INTRODUCTION

Coordinate-Based Voxel-Wise Meta-Analysis: Dividends Of Spatial Normalization. Report of a Virtual Workshop

Peter T. Fox, * **Angela R. Laird, and Jack L. Lancaster**

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

Spatial normalization transforms a brain image from its natural form ("native space") into a standardized form defined by a reference brain [Fox, 1995a]. The original motivation for introducing this technique was to allow the brain locations of task-induced functional activations to be reported in a "precise and unambiguous" manner, thereby "facilitating direct comparison of experimental results from different laboratories" [Fox et al., 1985]. The prospect of clear communication as a "dividend" from a community commitment to spatial normalization, however, proved largely unconvincing to the still-nascent brain mapping community of the middle 1980s. Improvement in the signal-to-noise ratio of functional brain maps that could be achieved by intersubject image averaging in standardized space [Fox et al., 1988; Friston et al., 1991] proved to be a very salient motivation, leading to widespread adoption of this data analysis standard. We estimate the human functional brain mapping (HFBM) literature reporting brain activations as x-y-z coordinates in standardized space to be no less than 2,500 articles (~10,000 experiments) with ~500 new articles (2,000 experiments) published per year (Fig. 1). Fortunately, regardless of the motivation for adoption of this standard, the widespread use of spatial standardization makes the HFBM literature fertile ground for quantitative meta-analysis methods based on spatial concordance [Fox and Lancaster, 1996a,b; Fox et al., 1998]. In reference to the title of this article, voxel-based, function-location meta-analysis can be considered a dividend that the HFBM community is now receiving from its long-term investment in the development and promulgation of community standards for data-analysis and, in particular, spatial normalization.

Meta-analysis is defined most generally as the post-hoc combination of results from independently performed studies to estimate better a parameter of interest. The original and by far the most prevalent form of meta-analysis pools studies with nonsignificant effects to test for significance in the collective, using the increase in n to increase statistical power [Pearson, 1904]. Effect-size meta-analyses have come under criticism for a variety of misuses, but are growing steadily in power and acceptance [Fox et al., 1998]. In the HFBM community, fundamentally new forms of meta-analysis are emerging, in which statistically significant effects are pooled and contrasted to estimate better such parameters as the spatial location, spatial distribution, activation likelihood, co-occurrence patterns, and underlying cognitive operations for specific categories of task. In the first published meta-analysis in cognitive neuroimaging, coordinates from three prior reports were tabulated and plotted to guide interpretation of results in a primary (non-meta-analytic) study [Frith et al., 1991]. Shortly thereafter, "stand-alone" HFBM meta-analyses began to appear in the literature [Buckner and Petersen, 1996; Fox, 1995b; Paus, 1996; Picard and Strick, 1996; Tulving et al., 1994]. To date, more than 50 meta-analyses of coordinate-based HFBM studies have appeared in the peer-reviewed literature. Although most of these meta-analyses are semiquantitative and statistically informal, this is changing.

The trend toward quantitative, statistically formal HFBM meta-analysis began with Paus [1996], who computed and interpreted means and standard deviations of the x-y-z addresses in a review of studies of the frontal eye fields. Fox et al. [1997, 2001] extended this initiative by correcting raw estimates of spatial location and variance for sample size to create scalable models of location probabilities (functional volumes models; FVM) and suggesting uses of such models
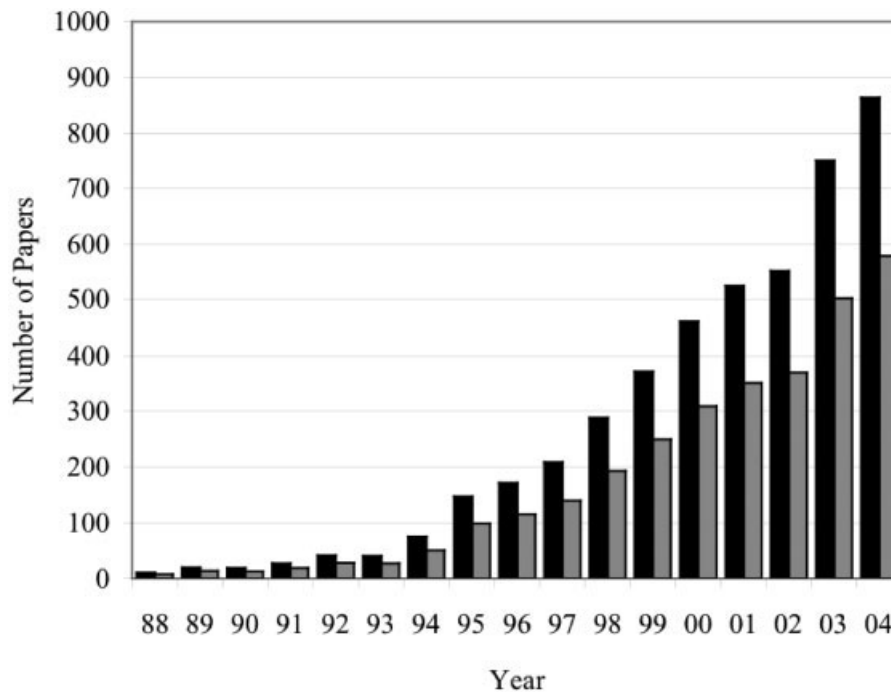
Figure 1.

Human functional brain mapping literature. Annual publication rates for all human brain mapping studies (dark bars) and for the subset suitable for quantitative meta-analysis (light bars) are shown. To quantify the total literature, Medline searches were carried out for each year using the criteria: Brain Mapping AND Human AND (fMRI or PET) AND Year (dark bars). To determine the average percentage of brain-mapping studies suitable for quantitative meta-analysis, 293 articles from six paradigm classes were reviewed: Stroop, n-back, picture naming, word generation, mental rotation, and pain tasks. Percentages ranged from 33% (pain) to 95% (picture naming) with an average value of 67%. This fraction (67%) was used to correct the total literature volume values (dark bars) downward to the meta-analyzable subset (light bars).

for data analysis. The FVM approach is limited, however, by the need for investigator identification of concordant sets of coordinates; that is, x-y-z coordinates must be identified as belonging to a specific functional region before inclusion in a meta-analysis. A method for computing whole-brain activation probability maps by meta-analysis was introduced by Xiong et al. [2000], where the likelihood of activation is determined as the percentage of studies in which individual voxels exceeded a significance threshold. The principal limitation of the voxel-wise "penetrance" method of Xiong et al. [2000] is the use of raw statistical parametric images (SPIs) as input data, rather than published coordinates. Nielsen and Hansen [2002] used the BrainMap database [Fox and Lancaster, 1996a,b, 2002; Fox et al., 2005; Laird et al., 2005a] to create meta-analytic models of the spatial distributions of activation clusters. This approach is fully automatic, looking for spatial clusters of activation foci and seeking outliers in the meta-data descriptors of clustered foci. This data-mining approach, however, is conceptually general rather than being developed specifically for HFBM experiments and consequently does not lend itself readily to testing specific neuroscientific hypotheses. None of these approaches fully tap the potential implicit in the HFBM reporting standards.

Coordinate-based, voxel-wise meta-analysis (CVM) overcomes many of these shortcomings [Chein et al., 2002; Turkeltaub et al., 2002; Wager et al., 2003a,b]. Input data are activation-location coordinates from conceptually related studies in the HFBM literature, e.g., all Stroop tasks. For each experiment included in the meta-analysis, the entire set of reported coordinates are placed within a 3D image matrix and blurred with a Gaussian filter approximating intersubject anatomical variability, thereby reconstructing the essence of the original SPI from which the coordinates were extracted. These "pseudo-SPIs" are concatenated to compute a voxel-wise estimation of activation likelihood for a family or category of tasks. As with FVM, a great advantage of CVM is that the tables of coordinates routinely reported by HFBM studies are its input data; raw data are not required. CVM, however, does not require user selection of comparable coordinates for modeling; rather, once a set of experiments is selected for meta-analysis the entire set of reported coordinates is used, thereby increasing the automation and objectivity of the analysis. Another advantage of CVM is that the output is a voxel-wise pseudo-SPI that can be compared directly with other CVM images (e.g., to contrast activation patterns for different categories of task) and with

SPIs (e.g., as an explicit confirmation of a CVM-based hypothesis). These comments are not intended to indicate that CVM is a fully mature method. For example, activation likelihood estimation (ALE), the most sophisticated, well-defended, and well-validated of the three original CVM methods, originally made no correction for multiple comparisons, nor could two ALE images be statistically compared [Turkeltaub et al., 2002]. ALE includes no means for weighting the computation by the number of subjects in each included study, although this has a strong effect on the reliability of observed activations [Lancaster et al., 2005]. Despite these limitations, the CVM approach was judged sufficiently promising by a diverse group of "brain-imaging experts" to motivate the present Meta-Analysis Special Issue of *Human Brain Mapping*.

The logistics underlying the development of this special issue deserve a brief account, as it proved to be a novel exercise in electronic collaboration and education: a "virtual workshop." A special issue devoted to meta-analysis was conceived by the authors (P.T.F., A.R.L., and J.L.L.), motivated by mutual, long-standing interests in spatial normalization, meta-analysis, and community standards for data analysis and data sharing. Because we judged ALE to be the present state-of-the-art for HFBM meta-analysis and wished to have the contributions meet a uniformly high standard, we envisioned that all included articles would use ALE, ALE equivalents, or ALE extensions. As we have argued previously that "meta-analysis should be performed by experts in the subject matter at hand rather than by statisticians…" [Fox et al., 1998], we proposed to solicit the participation of investigators with universally acknowledged expertise in a specific segment of the brain imaging literature, ideally, the author who introduced a specific experimental paradigm into functional brain imaging, regardless of prior experience with meta-analysis. To make the project less daunting to prospective contributors and to ensure a uniformly high quality of analysis, we further proposed that all meta-analyses would be carried out in our laboratory (by A.R.L.), with the literature review, experiment selection, and results interpretation carried out by the contributing authors. To maximize interaction among contributors, peer review would be provided by other contributors. This proposal was reviewed and endorsed by the Associate Editors of *Human Brain Mapping*, several of whom are contributing authors. This plan was implemented as proposed (with a few notable exceptions); the present special issue of the *Journal* is the result.

The CVM approach proved surprisingly versatile, powerful and user friendly to previously meta-analysis-naive imaging investigators and even to investigators who do not use imaging as their primary investigative tool. The diversity of the topics addressed and the design of the studies, even within the constraint that all studies used a CVM method, was impressive. Not unexpectedly, several of the studies addressed paradigms used widely in functional imaging as applied in normal controls, including the n-back task [Owen et al., 2005], the Stroop task [Derrfuss et al., 2005; Laird et al., 2005b], the Wisconsin Card-Sorting Task [Buchsbaum et al.,

2005], painful stimulation [Farrell et al., 2005], and saccade generation [Grosbras et al., 2005]. In each instance, the literature proved sufficiently rich that variations of the paradigm were compared and effects not identified readily in the primary literature were detected. For example, Laird et al. [2005b] identified somatotopy within the anterior cingulate gyrus based on the response required (spoken vs. button press) in the Stroop task. In a highly novel variation of this strategy, Price et al. [2005] contrasted picture-naming studies based on the baseline condition, using a conjunction analysis to isolate and differentiate the mental operations carried out during control conditions (high-level vs. low-level); the results go well beyond the original studies and even more importantly, have already been confirmed by a prospective functional magnetic resonance imaging (fMRI) study carried out exclusively for this purpose [Price et al., 2005]. Petacchi et al. [2005] pooled auditory control conditions from a wide variety of paradigms to test the hypothesis that the cerebellum plays a role in perceptual processing independently of movement planning or execution. Two studies contrasted normal subjects based on native language [Bolger et al., 2005; Tan et al., 2005]. Two studies contrasted patient populations to normal controls [Brown et al., 2005; Glahn et al., 2005]. In each case, meta-analysis confirmed cross-study concordance and pointed toward emerging effects and new hypotheses.

Accelerated evolution of meta-analysis methods was a less anticipated but very welcome outcome of our virtual workshop. As contributing authors became familiar with CVM methods, they made requests for additional functions to enhance their analyses. Price et al. [2005], for example, needed to assess the statistical significance of the difference between two ALE images to interpret their conjunctional meta-analysis. In response, Laird et al. [2005c] implemented and validated a permutation test for this purpose; once completed, this function was used by many of the contributing authors. Similarly, several contributors requested a correction for multiple comparisons and a volume-of-interest (VOI)-analysis tool, to determine which studies contributed to each likelihood focus; both of these were implemented [Laird et al., 2005c] and used by contributing authors. Neumann et al. [2005] contributed a highly original approach to network analysis, which already has been explored further and extended by Lancaster et al. [2005]. Overall, the impression is that CVM is strong foundation upon which to build and is amenable to many extensions and improvements.

Large-scale data sharing is yet another dividend of our virtual workshop. Collectively, the contributions to this issue have harvested more than 200 studies from the HFBM literature. Each study has been categorized in terms of its experimental conditions, subject type and number, imaging modality, etc. [Fox et al., 2005]; all published location coordinates (and in some instances, deeper-than-published coordinates) have been collated. With the cooperation and assistance of the contributing authors and the BrainMap staff (S. Farmer and A. Uecker, in particular), all of these

studies have been entered into the BrainMap database [Fox and Lancaster, 1996a,b, 2002; Fox et al., 2005; Laird et al., 2005a], extending its data volume to greater than 3,000 experiments. Every meta-analysis reported in this special issue thus can be recreated rapidly using data now available online. Three of the meta-analyses in this issue [Derrfuss et al., 2005; Lancaster et al, 2005; Neumann et al., 2005] were based entirely on data entered into BrainMap as part of this virtual workshop. One meta-analysis [Glahn et al., 2005] used data from another meta-analysis [Owen et al., 2005] as a starting point. All contributions were used to assess the filter functions of the experimental taxonomy of the Brain-Map database [Fox et al., 2005], vis-à-vis HFBM meta-analysis. We now invite members of the HFBM community to explore this rich data resource and to experiment with the meta-analysis methods described and demonstrated here.

Finally, our virtual workshop produced many educational dividends. For the most part, the contributing authors were meta-analysis naïve at the outset; now all are reasonably sophisticated in the benefits of voxel-based meta-analysis and in state-of-the-art meta-analysis methods. Although senior investigators conceived all of the meta-analyses, the processes of literature review and data coding were carried out in many instances by students and post-doctoral fellows. This was highly successful. Supervised meta-analysis is an outstanding vehicle for introducing students to the imaging literature in a very focused and goal-directed manner. Two of us (P.T.F. and A.R.L.) teach a graduate-level course in voxel-based meta-analysis for just this reason. In this context, the present special issue provides a primer of examples of meta-analyses by many of leading scientists in the field. We strongly encourage our colleagues to consider using meta-analysis as a highly interactive educational tool, with this issue and its associated data sets (above) as the teaching exercises.

REFERENCES

Bolger DJ, Perfetti CA, Schneider W (2005): A cross-cultural effect revisited: Universal structures plus writing system variations. Hum Brain Mapp 25:92–104.

Brown S, Ingham RJ, Ingham JC, Laird AR, Fox PT (2005): Stuttered and fluent speech production: an ALE meta-analysis of functional neuroimaging studies. Hum Brain Mapp 25:105–117.

Buchsbaum BR, Greer S, Chang WL, Berman KF (2005): Meta-analysis of neuroimaging studies of the Wisconsin card sorting task and component processes. Hum Brain Mapp 25:35–45.

Buckner RL, Petersen SE (1996): What does neuroimaging tell us about the role of prefrontal cortex in memory retrieval? Semin Neurosci 8:47–55.

Chein JM, Fissell K, Jacobs S, Fiez JA (2002): Functional heterogeneity within Broca's area during verbal working memory. Psychol Behav 77:635–639.

Derrfuss J, Brass M, Neumann J, von Cramon DY (2005): Involvement of the inferior frontal junction in cognitive control: meta-analyses of switching and Stroop studies. Hum Brain Mapp 25:22–34.

Farrell MJ, Laird AR, Egan GF (2005): Brain activity associated with painfully hot stimuli applied to the upper limb: a meta-analysis. Hum Brain Mapp 25:129–139.

Fox PT (1995a): Spatial normalization: origins, objectives, applications, and alternatives. Hum Brain Mapp 3:161–164.

Fox PT (1995b): Broca's area: motor encoding in somatic space. Behav Brain Sci 18:344–345.

Fox PT, Huang A, Parsons LM, Xiong JH, Zamarippa F, Rainey L, Lancaster JL (2001): Location-probability profiles for the mouth region of human primary sensory-motor cortex: meta-analysis and validation. Neuroimage 13:196–209.

Fox PT, Laird AR, Fox SP, Fox PM, Uecker AM, Crank M, Koenig SF, Lancaster JL (2005): BrainMap taxonomy of experimental design: description and evaluation. Hum Brain Mapp 25:185–198.

Fox PT, Lancaster JL (1996a): An atlas du cerveau sur internet. La Recherche 289:49–51.

Fox PT, Lancaster JL (1996b): Neuroscience on the net. Science 226:994–996.

Fox PT, Lancaster JL (2002): Mapping context and content: the BrainMap model. Nat Rev Neurosci 3:319–321.

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, Mintun MA, Reiman EM, Raichle ME (1988): Enhanced detection of focal brain responses using intersubject averaging and change-distribution analysis of subtracted PET images. J Cereb Blood Flow Metab 8:642–653.

Fox PT, Parsons LM, Lancaster JL (1998): Beyond the single study: function-location meta-analysis in cognitive neuroimaging. Curr Opin Neurobiol 8:178–187.

Fox PT, Perlmutter JS, Raichle ME (1985): A stereotactic method of anatomical localization for positron emission tomography. J Comput Assist Tomogr 9:141–153.

Friston KJ, Liddle PF, Frackowiak RS (1991): Comparing functional (PET) images: the assessment of significant change. J Cereb Blood Flow Metab 11:690–699.

Frith CD, Friston KJ, Liddle PF, Frackowiak RS (1991): Willed action and the prefrontal cortex in man: a study with PET. Proc R Soc Lond B Biol Sci 244:241–246.

Glahn DC, Ragland JD, Abramoff A, Barrett J, Laird AR, Velligan DI (2005): Beyond hypofrontality: a quantitative meta-analysis of functional neuroimaging studies of working memory in schizophrenia. Hum Brain Mapp 25:60–69.

Grosbras MH, Laird AR, Paus T (2005): Cortical regions involved in gaze production, attention shifts and gaze perception. Hum Brain Mapp 25:140–154.

Laird AR, Lancaster JL, Fox PT (2005a): BrainMap: the social evolution of a human brain mapping database. Neuroinformatics (in press).

Laird AR, McMillan KM, Lancaster JL, Kochunov P, Turkeltaub PE, Pardo JV, Fox PT (2005b): A comparison of label-based review and ALE meta-analysis in the Stroop task. Hum Brain Mapp 25:6–21.

Laird AR, Fox PM, Price CJ, Glahn DC, Uecker AM, Lancaster JL, Turkeltaub PE, Kochunov P, Fox PT (2005c): ALE meta-analysis: controlling the false discovery rate and performing statistical contrasts. Hum Brain Mapp 25:155–164.

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.

Lancaster JL, Laird AR, Fox PM, Glahn DC, Fox PT (2005): Automated analysis of meta-analysis networks. Hum Brain Mapp 25:174–184.

Nielsen FA, Hansen LK (2002): Modeling of activation data in the BrainMap database: detection of outliers. Hum Brain Mapp 15:146–156.

Neumann J, Lohmann G, Derrfuss J, von Cramon DY (2005): The meta-analysis of functional imaging data using replicator dynamics. Hum Brain Mapp 25:165–173.

Owen AM, McMillan KM, Laird AR, Bullmore E (2005): The n-back working memory paradigm: a meta-analysis of normative functional neuroimaging studies. Hum Brain Mapp 25:46–59.

Paus T (1996): Location and function of the human frontal eye-field: a selective review. Neuropsychologia 34:475–483.

Pearson K (1904): Report on certain enteric fever inoculation statistics. Br Med J 3:1243–1246.

Petacchi A, Laird AR, Bower JM (2005): The cerebellum and auditory function. an ALE meta-analysis of functional neuroimaging studies. Hum Brain Mapp 25:118–128.

Picard N, Strick PL (1996): Motor areas of the medial wall: a review of their location and functional activation. Cereb Cortex 6:342–353.

Price CJ, Moore CJ, Morton C, Laird AR (2005): Meta-analysis of picture naming: the effect of baseline. Hum Brain Mapp 25:70–82.

Tan LH, Laird AR, Fox PT (2005): Neuroanatomical correlates of phonological processing of Chinese characters and alphabetic words: a meta-analysis. Hum Brain Mapp 25:83–91.

Tulving E, Kapur S, Craik FIM, Moscovitch M, Huole S (1994): Hemispheric encoding/retrieval asymmetry in episodic memory: positron emission tomography findings. Proc Natl Acad Sci USA 91:2016–2020.

Turkeltaub PE, Eden GF, Jones KM, Zeffiro TA (2002): Meta-analysis of the functional neuroanatomy of single-word reading: method and validation. Neuroimage 16:765–780.

Wager TD, Smith EE (2003a): Neuroimaging studies of working memory: a meta-analysis. Cogn Behav Neurosci 3:255–274.

Wager TD, Phan KL, Liberzon I, Taylor SF (2003b): Valence, gender, and lateralization of functional brain anatomy in emotion: a meta-analysis of findings in neuroimaging. Neuroimage 19:513–531.

Xiong J, Rao S, Jerabek P, Zamarripa F, Woldorff M, Lancaster JL, Fox PT (2000): Inter-subject variability in cortical activations during a complex language task. Neuroimage 12:326–339.