

nature neuroscience

A debate over fMRI data sharing

Michael Gazzaniga, a prominent cognitive neuroscientist, has caused a stir in the neuroimaging community. Gazzaniga is the director of the newly established National fMRI Data Center, a public archive based at Dartmouth College that seeks to provide a repository for the enormous datasets generated by fMRI experiments. On the face of things, this would seem to be a welcome move, but Gazzaniga's announcement has alarmed many of his colleagues, primarily because of his suggestion that several journals—including *Journal of Cognitive Neuroscience*, of which he is the editor—were planning to require the release of primary data as a condition of publication.

In response to Gazzaniga's announcement, a circular letter, coordinated by Isabel Gauthier of Vanderbilt University and signed by more than forty researchers, was recently sent to the editors of various leading journals (including both *Nature* and *Nature Neuroscience*), expressing concerns about the proposed policy and asking them to clarify their positions. *Journal of Cognitive Neuroscience* intends to stand by its stated policy, but most other journals are taking a more cautious approach, and are distancing themselves from any suggestion that they might compel their authors to release data. Richard Frackowiak of University College London, who is editor of *Neuroimage*, put it bluntly in his public response to Gauthier: "...I would like to take a strategy of minimal response to this rather unfortunately launched proposal."

Although Gazzaniga may have misjudged the level of support he could expect from the community, he has clearly forced the pace in an important debate. Sharing of published data is standard practice in many other fields, and in a commentary on page 863 of this issue, Stephen Koslow of the US National Institute of Mental Health (NIMH) argues that neuroscience risks falling behind if it does not adopt similar practices. Gazzaniga's initiative represents a bold attempt to push the fMRI field in this direction, and as such, the debate deserves wide attention.

To understand the issues, and the difficulties inherent in archiving brain imaging data, it is important to consider the steps that go into a typical experiment. Functional MRI detects changes in blood oxygenation and blood flow that occur when brain tissue becomes more active. Subjects are therefore scanned as they perform various mental tasks, and by comparing the task of interest against some baseline, one can identify brain regions activated by the task. Within this general framework, however, there exist an enormous variety of protocols, both for performing the scans and for analyzing the resulting datasets.

The raw output of the scanner must first be converted into three-dimensional space, yielding a map of signal amplitude that is divided into voxels, volume elements that are typically a few millimeters across. Each subject may be scanned at two- or three-

second intervals for many minutes, yielding a time-series for each of many thousands of voxels. Once these four-dimensional maps are generated and corrected for noise and motion artifacts, they can be analyzed in many ways. The simplest is to compare the average activity between different cognitive tasks, in order to identify groups of voxels that are significantly activated (or deactivated) by a given task relative to a control condition. More complex methods include parametric designs, in which one looks for graded activation that is correlated with some cognitive variable; event-related analysis, to identify patterns of activity that are temporally correlated with the stimulus or the behavioral output; and higher-order analyses in which one looks for correlations between different brain regions in order to make inferences about their functional interactions.

Finally, the activations must be superimposed onto some standard anatomical frame of reference. The most common of these is Talairach coordinate space, but because of the variability between individual brains, Talairach coordinates do not correspond precisely to anatomical structures, and nor do the anatomical features of the cortex correspond precisely to its functional subdivisions. There is thus a chain of inferences involved in identifying a given activation with a specific brain region.

The datasets generated by a single study can run to gigabytes or even terabytes, many orders of magnitude more than can be depicted in printed form. Typically, what finally ends up in a published report is a list of locations (often in Talairach coordinates) that showed significant activation in the condition of interest. These coordinates represent only the center of each activation, and say nothing about its absolute magnitude or spatial distribution. Although it is common to discuss some of the activations in greater detail, and to illustrate them with images and/or graphs, the final report nevertheless represents a very reduced subset of the original data, filtered through a series of transformations and analyses that are often idiosyncratic. There is no consensus about the 'right' way do these analyses; each has its strengths and weaknesses, and new methods are constantly being developed as the field evolves.

This would seem to add up to a strong case for sharing the raw data, along with the associated analytical tools. Supporters of this view identify a number of advantages. Perhaps most importantly, it would allow meta-analysis, in particular to identify common features of tasks that activate specific brain regions. Marcus Raichle of Washington University in St. Louis describes this as "an incredibly powerful approach," particularly for brain structures whose functions are poorly understood. However, it is still extremely difficult to do this type of analysis; the lists of activations reported in the published literature are not rich enough to support detailed comparisons, and attempts to com-

pare raw data across multiple studies are often hampered by the lack of common data formats.

The ability to examine primary data would also allow researchers to investigate the robustness of published conclusions to different analytical methods and statistical significance thresholds. Any choice of threshold involves a subjective decision about the appropriate trade-off between false positives and false negatives, and activations that fail to reach a stringent significance criterion are often not reported. Moreover, as John Gabrieli of Stanford University notes, there may be a systematic bias to the published literature in this regard; it is difficult to write papers describing complex patterns of activation, and so there is often a temptation to set a conservative threshold, in order to reduce the number of activations and present a simpler story.

Not everyone is convinced, however, that these benefits will actually materialize as a result of the Dartmouth initiative. There are two major obstacles that must be overcome if the archive is to have any chance of success. First, it must be able to store detailed descriptions not only of the imaging data but also of the psychological context of each study. The field has moved beyond 'brain mapping' (although the name has stuck), and most fMRI studies are now designed to test specific hypotheses. Thus, if the data are to be interpreted, let alone replicated, they must be accompanied by detailed descriptions of stimuli, behavioral responses and, in some cases, psychological or clinical data about individual subjects. Codifying such diverse datasets represents a formidable computing problem, and it remains to be seen how effectively the Dartmouth team will be able to solve it.

The second prerequisite for success is that the archive must be easy to use. The sheer size of the datasets means that transmitting them is not trivial, and in addition to size, the file formats will inevitably be very complex, much more so than for (say) genomic data. The Dartmouth team warns that "some growing pains may be expected", and it is not yet clear how painful they will be. An early indication will come from a forthcoming special issue of *Journal of Cognitive Neuroscience*, edited by Mark D'Esposito (U.C. Berkeley) and devoted entirely to fMRI studies. The authors of the research papers have all agreed to deposit their data into the Dartmouth archive, effectively serving as its beta-testers. At the time of writing, however, no data have been deposited, and D'Esposito anticipates that it may be many months before the process is complete.

In addition to these technical challenges, several other concerns have been raised. Perhaps the most serious is that many datasets require months of analysis and lead to multiple publications. If the data are released at the time of the first publication, the authors could be 'scooped' to other discoveries arising from their work. The concern seems reasonable in principle, and few would question the importance of maintaining incentives for authors to do experiments. It is unclear how often this would happen in practice—to publish a scientifically credible analysis would require a depth of understanding that could probably only come from discussion with the authors—but in any case, it can easily be addressed by storing the data under an embargo for some period after publication (perhaps a few months) to allow the authors a period of exclusive access. Such arrangements are common in other

fields such as structural biology, and provided that the database can support timed access, individual journals can set time limits as they see fit.

The National fMRI Data Center is not itself in a position to enforce a policy of data sharing. In general, the strongest enforcement for such policies comes from funding bodies, which can insist on disclosure as a precondition for grant support. NIH, for example, requires this in fields such as genomic research and structural biology, and grants are designed to cover the costs of archiving data and physical samples. Steven Hyman, the director of NIMH, favors the idea of archiving fMRI data, at least in principle, commenting that "right now, based on published images, investigators really cannot tell whether their findings are more than crudely in agreement." Hyman emphasizes, however, that in all cases where sharing is mandated, it has been after extensive consultation with the community. Before any policy is made regarding fMRI data, he says, "it is important that the field be engaged in a discussion of what is to be shared, with what timing, and how."

That discussion will need to be broad and inclusive, both scientifically and geographically, and its scope will undoubtedly extend beyond the Dartmouth initiative (which calls itself a 'national' center and has an advisory board that is purely US-based). One important forum will be the Organization for Human Brain Mapping, which has established a task force on neuroinformatics, chaired by Jonathan Cohen of Princeton University and comprising researchers from both North America and Europe. Cohen emphasizes that his group has no mandate to impose standards, and might not even make specific recommendations; at this stage, its mission is to reach a consensus on the issues that need to be considered by the field as a whole if data sharing is to become a reality.

In the meantime, what role should journals take? In principle, every journal is free to impose whatever policy it chooses. In reality, however, editors must balance the interests of their readers with those of their contributors. A journal that fails to provide readers with enough information to evaluate its papers will lack scientific credibility, whereas one that makes excessive demands of authors will find itself losing submissions. Journals operate in a free market, and those whose policies fail to reflect the consensus within their communities will suffer scientifically and ultimately commercially.

How to balance these factors depends on the field. In genetics, for instance, most journals insist that DNA sequence data be publicly available at the time of publication. In structural biology, some journals insist on making crystal coordinates available at the time of publication, whereas others allow an embargo period (typically 6 to 12 months). In general, any data-sharing policy can only be effective if two criteria are met. First, there must be an efficient technological infrastructure so that data can be contributed, retrieved and reanalyzed without excessive effort. Second, there must be a political consensus within the field that the benefits of sharing data outweigh the costs. In the case of neuroimaging data, neither criterion has yet been met, in our view, and so although we have no objection to our authors depositing their fMRI data in the National Data Center (or any other similar archive), we shall not insist that they do so.