Beyond the single study: function/location metanalysis in cognitive neuroimaging

Peter T Fox, Lawrence M Parsons and Jack L Lancaster

Cognitive neuroimaging maps the brain locations of mental operations. This process is iterative, as no single study can fully characterize a mental operation or its brain location. This iterative discovery process, in combination with the location-reporting standard (i.e. spatial coordinates) of the cognitive neuroimaging community, has engendered a new form of metanalysis. Response locations from multiple studies have been analyzed collectively so as to better describe the spatial distribution of brain activations, with promising results. New hypotheses regarding elementary mental operations and their respective brain locations are being generated and refined via metanalysis. These hypotheses are being tested and confirmed by subsequent, prospective experiments. Function/location metanalysis is an important new tool for hypothesis generation in cognitive neuroimaging. This form of metanalysis is fundamentally different from the effect-size metanalyses prevalent in other literatures, with unique advantages and challenges.

Address
Research Imaging Center, University of Texas Health Science Center at San Antonio, 7703 Floyd Curl Drive, San Antonio, Texas 78284-6240, USA

Abbreviations
BA   Brodmann area
MT   middle temporal (area)
PET   position emission tomography
SMA   supplementary motor area

Introduction

Metanalysis is most generally defined as the *post hoc* combination of results from independently performed studies to better estimate a parameter of interest. Cognitive neuroimaging is a rapidly growing field (Figure 1) that has adopted spatial coordinates as a standardized, quantitative terminology for reporting the brain locations of mental operations [1–4]. This standard has prompted a new form of metanalysis: published studies are being quantitatively compared to determine whether apparently similar functions are performed by different brain locations and, conversely, whether apparently different functions are performed by a single brain region. While fitting within the broad definition of metanalysis, this type of metanalysis, here termed effect-location metanalysis, differs substantially from effect-size metanalyses performed in other disciplines, with many theoretical and methodological implications.

There are four main purposes of this review: first, to characterize metanalysis as performed in disciplines other than human functional brain mapping and to draw lessons therefrom; second, to characterize metanalysis as performed in human brain mapping and to present exemplars thereof; third, to characterize the strategies and methods for metanalysis that have been developed for human brain mapping; and, fourth, to point out some of the methodological challenges awaiting the emerging discipline of metanalysis in human brain mapping. Throughout this review, it is argued that metanalyses performed in human brain mapping (effect-location metanalyses) are fundamentally different from the effect-size metanalyses performed in other disciplines.

The metanalysis controversy

The original, and by far the most prevalent, form of metanalysis pools studies with nonsignificant effects to test for significance in the collective [5]. *Vis-à-vis* the above definition, such metanalyses re-estimate a statistical parameter: the p-value. Metanalyses of effect size frequently come under criticism. The most well-known pitfall of effect-size metanalysis is publication bias [6,7], also known as the 'file drawer effect', whereby negative results are omitted from the analysis; this bias is attributable to the reluctance of authors, reviewers, and editors to publish negative results. Selective omission of negative results falsely inflates a *post hoc* estimate of effect size. Thus, an ideal effect-size metanalysis (see e.g. [8]) must have access to all studies of the effect—published and unpublished.

Failure to correct for publication bias is not the only criticism of effect-size metanalysis. In the late seventies and early eighties, the prospect of mining important new results from the published literature using rather simple statistical machinery was cast in an overly optimistic light. Not unexpectedly, this attracted a rush of papers that committed diverse statistical and experimental errors. In the spirit of a frontier period for metanalysis, editors published first and asked questions later [9–13]. As a consequence, the reputation of metanalysis was tarnished rather broadly among statisticians and more selectively in fields in which metanalyses of poor quality were published. Fortunately, metanalysis has largely recovered its credibility and is growing at a rapid pace [9]. Nevertheless, the aspiring metanalyst of cognitive neuroimaging is well advised to become familiar with the concept, history, methods and pitfalls of metanalysis [14,15••,16•,17••,18••].

Published guidelines applicable to function/location metanalysis in cognitive neuroimaging are difficult to find, as most critiques target metanalyses of effect-size (rarely

Figure 1
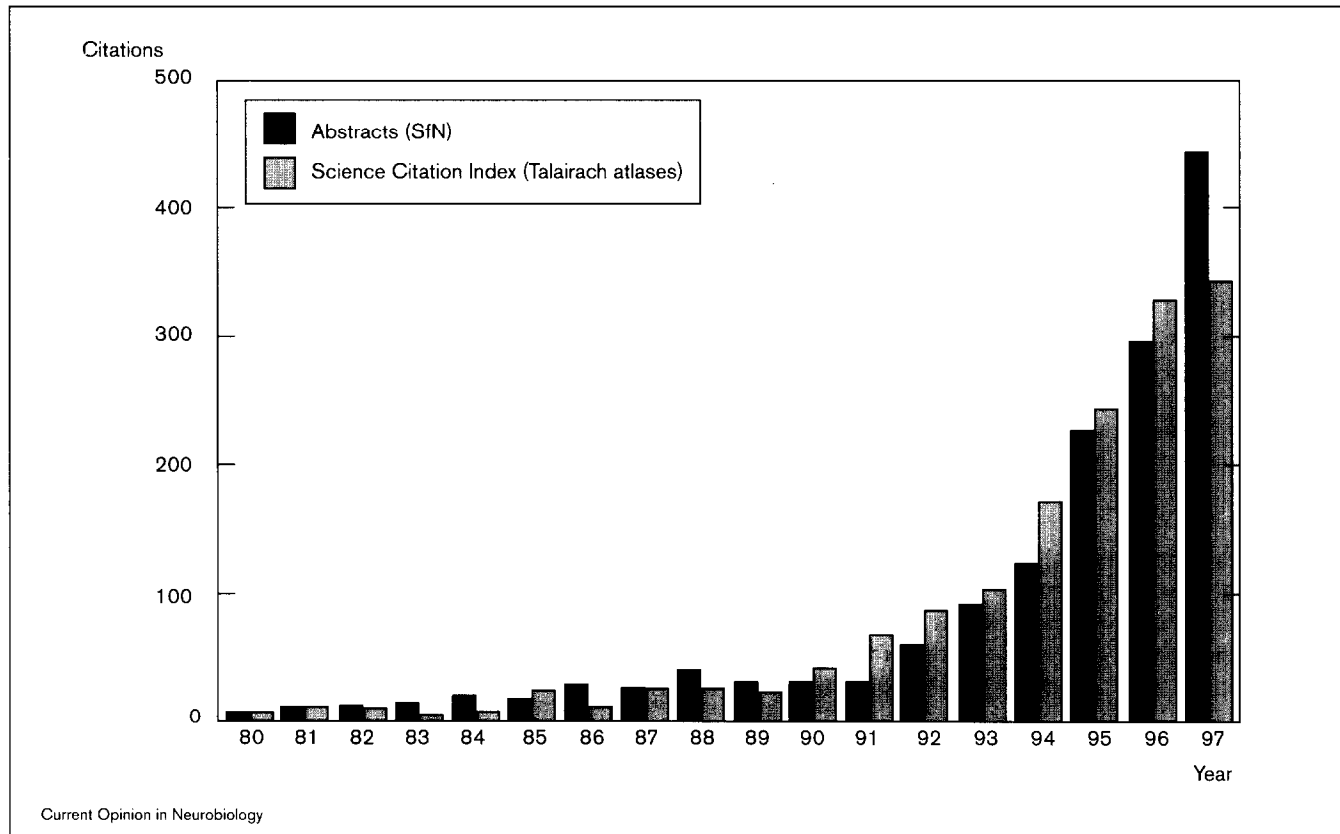
Two indices of the increased number of reports of human brain mapping are illustrated. Grey bars chart the number of publications using spatial coordinates, as determined by searching for citations (in Science Citation Index) of the Talairach atlases [38,39]. (The 1997 value is projected from bi-monthly values for January through October 1997.) Black bars chart the number of abstracts on functional mapping of the human brain reported at the annual meeting of the Society for Neuroscience (SfN).

performed in neuroimaging) or make domain-specific recommendations. The most useful tenet is that "the metanalysis is itself a study requiring careful planning and execution" [11]. As an experiment, a metanalysis should address a specific question, not troll through data hoping to land significant effects [10,15••]. All pertinent data should be identified, just as all entries into a clinical trial must be recorded. All data exclusions should be made prior to data analysis and only in accordance with pre-established exclusion criteria. Quality controls should be applied judiciously. Data quality, for example, can be enhanced by adhering to standards for sample size, analysis methods, and reporting conventions [12,13]. Note, however, that exclusions based on quality constitute a form of sampling bias [15••]. Analysis methods must be established before starting the study and should be appropriate to the research question and the nature of the data. Appropriate methodology, however, does not guarantee a valid metanalysis.

A frequent recommendation, also in keeping with the above tenet, is that metanalysis should be performed by experts in the subject matter rather than by statisticians [12,15••,16•]. Many abuses of metanalysis have been committed by statisticians fishing in unfamiliar waters [12]. Hypothesis generation and refinement, the most common roles of metanalysis in cognitive neuroimaging, are usefully performed only by subject-matter experts. Furthermore, laboratories prepared to test new hypotheses (with new imaging experiments) benefit most immediately from their creation. Not surprisingly, virtually all function/location metanalyses have been performed by subject-matter experts (not by statisticians or novices), and the majority have been performed by cognitive neuroimagers. Overall, the guidelines implied by the admonition that a metanalysis is an experiment have been spontaneously adopted by those shaping the new field of effect-location metanalysis (see below), probably as a result of the experimental sophistication of the parties involved.

Brain-mapping metanalyses are also buffered from traditional criticisms of metanalysis by focusing on effect location rather than on effect size. Effect-location metanalyses have employed statistically significant effects, with the parameter of interest being location rather than magnitude. Selective omission of subsignificant effects strongly biases metanalyses of effect size, but has a minimal and unbiased

influence on effect location. Inclusion of subsignificant effects, on the other hand (to avoid a publication bias), would artificially inflate the estimate of location variance, with minimal influence on estimated mean location. Another frequently offered recommendation for metanalysis is that the studies be of uniformly high quality and apply comparable methodologies. Without exception, the brain-mapping metanalyses discussed below combine only statistically significant effects reported in peer-reviewed journals, and the effects were extracted using analysis (statistical parametric imaging) and reporting (spatial coordinates) methods that meet widely accepted community standards. In this context, cautions of 'garbage in, garbage out', and recommendations, for example, that metanalyses compute quality scores for weighting data [19] are patently superfluous. Nevertheless, however rigorous a metanalysis may be, it must be viewed as only a step in the discovery process, rather than as an endpoint.

Dialectical metanalysis: an innovation in cognitive neuroimaging

Cognitive neuroimaging characterizes the information-processing tasks (mental operations) performed in specific brain locations and circuits. No single paradigm can perfectly isolate a cognitive operation. Nor can images from one set of subjects fully characterize the spatial location and distribution of a brain area in the population. Thus, the discovery process is intrinsically iterative, progressing through a gradual convergence of experiments from a number of independent laboratories. Driven by the need for converging data and grounded in the convention of reporting activation locations as spatial coordinates, an exciting form of scientific dialogue has emerged. Spatial clusterings across studies (or a lack thereof) are being aggressively explored. Overlap of responses from apparently dissimilar tasks suggests an underlying commonality of mental operation. Segregation of responses from apparently similar tasks implies a fundamental dissociation in processing.

In what was probably the first published function/location metanalysis in cognitive neuroimaging, Frith et al. [20] graphically plotted and tabulated coordinates from three reports of cortical activation present when internally generated motor or speech response conditions are contrasted to motor or speech response that were fully specified by immediate stimulus conditions. The authors used the analysis to guide interpretation of their results. In a similar manner, de Jong et al. [21] tabulated six PET studies of the middle temporal cortical region MT, computing location mean and variance, to interpret observed PET activations during the visual perception of movement in depth. The practice of including informal function/location metanalyses in the discussion section of brain-mapping reports in now nearly routine. The clear value of this practice has led to the publication of dedicated metanalyses, many of which introduce new hypotheses.

Fox [22] performed a metanalysis of nine studies reporting ten experiments activating frontal operculum (Broca's area). His analysis compared group mean activations from six independent PET studies involving overt and covert speech/oral movement and three PET studies with overt and covert non-oral tasks (i.e. actual hand movements, covert arm/hand movements and foot vibration). Overlap of the two sets of activations suggested that the frontal operculum supports not only the programming of oral/speech movements but also the programming of limb movements. This analysis prompted Fox to propose that the frontal operculum can be considered a ventral pre-motor area specialized for somatically (proprioceptively) defined movements, in contrast to the dorsal pre-motor area, which may be specialized for movements defined in reference to locations within a three-dimensional (3D) space. In a retrospective confirmation of this proposal, image averaging was applied to an early PET study of saccadic eye movements, contrasting saccades to visual targets with saccades in the dark (referenced chiefly to a proprioceptively defined location). Saccades to visual targets activated both dorsal and ventral lateral pre-motor cortex, whereas saccades in the dark activated only the ventral pre-motor region.

Tulving et al. [23] conducted a metanalysis of PET studies of the encoding and retrieval processes applied to information about specific episodes or events. The researchers tabulated the hemisphere in which prefrontal activation occurred in fourteen published papers of healthy subjects performing memory tasks. In these studies, there were nine conditions isolating the encoding of episodic information and eight isolating the retrieval of episodic information. Although there had long been indications from brain-damaged humans that the prefrontal cortex is involved in episodic memory, this metanalysis revealed an hemispheric asymmetry such that the left prefrontal cortex was activated during the encoding of episodic information and the right prefrontal cortex was activated during the retrieval of episodic information. The investigators built these findings into the 'hemispheric encoding/retrieval asymmetry' hypothesis, which has had a significant role in subsequent research on the neural basis of memory.

Paus [24••] recently has reviewed eight neuroimaging reports investigating the frontal eye fields, the principal area in cerebral cortex involved in oculomotor control (Figure 2). These reports provided 15 studies that employed eye movement tasks under four conditions: externally paced, internally paced, reflexive, or cognitive (i.e. with contingent or memorized targets). There was a tight clustering in the location of the group mean average activation of the different studies in both the rostro-caudal (y-axis) and dorso-ventral (z-axis) direction, but a much greater spread in the medio-lateral (x-axis) direction. In addition, a simple arithmetic average of the stereotaxic coordinates across all studies was computed. The metanalysis of activation location for each of the

Figure 2
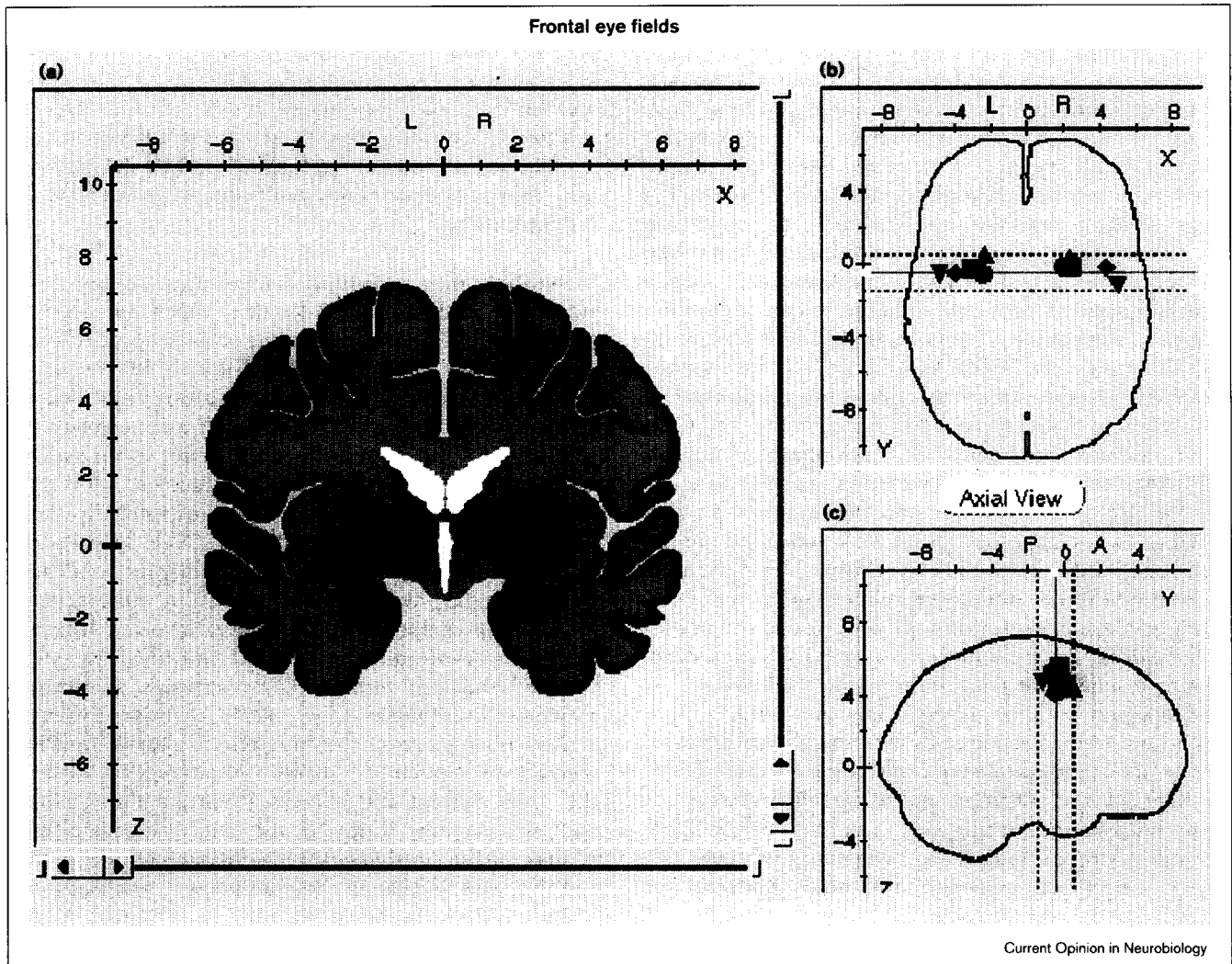
Depicted here is Paus' metanalysis of the frontal eye fields [24**]. The poor clustering (high variance) in the x (medial–lateral) axis **(a,b)** is in strong contrast with the tight clustering (low variance) in the y (anterior–posterior) axis **(a,c)** and z (superior–inferior) axis **(a,b)**. A spatial segregation of function was inferred by Haxby (JV Haxby, personal communication), with spatial working memory (and saccades to memory targets) being lateral, and oculomotor control (and saccades to visual targets) being medial. This functional segregation was subsequently confirmed by Courtney *et al.* [25].

four kinds of tasks indicated that the frontal eye fields were in the precentral sulcus, not in Brodmann area 8 as commonly believed, and were primarily involved in visuomotor rather than cognitive aspects of oculomotor control. These implications were consistent with Paus' subsequent review [24**] of six studies of humans with prefrontal lesions performing tasks similar to those included in his metanalysis.

Haxby and colleagues [25] hypothesized that the medio-lateral spread in activation sites for frontal eye fields in Paus' metanalysis [24**] were related to differences across tasks in the demand for spatial working memory. They also noted in an informal metanalysis that neuroimaging studies of spatial working memory that included dorsal and posterior frontal cortex consistently found activation

in the superior frontal sulcus; however, a spatial working memory role for the activation was discounted because researchers assumed the activations were in pre-motor or frontal eye field regions. Haxby and colleagues [25] predicted that, in light of the fact that in monkeys there is a spatial working memory area just anterior to the frontal eye fields, there should be an analogous spatial working memory area in humans. If so, then the location of the hypothesized area should be adjacent to the area in the precentral sulcus that Paus' metanalysis [24**] indicated human frontal eye fields were located. This prediction was confirmed by explicit experimental test using functional magnetic resonance imaging (fMRI) [25].

Picard and Strick [26**] conducted a grand metanalysis of tasks activating the medial wall of the frontal areas

of the left cerebral hemisphere. The 29 studies that met their criteria involved arm movement, eye movement, and speech, and, in combination, contained 46 tasks activating the supplementary motor area (SMA) and 69 tasks activating anterior cingulate cortex. These brain areas have long been known to be involved in several aspects of motor control as well as attentional processes related to response generation. The various experimental tasks were grouped into broad, relative categories: simple and complex. The simple category included tasks requiring the most basic spatial and temporal organization of movement and tasks that were overlearned or highly practiced. The complex category included tasks with additional motor or cognitive demands, such as the selection of a motor response or the acquisition of a conditional association. The investigators classified tasks rather than movements in order to examine the aspects of pre-motor function related to the integration of 'rules' and 'conditions' into movement production. They then conducted a metanalysis of activated locations for the two task categories and two brain areas. The metanalysis revealed that rostral areas subserved simpler tasks and caudal areas subserved more complex tasks, in both SMA and anterior cingulate. Within this broad dichotomy, there was somatotopy (face rostral, arm caudal) in both areas. Thus, the metanalysis implied the existence of four distinct functional areas in the medial wall, two in SMA and two in anterior cingulate cortex. These implications were supported by correspondences observed in a review of monkey neurophysiological studies [26••]. In light of the variation of methodology across the many tasks in the metanalysis, Picard and Strick [26••], neurophysiologists rather than neuroimagers, note that "the clustering of activation sites, which underlies our interpretations, is remarkable."

Buckner and Petersen [27••] performed two different metanalyses of neuroimaging activations from studies of tasks involving long-term memory retrieval. In the first metanalysis, they found that the same memory task performed in three independent studies each activated two 10 mm areas in prefrontal cortex, demonstrating the reliability of PET methods in localizing and resolving different functional areas in prefrontal cortex. In a related but more comprehensive examination of PET effect reliability, Poline et al. [28] conducted a 'prospective' metanalysis in which 12 laboratories (many speaking different native languages) performed the same study contrasting covert verb generation to eyes closed rest. The aim of the study was to evaluate the consistency of activations observed in the independent studies. A metanalysis of the results indicated "highly consistent results" across studies. Buckner and Petersen [27••] conducted a second metanalysis of ten studies of tasks in which subjects retrieved words or information from semantic memory (Figure 3) and eight studies of tasks in which subjects retrieved information from a specific study episode. The two kinds of memory tasks, semantic and episodic, were selected because a variety of behavioral,

neuropsychological, and animal neurophysiological data suggested they are distinct and dissociable memory functions. This suggestion was confirmed by the metanalysis [27••] in which semantic memory tasks over a wide range of material led to activations clustered in a circumscribed area of left inferior prefrontal cortex, whereas episodic memory tasks over a wide range of material all activated a circumscribed area in right anterior prefrontal cortex as well as more posterior prefrontal areas which varied across different stimuli.

Metanalysis from individual subject data has been considered a seldom-realized 'ideal', which (at least in other fields) must "remain in the realm of luxury" [13]. After only five years of function/location metanalysis in cognitive neuroimaging, there are already five such investigations [22,29•,30•,31••,32•]. Shulman and colleagues [29•,30•,31••] performed three separate re-analyses of the same nine within-lab individual subject PET data. The nine studies involved a variety of primarily visual information-processing tasks in visual search, spatial attention, memory, language, and mental imagery paradigms. Four (of nine) involved visual processes without motor or linguistic components; four others involved language processing with vocal responses; and the last study involved language with a manual response. In the first metanalysis, Shulman et al. [29•] examined whether active visual processing (contrasted to passive viewing) increases activations in medial visual regions early in the visual system and decreases activation in auditory and somatosensory cortex. In the re-analyses, six of nine studies produced significant modulation of the medial visual areas, indicating that top-down processes can affect early visual cortex, and two of five other studies produced decreases in left Brodmann area (BA) 41/42. In a 'mega-analysis' that pooled all of the studies, small reliable decreases were observed in insula, parietal operculum, and BA 40, areas that probably support somatosensory information. These analyses suggest then that precortical input to task-irrelevant sensory cortical areas is not broadly suppressed.

In the second re-analyses of these data, Shulman et al. [30•] found no consistent activation in the cerebral cortex outside the visual cortex, suggesting that many non-task-specific information operations (e.g. arousal, goal setting, operation sequencing, decision making) do not activate cortex, at least not strongly enough to be detected by current PET methods. There were activations in left and medial cerebellum related to both overt and covert responses, possibly implying support for higher level aspects of response generation. These activations were correlated with observed right thalamic activations. Activations in right cerebellar regions were not affected by the presence or absence of overt or covert motor responses but were affected by within-experiment variables that held motor response constant. The right cerebellar focus and observed left thalamic activations both increased with

Figure 3
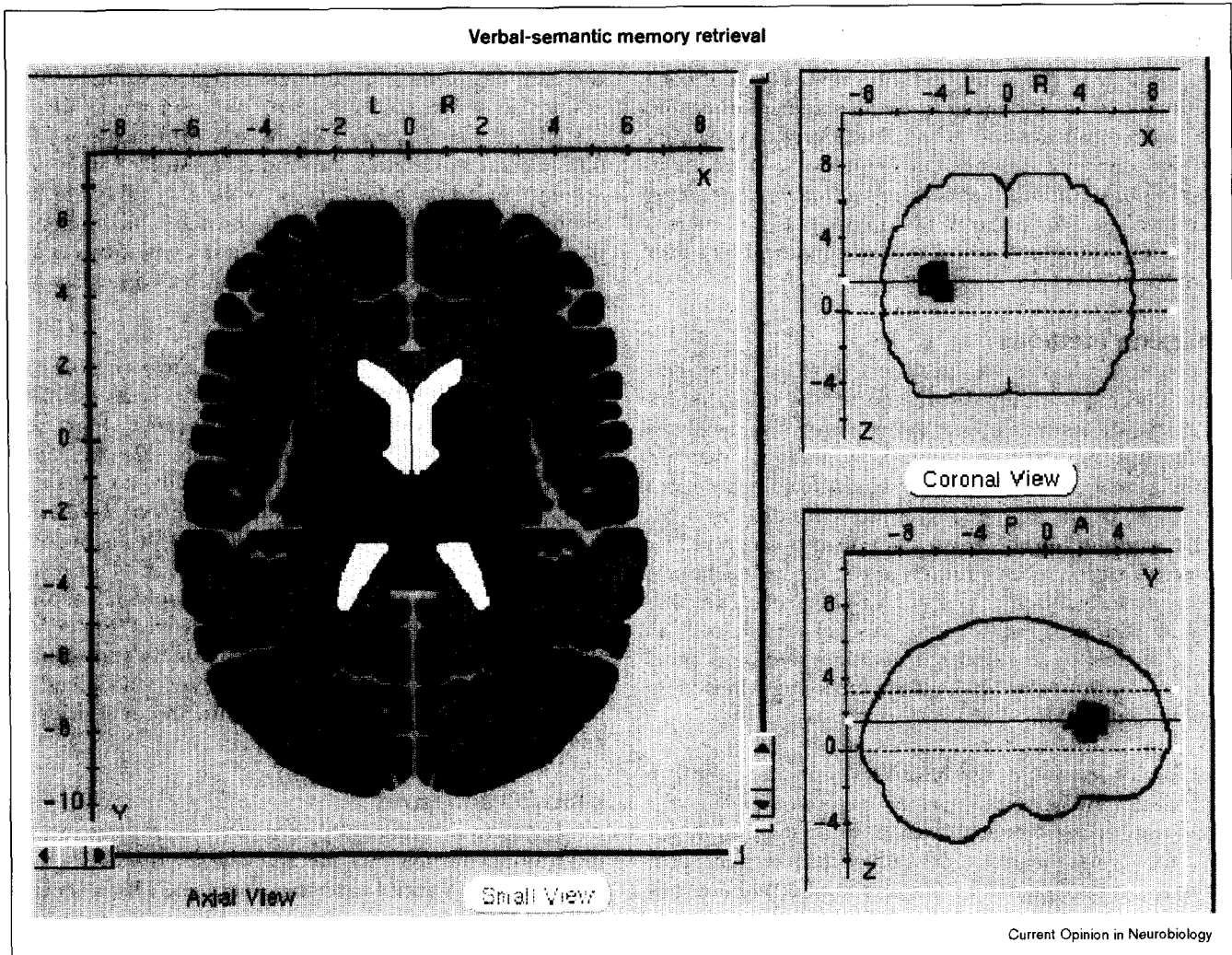
The eight studies of verbal-semantic memory retrieval reviewed by Buckner *et al.* [27••] illustrate extremely tight clustering in all three spatial axes. Buckner *et al.* infer that the diverse tasks reported by these studies all engaged a specific mental operation, which they characterize as verbal-semantic memory retrieval.

task complexity, suggesting a role for these regions in management of information processing.

In the third re-analysis of these data, Shulman *et al.* [31••] examined decreases in activation during the active tasks and found consistent decreases in posterior cingulate/precuneous, bilateral inferior parietal cortex, left dorsolateral frontal cortex, left lateral inferior frontal cortex, left inferior temporal gyrus, medial frontal regions, and right amygdala. These decreases in the contrast between active and passive tasks may reflect either decreased activity resulting from active task processes that generalize over task or increases attributable to processes specific to the passive task, such as unconstrained verbally mediated thoughts and monitoring of the external environment, body, and emotional state.

Metanalysis has also been used to estimate another neuroimaging parameter. Group-mean averaging, widely used in human brain mapping to increase response signal:noise ratio, has the disadvantage that per-subject spatial variance about the mean location is unknown. Lacking a value for spatial variance, there is no ready measure by which to gauge whether two activated brain areas are far enough apart to be considered different. Huntun *et al.* [32•] performed a re-analysis of PET data produced within a single laboratory in order to assess functional and anatomical variability. The re-analysis compared two independent groups of subjects performing the same tasks. The researchers tested whether activations present in one group of subjects would replicate in a second group performing the same tasks and found that the activation in the second group was well predicted by

the strength of activation in the first group in each case. The researchers also evaluated how closely the location of peak activations clustered across subjects and found the variability for activation sites was 6.6 mm. They also found that the activation sites followed a normal distribution about the mean in each coordinate direction. The size and distribution of the activation-site variability was homogeneous across cerebral cortex and cerebellum. Huntun's observations that response locations from different individuals form a narrow (2–6 mm standard deviation) Gaussian spatial distribution has been confirmed by Fox et al. [33••] and Hasnain et al. [34•].

Emerging methods

Response coordinates are ideal data for parametric metanalysis, being normally distributed sets of real-numbers [32•,33••,34•]. As seen above, the majority of metanalyses exploit this property in a relatively informal manner, graphing multiple studies using standardized coordinates to visualize clusterings and segregations. To date, no studies have applied parametric statistical tests to prove, for example, that the mean locations of two mental operations are significantly different or that a psychophysical variable co-varies with response location. This gap is more likely to be attributable to an absence of methods and precedents than to a lack of data. Methods are emerging which are a prelude to more formal statistical analyses.

Community-accessible databases are important aids to metanalysis. Some authorities have recommended that community-accessible databases provide only summary data (e.g. group-mean values) rather than original data, avoiding issues of confidentiality and intellectual property [13]. BrainMap® [35•] is a community database of the human brain-mapping literature, which provides experimental conditions (tasks), brain locations (spatial coordinates), imaging methods, statistical methods and effect sizes, subject population and sample sizes, and citations. BrainMap® (http://ric.uthscsa.edu/services) was created to facilitate function/location metanalyses and to serve as a hub for federating emerging databases of human neuroscience. Talairach Daemon™ (http://ric.uthscsa.edu/projects) is a database of anatomical (rather than functional) spatial probabilities, which has already been federated with (linked to) BrainMap® [36]. Links between BrainMap® and the CARET (http://v1.wustl.edu/caret.html) visualization system of Drury and Van Essen [37••] are in progress. Tools for computing multi-study mean and variance, both raw and weighted by sample size (functional volumes modeling, below) are in place or in progress.

Drury and Van Essen [37••] recently introduced a procedure and software (CARET) by which volume coordinates (3D) are transformed to folded-surface coordinates (2.5D) and then unfolded to coordinates on a flattened map of the cortex (2D). In essence, this is a spatial normalization that corrects for the foreshortening of cortical distances created by folding. For perisulcal sites,

surface distances will be much greater (by a factor of 2–3) than 3D distances, making functional areas more readily differentiated from one another both visually and statistically. By this method, the investigators compared the locations of right hemisphere activations observed in a study in which subjects heard and read music and in a study in which subjects heard and read single words. The metanalysis vividly illustrated clustering within and segregation between task types. While this study did not perform a statistical analysis to demonstrate the difference in mean location between the two tasks types, this could readily be done if site mean and variance are calculated in this flattened space.

Functional volumes modeling is a strategy introduced by Fox et al. [33••] for calculating a best estimate of response location and location variance from the published literature. Reported locations are weighted by the number of subjects in each study, to obtain a weighted mean location. The variance among studies is computed in a similar manner and corrected using the best available estimate of per-subject variability. The result of these computations is a 'probabilistic volume', setting confidence limits for 3D functional regions (Figure 4). Probabilistic volumes can be used for a variety of purposes, including statistical analysis, metanalysis, modeling of functional areas and functional systems, and neuroscience education.

Unresolved issues

As cognitive neuroimaging evolves toward statistically more formal metanalyses, methodological issues will arise, some of which can be readily foreseen. Discrepancies between laboratories in spatial coordinate computation and reporting will need to be identified and rectified; for example, corrections for inter-study differences in sample size and statistical analysis methods will need to be developed. By far the most daunting issue, however, is selecting the results to be included in a metanalysis, that is, unbiased literature-sampling strategies.

Metanalyses in cognitive neuroimaging address function/location correspondences. Virtually every activation task engages a considerable number of elementary mental operations and, thereby, activates a constellation of functional areas. Even the most closely matched control condition corrects for (subtracts off) only a fraction of these areas. Thus, each reported experiment will be contaminated by activations irrelevant to the operation/region of interest. These irrelevant activations must be excluded from the metanalysis. In some instances irrelevant activations will be readily identified; in other instances, this distinction is less clear and the exclusion process may introduce bias. For example, motor-area activations linked to button press responses to confirm task performance in a visual-spatial attention task will be far removed from occipital and parietal attentional effects and can be safely discarded. On the other hand, SMA responses may lie close to

Figure 4
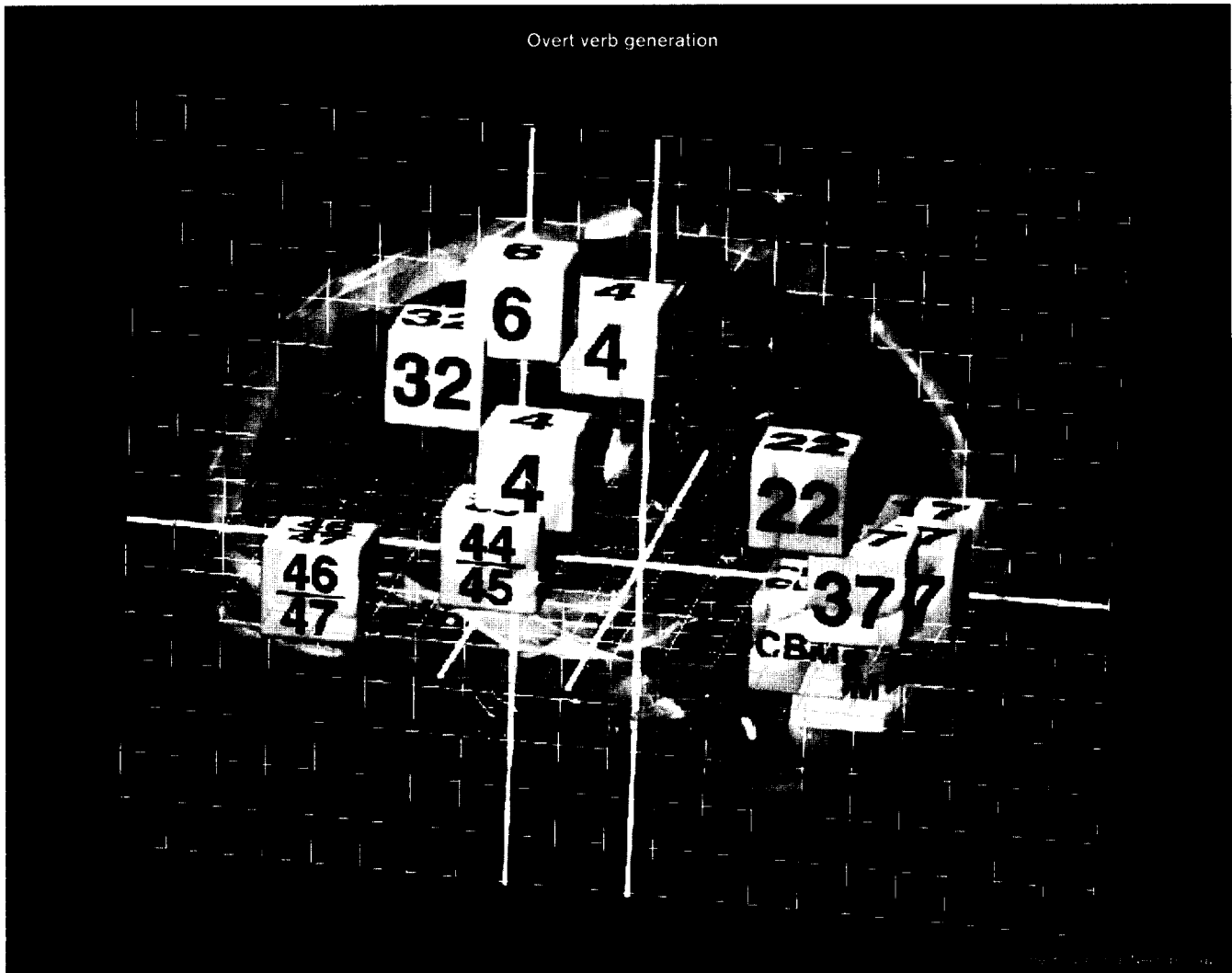
A functional volumes model [33••] of the principal brain areas engaged by overt verb generation was computed by metanalyses of each of the individual areas. Where possible, published metanalyses (e.g. [26••,27••,33••]) were used. Each parallelepiped represents the mean location ±2 sd (population), predicting the 95% confidence intervals for the population of individual subjects performing this task. The numbers are Brodmann designations. CBM, cerebellum.

attention-related activations in the anterior cingulate; here, the metanalyst must make a difficult decision and may bias the sample.

Guidelines for sampling can be envisioned. Sampling is a three-tier process: first, there is sampling among papers; second, among experiments within a paper; and third, among response locations within an experiment. Paper sampling could be based on adherence to analysis and reporting standards, statistical threshold, observation sample size and the like. Experiment sampling could be based on a cognitive analysis of the task state. If the mental operation of interest is probably engaged, the task is included. Discarding individual responses must be justified by the control state, which fails to remove irrelevant activations. This still leaves the

problem of selecting among adjacent, functionally related responses (e.g. distinguishing the primary motor mouth representation from Broca's area [22,33••]). What are the criteria? Centers experimenting with metanalyses would be well advised to address these issues immediately, rigorously, and openly. Collectively achieved solutions are far more likely to gain community acceptance than those adopted in isolation.

Conclusions

The field of human brain mapping is large and rapidly expanding in scope and volume. Synthesis of observations from different laboratories and different methods is needed. Metanalysis is filling this need. Initially performed rather informally, metanalyses in cognitive neuroimaging are advancing in sophistication, comprehen-

sivity and number. Metanalyses have already explored a wide range of cognitive domains, including vision, semantics, language, music, movement planning, memory encoding and retrieval, and spatial working memory. Being exploratory in nature, metanalyses naturally serve to suggest and refine hypotheses, which are prospectively tested by traditional imaging experiments. Several examples of new hypotheses generated through metanalysis have already emerged (above). In at least one instance [24**,25] the dialectical loop (hypothesis → mapping experiments → metanalysis → new hypothesis → new mapping experiments) has been closed by experimental confirmation of the new hypothesis.

Function/location metanalysis appears to be theoretically and methodologically different from effect-size metanalysis, commonly performed in other fields. These differences protect effect-location metanalysis from many of the pitfalls and criticisms addressed to effect-size metanalyses. Formal statistical methods for function/location metanalyses remain rudimentary. The complexity of neuroimaging experiments and data gives rise to equally complex sample-selection issues, for which rigorous procedures will need to be developed.

Acknowledgements

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Fox PT: **Spatial normalization: origins, objectives, applications and alternatives.** Hum Brain Mapp 1995, 3:161-164.

2. Lancaster JL, Glass TG, Lankipalli BR, Downs H, Mayberg H, Fox PT: **Spatial normalization.** Hum Brain Mapp 1995, 3:209-223.

3. Collins DL, Holmes CJ, Peters TM, Evans AC: **Automatic 3-D model-based neuroanatomical segmentation.** Hum Brain Mapp 1995, 3:190-208.

4. Friston JK, Ashburner J, Frith CD, Poline JB, Heather JD, Frackowiak RSJ: **Spatial registration and normalization of images.** Hum Brain Mapp 1995, 3:165-189.

5. Pearson K: **Report on certain enteric fever inoculation statistics.** Br Med J 1904, 3:1243-1246.

6. Felson DT: **Bias in meta-analytic research.** J Clin Epidemiol 1992, 45:885-892.

7. Easterbrook PJ, Berlin JA, Gopalan R, Matthews DR: **Publication bias in clinical research.** Lancet 1991, 337:867-872.

8. Simes RJ: **Publication bias: the case for an international registry of clinical trials.** J Clin Oncol 1986, 4:1286-1291.

9. Olkin I: **A critical look at some population meta-analytic methods.** Am J Epidemiol 1994, 140:297-299.

10. Nony P, Cucherat M, Haugh MC, Boissel JP: **Critical reading of the meta-analysis of clinical trials.** Therapie 1995, 50:339-351.

11. Jones DR: **Meta-analysis: weighing the evidence.** Stat Med 1995, 14:137-149.

12. Bailar JC: **The practice of meta-analysis.** J Clin Epidemiol 1995, 48:149-157.

13. Olkin I: **Statistical and theoretical considerations in meta-analysis.** J Clin Epidemiol 1995, 48:133-146.

14. Sacks HS, Berrier J, Reitman D, Ancona-Berk VA: **Meta-analyses of randomized controlled trials.** N Eng J Med 1997, 316:450-455.

15. Conn HO: **Interpretation of data from multiple trials: a critical**
•• **review.** J Int Med 1997, 241:177-183.
This paper provides criteria for designing and performing a metanalysis to assure high quality. The limitations are that metanalysis is considered only in the narrow sense of assessing effect significance (i.e. estimating p-values) and that source material is limited to randomized clinical trials. Nevertheless, many of the recommendations and cautions can be extrapolated to metanalysis in human brain mapping.

16. Friedman HP, Goldberg JD: **Meta-analysis: an introduction and**
•   **a point of view.** Hepatology 1996, 23:917-928.
This paper provides a good introduction to metanalysis, including statistical models and sources of controversy. The limitations are that metanalysis is considered only in the narrow sense of assessing effect significance (i.e. estimating p-values) and that source material is limited to randomized clinical trials.

17. Petitti DB: **Meta-analysis and endocrinology.** Endocrinol Metab
•• Clin North Am 1997, 26:31-44.
Includes an excellent discussion of some limitations of metanalysis, including publication bias, quality control and data heterogeneity. The limitations are that metanalysis is considered chiefly in the narrow sense of assessing effect significance (i.e. estimating p-values) and that source material is limited to clinical trials.

18. Mosteller F, Colditz GA: **Understanding research synthesis**
•• **(meta-analysis).** Annu Rev Public Health 1996, 17:1-23.
This is a recent, clear, comprehensive exposition of effect-size metanalysis applications, methods and pitfalls.

19. D'Agnostino RB, Weintraub M: **Meta-analysis: a method for synthesizing research.** Clin Pharmacol Ther 1995, 58:605-616.

20. Frith CD, Friston K, Liddle PF, Frackowiak RSJ: **Willed action and the prefrontal cortex in man: a study with PET.** Proc R Soc Lond [Biol] 1991, 244:241-246.

21. de Jong BM, Shipp S, Skidmore B, Frackowiak RSJ, Zeki S: **The cerebral activity related to the visual perception of forward motion in depth.** Brain 1994, 117:1039-1054.

22. Fox PT: **Broca's area: motor encoding in somatic space.** Behav Brain Sci 1995, 18:344-345.

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

24. Paus T: **Location and function of the human frontal eye-field: a**
•• **selective review.** Neuropsychologia 1996, 34:475-483.
The first 'stand-alone' metanalysis to compute mean location and variances (unweighted), in the context of drawing new theoretical inferences from published data.

25. Courtney SM, Petit L, Ma Maisog J, Ungerleider LG, Haxby JV: **An area specialized for spatial working memory in human frontal cortex.** Science 1998, 279:1347-1350.

26. Picard N, Strick PL: **Motor areas of the medial wall: a review**
•• **of their location and functional activation.** Cereb Cortex 1996, 6:342-353.
The largest metanalysis to date, this study advances important new hypotheses on the functional organization of the medial motor areas. Performed by neurophysiologists (rather than neuroimagers), this study illustrates the accessibility of the coordinate-based literature to scientists in neighboring disciplines.

27. Buckner RL, Petersen SE: **What does neuroimaging tell us**
•• **about the role of prefrontal cortex in memory retrieval?** Semin Neurosci 1996, 8:47-55.
Metanalysis leads to the proposal that different regions of the prefrontal cortex serve different roles in long-term memory. This paper reports two, independent metanalyses. One metanalysis demonstrates the consistency of functional activations across studies for the same task, thereby constituting a partial validation of function/location metanalysis. The other metanalysis assesses functional/operational commonalities of several working-memory tasks. The metanalysis leads to the hypothesis that episodic working memory tasks activate a different functional area (right anterior prefrontal) than do semantic tasks (left interior prefrontal), but that within the spectrum of

semantic tasks, the type of material has no influence on function location. This is an excellent example of hypothesis generation through metanalysis.

28. Poline J-B, Vandenberghe R, Holmes AP, Friston KJ, Frackowiak RSJ: **Reproducibility of PET activations studies: lessons from a multi-center European experiment.** *Neuroimage* 1996, 4:425-437.

29. Shulman GL, Corbetta M, Buckner RL, Raichle ME, Fiez JA,
• Miezin FM, Petersen SE: **Top-down modulation of early sensory cortex.** *Cereb Cortex* 1997, 7:193-206.
See annotation [31••].

30. Shulman GL, Corbetta M, Buckner RL, Fiez JL, Miezin FM,
• Raichle ME, Petersen SE: **Common blood flow changes across visual tasks: I. Increases in subcortical structures and cerebellum but not in nonvisual cortex.** *J Cogn Neurosci* 1997, 9:623-646.
See annotation [31••].

31. Shulman GL, Fiez JL, Corbetta M, Buckner RL, Miezin FM,
•• Raichle ME, Petersen SE: **Common blood flow changes across visual tasks: II Decreases in cerebral cortex.** *J Cogn Neurosci* 1997, 9:647-662.
These three papers [29•,30•,31••] comprise the largest metanalysis of original data (not summary data) performed to date, setting a new standard for data quality and statistical rigor. These studies offer a variety of important hypotheses ranging from functional organization to hemodynamic physiology. The discussion of the mechanisms underlying task-related decreases in blood flow [31••] is extremely thoughtful.

32. Huntun DL, Miezin FM, Buckner RL, van Mier HI, Raichle ME,
• Petersen SE: **An assessment of functional-anatomical variability in neuroimaging studies.** *Hum Brain Mapp* 1996, 4:122-139.
This paper provides the best estimate, so far, of inter-subject variability in brain functional organization.

33. Fox PT, Lancaster JL, Parsons LM, Xiong JH, Zamarripa F:
•• **Functional volumes modeling.** *Hum Brain Mapp* 1997, 5:306-311.

The first effort to correct estimates of spatial location and variance for sample size. Suggests a variety of applications of spatial models, including neural-systems modeling, assessing neural plasticity and a method of hypothesis-driven analysis of statistical parametric images.

34. Hasnain M, Fox PT, Woldorff MG: **Intersubject variability of**
• **functional areas in the human visual cortex.** *Hum Brain Mapp* 1998, 6:in press.
Similar to the study by Huntun *et al.* [32•], this study provides an estimate of the inter-subject variability of functional areas, also finding uniformity of spatial variance within the visual system.

35. Fox PT, Lancaster JL: **Neuroscience on the net.** *Science* 1996,
• 226:994-996.
Presents the BrainMap® database in the context of an overview of the evolution and utility of databases in neuroscience. Touches on the role of databases in metanalysis and the long-term need for database federation.

36. Lancaster JL, Rainey LH, Summerlin JL, Frietas CS, Fox PT, Evans AC, Toga AW, Mazziotta JC: **Automated labeling of the human brain: a preliminary report on the development and evaluation of a forward-transform method.** *Hum Brain Mapp* 1997, 5:238-242.

37. Drury HA, Van Essen DC: **Functional specialization in human**
•• **cerebral cortex analyzed using the visible man surfaced-based atlas.** *Hum Brain Mapp* 1997, 5:233-237.
This study introduces a surface-flattened atlas of the human brain (created from the Visible Man), in the context of an illustrative metanalysis. An excellent example of the value of inter-database linkages.

38. Talairach J, Szikla G, Tournoux P: *Atlas d'Anatomie Stereotaxique du Telencepahle.* Paris: Masson; 1967. [Title translation: Stereotaxic atlas of the anatomy of the telencephalon.]

39. Talairach J, Tournoux P: *Co-planar Stereotaxic Atlas of the Human Brain.* New York: Thieme Medical Publishers, Inc; 1988.