probability. As the library of FVM models grows, the percentage of observed activations to which a label could be assigned would also grow.

Neuroscience-education applications are also readily conceived. FVM models synthesize large volumes of data in a manner readily visualized in 2-D and 3-D. Comprehension of models of functional areas performing specific mental operations and of neural systems performing complex tasks is far easier with synthetic, solid models than using the neuroimaging literature itself. As FVM models can be created only for replicated observations and explicitly incorporate the number of studies and subjects establishing each functional area, the student has a clear gauge of the depth of data supporting the model. FVMs can also be used to query the BrainMap™ database, as it supports bounding-box searches of any dimensions. BrainMap™ can be used to

retrieve the original input data used to create the FVM, as reported responses have unique identifiers which can be saved as a list linked to the FVM model. Response identifiers, in turn, can retrieve full citational information, experimental-condition descriptions, imaging methods details, other associated responses, and the like. Thus, the FVM model can serve as a path leading the student progressively deeper into the functional-imaging literature.

Diagnostic and therapeutic targeting is an additional application of FVM. By targeting, we mean aiming a restricted-field measurement or treatment device, such as magnetoencephalography (MEG) or transcranial magnetic stimulation (TMS). The location and confidence bounds for a functional area, expressed by an FVM, can be projected onto a spatially normalized MRI. Applying the inverse spatial normalization transformation, the FVM can be projected onto a nonnormalized MRI and, thereby, onto a person's head. Thus, the location of functional areas can be identified for physiological measurements, such as MEG, or for transcranial interventions, such as TMS.

In summary, FVM is a method for modeling the spatial distribution of brain functional areas. Initial validations of the modeling construct are promising. A variety of applications can be proposed.

REFERENCES

- Andreasen NC, O'Leary DS, Cizadlo T, et al. (1995): Remembering the past: Two facets of episodic memory explored with positron-emission tomography. *Am J Psychiatry* 152:1576-1585.
- Bookheimer SY, Zeffiro TA, Blaxton T, Gaillard W, Theodore W, et al. (1995): Regional cerebral blood flow during object naming and word reading. *Hum Brain Mapping* 3:93-106.
- Buckner RL, Raichle ME, Petersen SE (1995): Dissociation of human prefrontal cortical areas across different speech production tasks and gender groups. *J Neurophysiol* 74:2163-2173.
- Fox PT, Lancaster JL (1995): Neuroscience on the net. *Science* 266:994-996.
- Fox PT, Mintun MA, Reiman EM, Raichle ME (1988): Enhanced detection of focal brain responses using intersubject averaging and distribution analysis of subtracted PET images. *J Cereb Blood Flow Metab* 8:642-653.
- Fox PT, Ingham RJ, Ingham JC, Hirsch TB, Downs JH, Martin C, Jerabek P, Glass T, Lancaster JL (1996): A PET study of the neural systems of stuttering. *Nature* 382:158-162.
- Friston KJ (1997): Testing for anatomically specified regional effects. *Hum Brain Mapping* 5:133-136.
- Lancaster JL, Glass TG, Lankipalli BR, Downs H, Mayberg H, Fox PT (1995): A modality-independent approach to spatial normalization of tomographic images of the human brain. *Hum Brain Mapping* 3:209-223.
- Lancaster JL, Chan E, Mikiten S, Nguyen S, Fox PT (1997): Brain-Map[®] search and view. *Neuroimage* 5:634.
- Paus T, Petrides M, Evans AC, Meyer E (1993): Role of the human anterior cingulate cortex in the control of oculomotor, manual, and speech responses: A positron emission tomography study. *J Neurophysiol* 70:453-469.
- Petersen SE, Fox PT, Posner MI, Mintun M, Raichle ME (1988): Positron emission tomographic studies of the cortical anatomy of single word processing. *Nature* 362:342-345.
- Petrides M, Alivisatos B, Meyer E, Evans AC (1993): Functional activation of the human frontal cortex during the performance of verbal working memory tasks. *Proc Natl Acad Sci USA* 90:878-882.
- Talairach J, Tournoux P (1988): *Co-Planar Stereotaxic Atlas of the Human Brain*, New York: Thieme Medical Publishers, Inc.