

Functional Volumes Modeling: Scaling for Group Size in Averaged Images

Peter T. Fox, Aileen Y. Huang, Lawrence M. Parsons, Jin-Hu Xiong,
Lacey Rainey, Jack L. Lancaster

Research Imaging Center, University of Texas Health Science Center, San Antonio, Texas

Abstract: Functional volumes modeling (FVM) is a statistical construct for metanalytic modeling of the locations of brain functional areas as spatial probability distributions. FV models have a variety of applications, in particular, to serve as spatially explicit predictions of the Talairach-space locations of functional activations, thereby allowing voxel-based analyses to be hypothesis testing rather than hypothesis generating. As image averaging is often applied in the analysis of functional images, an important feature of FVM is that a model can be scaled to accommodate any degree of intersubject image averaging in the data set to which the model is applied. In this report, the group-size scaling properties of FVM were tested. This was done by: (1) scaling a previously constructed FV model of the mouth representation of primary motor cortex (M1-mouth) to accommodate various degrees of averaging (number of subjects per image = $n = 1, 2, 5, 10$), and (2) comparing FVM-predicted spatial probability contours to location-distributions observed in averaged images of varying n composed from randomly sampling a 30-subject validation data set. *Hum. Brain Mapping 8:143–150, 1999.* © 1999 Wiley-Liss, Inc.

Key words: FVM; brain; M1-mouth

INTRODUCTION

Functional volumes modeling (FVM) is a strategy by which spatial probability contours for brain functional areas can be derived from quantitative metaanalysis of a converging body of group-mean, brain-activation studies [Fox et al., 1997]. Within the growing family of quantitative metaanalysis methods appearing in the brain-imaging literature [reviewed in Fox et al., 1998], FVM is the most mathematically formalized and potentially the most useful as an analytic tool. That an FVM entirely derived from published, group-mean input

data can predict the mean location and spatial distribution of responses in a sample of individual subjects has recently been demonstrated, modeling the mouth representation of primary motor-sensory cortex (M1-mouth) [Fox et al., in review]. What remains to be demonstrated is whether FV models can be scaled to accommodate intersubject averaging in the data sets to which they are applied.

A variety of applications are envisioned for FV models. In particular, FVM can serve as spatially explicit predictions in Talairach space of functional activations expected in a brain-mapping experiment. Having formal a priori predictions of expected responses will permit voxel-based analysis of the resulting data to be explicitly hypothesis testing. In addition to increasing scientific rigor, this will increase statistical power by restricting the total volume within which responses are sought, thereby lessening the severity of correction for multiple comparisons.

Contract grant sponsor: NIH; Contract grant numbers: MH/DA52176, DC03689, LM6858; Contract grant sponsor: EJLB Foundation; Contract grant sponsor: Mind Science Foundation.

*Correspondance to: Peter T. Fox, Research Imaging Center, UTHSCSA, 7703 Floyd Curl Drive, San Antonio, TX 78284-6240.

Received for publication 22 June 1999; accepted 12 July 1999

TABLE I. M1-mouth input data*

Citation	Task	n	Left			Right		
			X axis	Y axis	Z axis	X axis	Y axis	Z axis
Petersen et al., 1988	speech	17	-42.8	-15.1	43.6	50.3	-9.9	38.3
Paus et al., 1993	speech	8	-51.0	-11.0	33.0	62.0	-4.0	22.0
Petrides et al., 1993	speech	10	-50.0	-11.0	38.0	44.0	-6.0	36.0
Bookheimer et al., 1995	speech	16	-45.0	-10.0	38.0	42.0	-4.0	44.0
Braun et al., 1997	speech	20	-48.0	-16.0	28.0	44.0	-16.0	28.0
	Total	71						

* M1-mouth location coordinates reported in five, group-mean, brain-activation studies contrasting overt speech to a nonspeaking condition are shown. Values are reported by number of subjects (n), cerebral hemisphere (left, right), and coordinate axis (x = left-right; y = anterior-posterior; z = superior-inferior).

As image averaging is often applied in the analysis of functional images, an important requirement of FVM is that a model can be scaled to accommodate any degree of intersubject averaging in the data set to which the model is applied. The Central Limit Theorem predicts well-behaved reduction of the standard error of the mean as a sample size increases. That is, the variance among a sample of m means, each computed from n individual values, will be lower than the variance in a sample of individual values by a factor of $1/n$. Whereas well-behaved reduction of the standard error of the mean is well established for numeric averaging, it is not established for location coordinates within averaged images. Image averaging is well behaved with respect to intensity values, as these are true numeric averages computed at each voxel [Fox et al., 1988]. However, with respect to the locations of local maxima, the effects of image averaging are not so straightforward. A local-maximum in a functional image is determined by a 3D intensity contour. In an multisubject image, the intensity contour of a local maximum is a function of the actual intensity values in each individual image and by 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. As a result, there is no mathematical guarantee that the decrease in variance among response loci in averaged images will be the same as that obtained by averaging locations coordinates from individual-subject images. The purpose of this report was to empirically validate: (1) the model assumptions underlying FVM group-size scaling, and (2) FVM

predictive accuracy for group-mean data across a range of group sizes.

METHODS

Literature metanalysis of M1-mouth and functional volume modeling

The mean location and location variability of M1-mouth were estimated from the coordinate-referenced, group-mean, brain-activation literature. Candidate studies were limited to those: (1) on normal subjects, (2) using overt oral tasks. Eight such studies were identified [Petersen et al., 1988; Paus et al., 1993; Petrides et al., 1993; Andreasen et al., 1995; Bookheimer et al., 1995; Fox et al., 1996; Braun et al., 1997; Murphy et al., 1997]. Of these, Fox et al. [1996] was discarded because the subject sample partially overlapped the validation sample (below). Andreasen et al. [1995] and Murphy et al. [1997] were discarded because they failed to differentiate between M1-mouth and ventral pre-motor cortex (i.e., BA 6/44, or Broca's area), likely due to excessive smoothing. Five studies were retained for the metanalysis. They ranged in group-size (n) from 8 to 20 and totaled 71 subjects (Table I). Using these input data and the FVM construct [Fox et al., 1997], population (i.e., for single-subject, nonaveraged images) location-probability profiles were computed for each cardinal axis (x, y, z) and each cerebral hemisphere. These population profiles were then scaled down to the group sizes predicted for statistical parametric images (SPI), which averaged images from 2, 5, or 10 individual subjects (here termed SPI[2], SPI[5], SPI[10]).

Validation data

Three cohorts of 10 subjects each participated in one of three speech-motor protocols at the Research Imaging Center (UTHSCSA). All three protocols acquired three scans per condition, which allowed intrasubject image averaging and per-subject response localization. High-resolution, 3-D (1 mm³ voxels), T1-weighted anatomical MRI was obtained in each subject, for purposes of anatomical normalization. PET-MRI registration and spatial normalization were performed using the Lancaster et al. [1995] algorithm, as implemented in the SN[™] software (Research Imaging Center, UTHSCSA, San Antonio, Texas; ric.uthscsa.edu/projects/). This algorithm employs a nine-parameter, affine transformation to normalize images relative to the atlas of Talairach and Tournoux [1988].

Statistical parametric images (SPIs) were created using the Fox et al. [1988] algorithm, as implemented in the MIPS[™] software (Research Imaging Center, UTHSCSA, San Antonio, Texas; ric.uthscsa.edu/projects/). This algorithm uses the pooled variance of all brain voxels as the reference for computing significance, rather than computing the variance at each voxel. This procedure allows formation of single-subject SPI, even without intrasubject averaging and is more reliable for small samples than the voxel-wise variance methods of Friston et al. [1991] and others [Strother, 1997]. Fifty-four SPIs were created: 30 single-subject SPIs (SPI[1]), 15 SPI[2], 6 SPI[5] and 3 SPI[10]. For each n ($n = 1, 2, 5, 10$), data sets were mutually exclusive. That is, sampling was without replacement.

The M1-mouth locations (left and right) were determined in each SPI, as follows. The 30-subject SPI was scanned with a local-maximum search algorithm [Mintun et al., 1989], to determine the locations of the M1-mouth areas (left and right) in the image with the highest possible signal to noise (i.e., the 30-subject SPI). In this 30-subject image, several functional regions comprising the speech-motor system were readily identified, including: the left and right M1-mouth representations (BA4); supplementary motor area (medial BA6); the left ventral premotor region (BA6/44); and the left insula. The two M1-mouth locations were used to limit the search domain within the 30, single-subject SPIs. Potential left and right M1-mouth loci were automatically detected as the most intense (highest positive z score) local maximum within a 3-cm radius of the grand-average M1-mouth locus. The large detection radius avoided artificially limiting the variance observed among loci.

For the 30 SPI[1], the most intense response within the search range was not appropriate in location for

M1-mouth in nineteen instances; 10 in the left hemisphere; 9 in the right hemisphere. For example, the supplementary motor area was captured by the search algorithm in 1 case. Ventral premotor cortex (in humans, “Broca’s area”) was captured by the search algorithm in 14 cases. Superior temporal (in humans, “Wernicke’s area”) was captured in two cases and superior premotor was captured in two cases. When the most intense response was inappropriately located, the next most intense response was located appropriately for M1-mouth in 11 instances (8 left; 3 right). In eight instances (2 left; 6 right), no M1-mouth response could be identified. In the 15 SPI[2], a M1-mouth response could not be identified in two instances: once on the left; once on the right. In the six SPI[5], an M1-mouth response could not be identified in the left hemisphere of one image. In the SPI[10], left and right M1-mouth responses were identified in all three images.

Statistical analysis

M1-mouth response loci resulting from the above-described imaging and image processing steps were tested to determine: (1) whether image averaging was well-behaved with respect to response means and variances, (2) whether the spatial distributions of image-averaged data differed significantly from Gaussian, (3) whether the FVM for M1-mouth cortex correctly predicted the spatial distribution (mean and population percentiles) of SPI[n].

Image averaging

The reduction of variance among M1-mouth response loci achieved by image averaging was compared to that achieved by coordinate averaging, using the same data sets as input for both averaging procedures. The numeric means of groups of 1, 2, 5, and 10 SPI[1] were compared to the loci from SPI[1], SPI[2], SPI[5] and SPI[10]. Comparisons were graphical and statistical. A Hotelling’s T₂ test was used to test the null hypothesis that means achieved by averaging images (SPI[1]) did not differ from those achieved by averaging coordinates (SPI[n]).

Normality

Fox et al. (in review) demonstrated that M1-mouth location-distribution profiles of individual-subject data for each axis, in each hemisphere is normal and can be pooled to generate a larger sample size. Thus to assess the normality of the spatial distribution of response locations within averaged images, responses for each n were analyzed by pooling the data across axes (x, y, z)

TABLE II. Effects of image averaging on response locations*

n	Left				Right			
	m	X	Y	Z	m	X	Y	Z
1	28	-45 ± 4.4	-10 ± 6.2	38 ± 5.0	24	50 ± 3.6	-7 ± 6.0	37 ± 5.5
2	14	-45 ± 2.7	-10 ± 4.9	39 ± 3.6	14	50 ± 4.1	-8 ± 4.0	36 ± 5.4
5	5	-46 ± 2.2	-10 ± 2.2	40 ± 1.4	6	50 ± 2.5	-9 ± 3.2	37 ± 3.6
10	3	-45 ± 1.2	-8 ± 1.8	40 ± 1.8	3	50 ± 3.1	-9 ± 0.7	37 ± 2.7
30	1	-46	-8	40	1	52	-8	38

* M1-mouth location coordinates ± one standard deviation for single-subject data (n = 1) and for 3 group sizes (n = 2, 5, 10) are shown. m is the number of images from which the mean and standard deviation were computed.

and hemisphere (left and right). A pooled histogram for each *n* was created by normalization of each of the six, individual-axis data sets to a zero-mean and unit standard deviation histogram.

Fit of data to model

The hypothesis that FV models derived from the group-mean SPI literature will predict the spatial distributions of SPI[n] for any *n* was assessed graphically and by descriptive statistical parameters. Graphical assessment used the BrainMap® database and user-interface tools, as follows. M1-mouth response loci for each level of averaging (n = 1, 2, 5, and 10) were entered into the BrainMap® database. Corre-

sponding FVM for each level of averaging (n = 1, 2, 5, and 10) were created using the M1-mouth literature (Table I). Response distributions, relative to the bounds of the FVM, were visualized. As a descriptive statistic, the percentage of responses lying within the 95th confidence bound along each axis was determined for each *n*.

RESULTS

Image-averaging

The effects of image averaging were assessed by comparing response loci from 15 SPI[2], 6 SPI[5] and 3 SPI[10] with numerical averages of the same data sets

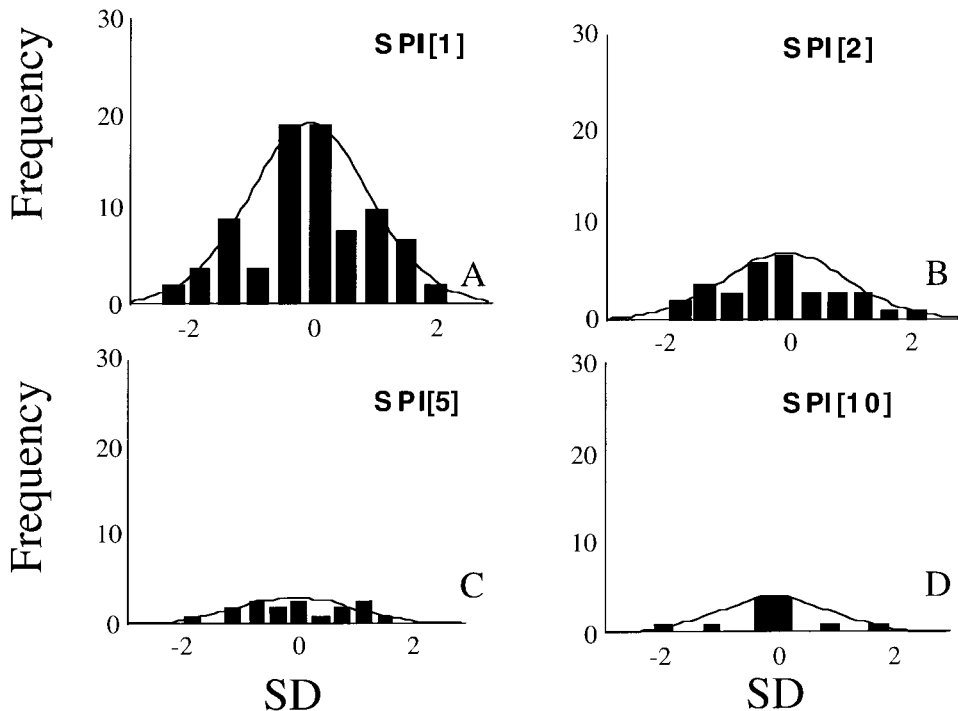
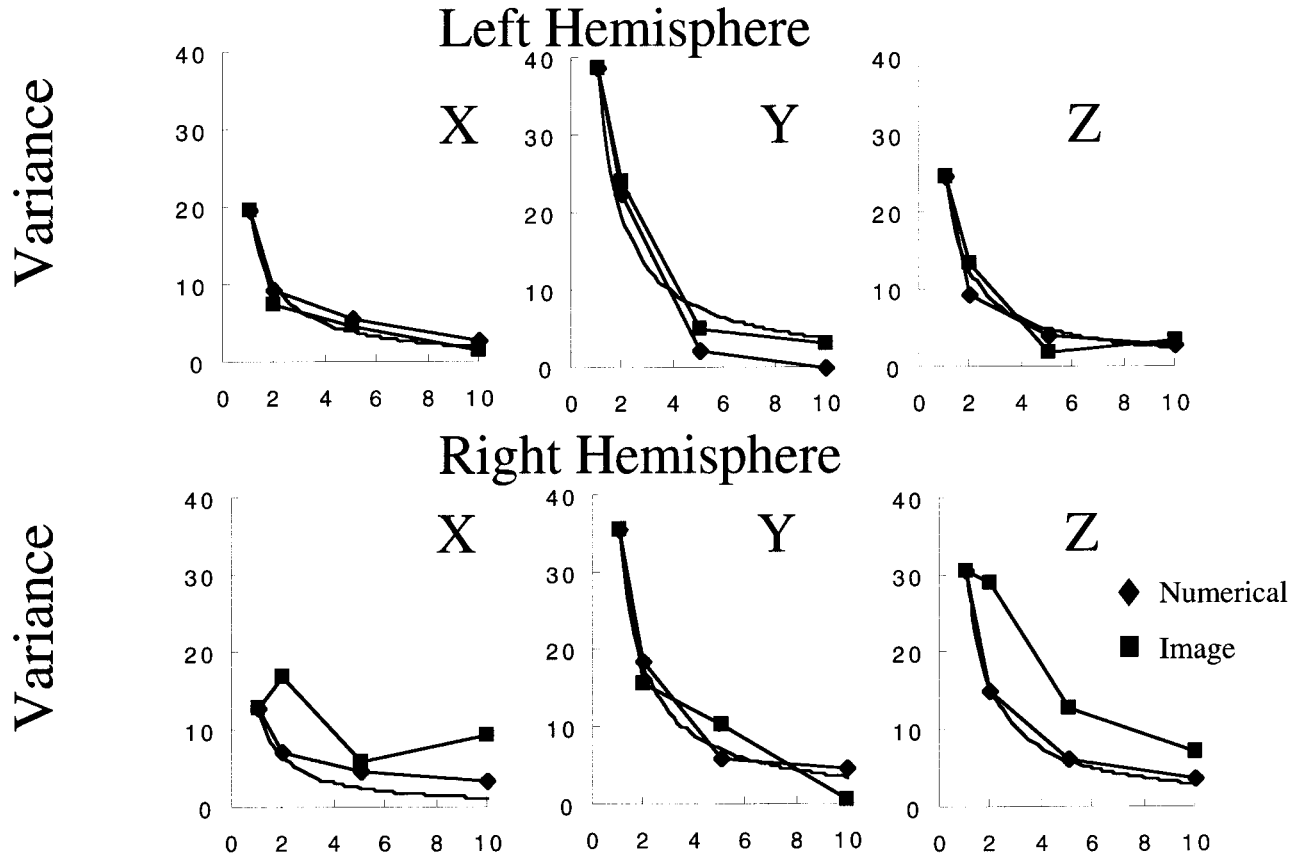


Figure 1. Reduction in variance with image averaging is shown. The decline in variance, relative to per-subject (SPI[1]) variance, is shown as a solid line. The decline in variance, relative to SPI[1] with simple, numeric averaging of per-subject coordinates, is shown as a dotted line. The decline in variance, relative to SPI[1], achieved by image averaging, is shown as a dashed line.



Number of Individual Images Per Averaged Image

Figure 2.

Pooled histograms of the spatial distributions of the image-averaged M1-mouth loci: (a) 30 SPI[1], b) 15 SPI[2], (b) 6 SPI[5], (c) 3 SPI[10]. Overall, the image averaged responses form spatial distributions that are not significantly different from a Gaussian.

and with the inverse function ($1/n$). Means and variances were computed across groups. Location means were not effected by image averaging or numeric averaging (Table II; Fig. 1). In no instance did location between nonaveraged (SPI[1]) and averaged data ($P > 0.70$; Hotteling's T-test) differ. Mean location did not statistically differ between numeric averaging and image averaging ($P > 0.3$ for $n = 2, 5, \text{ and } 10$).

Variability was reduced relative to SPI[1], as a function of $1/n$ for both image averaging (SPI[n]) and for numeric averaging, also as predicted (Table II, Fig. 1). Well-behaved image averaging suggests that the spatial distribution (across subjects) of M1-mouth is unimodal and that the mean location is at or near the mode. Although these are properties of a Gaussian distribution, well-behaved image averaging is not per se a demonstration of normality.

Profile normality

The M1-mouth location-distribution profiles observed for the different group sizes appeared normal (i.e., Gaussian), within the limits of sample size (Fig. 2, Table III). This observation was statistically tested by computing the skew and kurtosis of the four profiles ($n = 1, 2, 5, 10$). All values were near zero, as expected for a Gaussian distribution. None differed significantly from zero (gamma-one and gamma-two statistics) [Zar, 1996]. The gamma and K2 D'Agostino-Pearson statistics for the pooled distributions also were not significantly different from Gaussian ($P > 0.5$). These observations support the FVM modeling assumption that location-distribution profiles are normal. Whereas averaging would be expected to increase the normality of the

TABLE III. Profile normality*

n	Left			Right				
	m	X $\gamma_1; \gamma_2$	Y $\gamma_1; \gamma_2$	Z $\gamma_1; \gamma_2$	m	X $\gamma_1; \gamma_2$	Y $\gamma_1; \gamma_2$	Z $\gamma_1; \gamma_2$
1	28	-0.1; 0.3	0.3; -1.0	-0.5; -0.2	24	0.3; -0.2	0.2; -0.6	0.0; 1.4
2	14	0.2; -0.8	0.3; -0.5	0.9; 0.3	14	-0.8; 0.6	0.5; -0.4	-0.5; -0.5
5	5	-0.1; 1.8	1.7; 0.7	0.7; 1.0	6	-1.2; 0.3	1.7; 3.0	0.2; 2.0

* Profile normality is an assumption of FVM modeling. The normality of the spatial distributions of the M1-mouth locations in each spatial axis (X, Y, Z) for single-subject data (n = 1) and for 2° of averaging (n = 2, 5) are shown. Skew (γ_1 statistic) and kurtosis (γ_2 statistic) were near zero, the expected values for a normal distribution, for all three axes in both hemispheres.

data sets, the small sample sizes at larger group size likely prevented this effect from being appreciated.

Validation of M1-mouth FVMs

Literature-derived FV models of the spatial distribution of M1-mouth loci were validated by comparisons to the data sets just described, as follows. Spatial probability distributions for SPI[1], SPI[2], SPI[5], and SPI[10] were computed (from the literature-derived, group-mean input data, above) and expressed as 95% confidence limits (z = 2.0) for each axis (x, y, z). The fraction of subjects falling within these bounds was generally high (Table IV). For the higher degrees of averaging (e.g., SPI[10]), small sample sizes made the computed fractions more variable. Nevertheless, there appeared to be no systematic or progressive errors in the spatial probability computations.

DISCUSSION

The analyses here presented confirmed the two principal modeling assumptions and the predictive power of functional volumes modeling [Fox et al., 1997], as regards its application to group-mean data sets. First, the mean location of the M1-mouth response

was not altered by averaging. Second, the variability of response locations among sets of averaged images declined with the number of individual-subject images per average, as predicted. Third, spatial probability contours computed using these two assumptions were validated by comparison to data sets not utilized in their creation.

The validations performed here were limited by the relatively small number of images that were created for each averaging level. For example, only three 10-subject images (SPI[10]) were created. This limitation was due conjointly to the size of the validation data set (30 individual subjects) and to the decision to sample without replacement when creating the subject groups for averaging. Despite this limitation, the data provide strong support of the two assumptions of well-behaved image averaging and of the predictive power of FVM when applied to averaged images.

Theory and scope

Functional volumes modeling takes advantage of the normality of the spatial distribution of functional activations within a Cartesian space. Brain functional areas are described as probability distributions about a mean address, with no need to resort to gross anatomi-

TABLE IV. Fit of FV models to validation data*

n	Left			Right		
	X	Y	Z	X	Y	Z
1	100% (28/28)	89% (25/28)	86% (24/28)	100% (24/24)	92% (22/24)	92% (22/24)
2	100% (14/14)	79% (11/14)	93% (13/14)	79% (11/14)	93% (13/14)	86% (12/14)
5	100% (5/5)	80% (4/5)	80% (4/5)	50% (3/6)	83% (5/6)	67% (4/6)
10	100% (3/3)	33% (1/3)	67% (2/3)	33% (1/3)	100% (3/3)	67% (2/3)

* Percentage of M1-mouth responses falling within the FVM-predicted bounds are shown for each cardinal axis (X, Y, Z) in each cerebral hemisphere (right and left) for single-subject data (n = 1) and for 4 degrees of averaging (n = 2, 5, 10, 15).

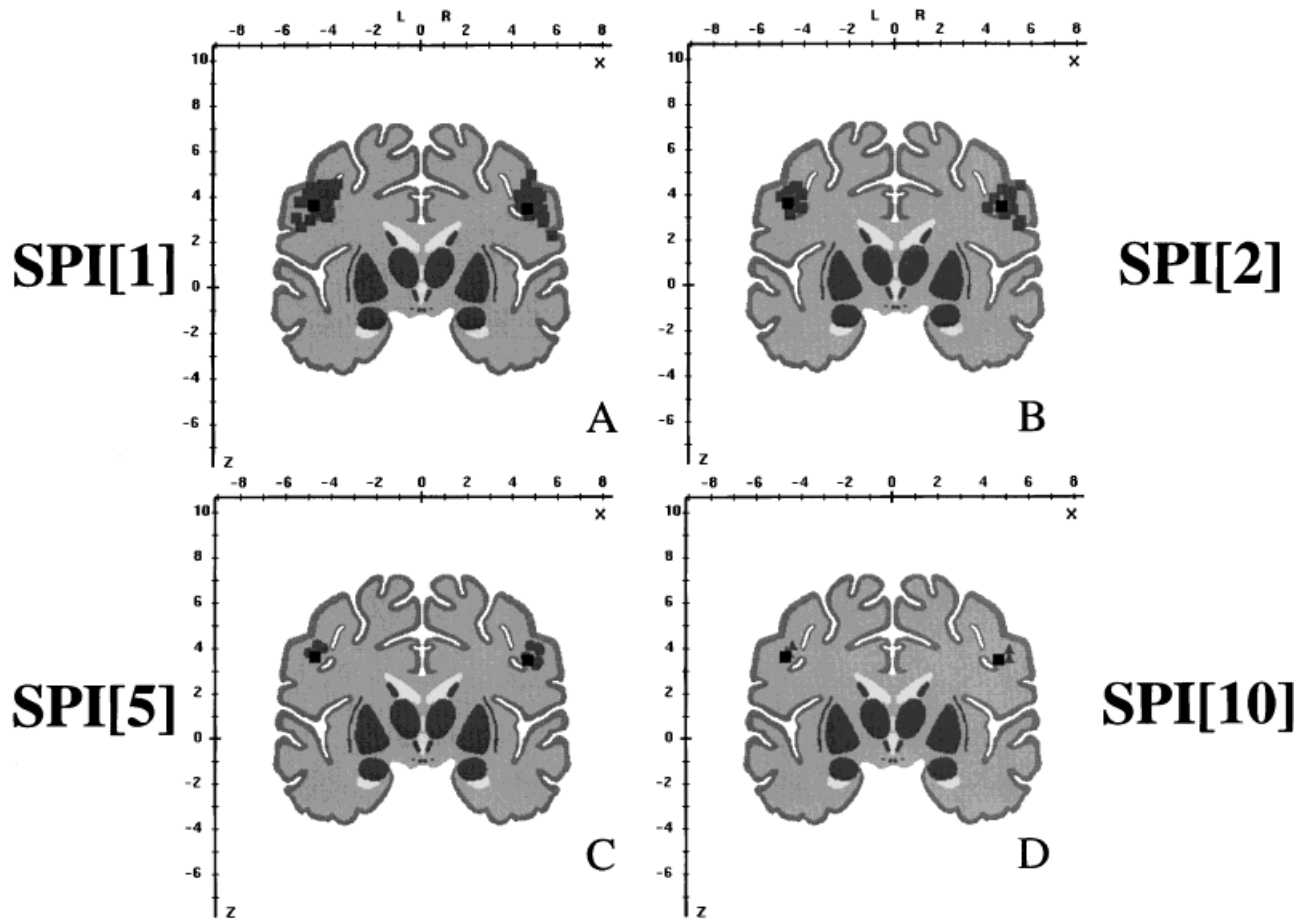


Figure 3.

Goodness-of-fit of validation data to the meta-analytically derived, spatial probability contours for MI mouth are illustration. **A** illustrates 28 individual-subject response loci (SPI[1]). **B** illustrates 20 two-subject response loci (SPI[2]). **C** illustrates 8 five-subject (SPI[5]) response loci. **D** illustrates 3 ten-subject (SPI[10]) re-

sponse loci. In A–D, boxes illustrate 95% confidence bounds appropriate to the number of subjects per image (n). Plotted in BrainMap [Fox and Lancaster, 1996]. All values are in mm referable to Talairach and Touroux [1988].

cal landmarks. In FVM, the estimate of per-subject variance is derived from interstudy variance, correcting for sample size and interlaboratory methodological error. Thus group-mean studies can be used to estimate per-subject variance. This requires, however, that each area to be modeled be reported in a sufficiently large number of independent studies to allow a good estimate of variance. As shown in the present study, it does not require that these studies be exact duplications or even that they use the same activation paradigm. It merely requires that whatever activation paradigm is employed be sufficient to recruit the area to be modeled. The corpus of group-mean studies reporting function locations in standardized coordinates is now quite large and is growing at a rate of approximately one per day [Fox, 1997]. Thus a large

number of functional areas should be readily modeled using the available data.

Utility

Several motivations for performing quantitative meta-analysis of functional areas can be identified. First and foremost, meta-analytic models serve as accurate, concise, intuitive formulations of accumulated knowledge, as is illustrated here. In addition, models can serve as tools for automated image analysis, image interpretation, and data retrieval. As an example of image interpretation, a location-probability model (either structural or functional) can be used to assign a most-likely name and an likelihood value to a feature within a brain image (e.g., an activated location).

Structure-location names and probabilities are already being used for this purpose [Lancaster et al., 1997]. Functional spatial-probability contours could be used in a similar manner. Functional location-probability contours can be used as a regions-of-interest, to specify locations within an activation image hypothesized to be engaged by a task. By this strategy, analysis would ask whether or not the hypothesized areas (defined by FVM as location-probability boundaries) were activated under a specific set of conditions, thereby directly addressing the recurring criticism that voxel-based image analyses are intrinsically hypothesis generating rather than hypothesis testing [Ford, 1986; Friston et al., 1991, Worsley et al., 1992]. This strategy would also increase statistical power, by reducing the analyzed volumes, thereby reducing the severity of the correction for multiple comparisons [Friston et al., 1997]. Precise description of location-probability distributions for the normal-subject population provides a powerful tool for identifying aberrant organizations, such as are likely occur with developmental and acquired brain lesions. Spatial probability models can also be used to guide experimental or therapeutic interventions. For example, Paus and colleagues [1997] have used probabilistic estimates of mean location to guide delivery of transcranial magnetic stimulation. Finally, retrieval of studies activating a specific brain location from a database, such as BrainMap® [Fox and Lancaster, 1996], can be readily and powerfully performed by means of location-probability bounds.

REFERENCES

- Andreasen NC, O'Leary DS, Cizadlo T, Arndt S, Rezai K, Watkins GL, Ponto LLB, Hichwa RD. 1995. Remembering the past: two facets of episodic memory explored with positron emission tomography. *Am J Psych* 152:1576–1585.
- Bookheimer SY, Zeffiro TA, Blaxton T, Gaillard W, Theodore W. 1995. Regional cerebral blood flow during object naming and word reading. *Hum Brain Mapp* 3:93–106.
- Braun AR, Varga M, Stager S, Schulz G, Selbie S, Maisog JM, Carson RE, Ludlow CL. 1997. Altered patterns of cerebral activity during speech and language production in developmental stuttering. *Brain* 120:761–784.
- 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. 1996. Un atlas du cerveau sur internet. *La Recherche* 289:49–51.
- Fox PT, Lancaster JL, Parsons LM, Xiong JH, Zamarripa F. 1997. Functional volumes modeling: theory and preliminary assessment. *Hum Brain Mapp* 5:306–311.
- Fox PT, Huang AY, Parsons LM, Xiong JH, Zamarripa F, Rainey L, Lancaster JL. 1999. Location-probability profiles for the mouth region of human primary motor-sensory cortex. *J Neurosci* (in press).
- 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.
- Fox PT, Mintun MA. 1989. Noninvasive functional brain mapping by change-distribution analysis of average PET images of H₂ ¹⁵O tissue activity. *J Nucl Med* 30:141–149.
- Fox PT, Parsons LM, Lancaster JL. 1998. Beyond the single study: function/location metanalysis in cognitive neuroimaging. *Current Opinion Neurobiol* 8:178–187.
- Friston KJ, Frith CD, Liddle PF, Frackowiak RSJ. 1991. Comparing functional (PET) images: the assessment of significant change. *J Cereb Blood Flow Metab* 11:690–699.
- Friston KJ. 1997. Testing for anatomically specified regional effects. *Hum Brain Mapp* 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 Mapp* 3:209–223.
- Mintun MA, Fox PT, Raichle ME. 1989. A highly accurate method of localizing regions of neuronal activation in the human brain with positron emission tomography. *J Cereb Blood Flow Metab* 9:96–103.
- Murphy K, Corfield DR, Guz A, Fink GR, Wise RJS, Harrison J, Adams L. 1997. Cerebral areas associated with motor control of speech in humans. *J Appl Physiol* 83:1438–1447.
- 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* 90:878–882.
- Ramsey NE, Kirkby BS, Van Gelderen P, Berman KF, Duyn JH, Frank JA, Mattay VS, Van Horn JD, Esposito G, Moonen CTW, Weinberger DR. 1996. Functional mapping of human sensorimotor cortex with 3D BOLD correlates highly with H₂ ¹⁵O PET rCBF. *J Cereb Blood Flow Metab* 16:755–764.
- Strother SC, Lange N, Anderson JR, Schaper KA, Rehm K, Hansen LK, Rottenberg DA. 1997. Activation pattern reproducibility: measuring the effects of group size and data analysis models. *Hum Brain Mapp* 5:312–316.
- Talairach J, Tournoux P. 1988. Co-planar Stereotaxic Atlas of the Human Brain. New York: Thieme.
- Worsley KJ, Evans AC, Marrett S, Neelin P. 1992. Determining the number of statistically significant areas if activation in the subtracted studies from PET. *J Cereb Blood Flow Metab* 12:900–918.
- Zar JH. 1996. *Biostatistical Analysis*, 3rd ed. Upper Saddle River, NJ: Prentice-Hall.