ELSEVIER

NeuroImage

www.elsevier.com/locate/ynimg
NeuroImage 37 (2007) 1042–1044

Commentary

Functional magnetic resonance imaging: Measuring versus estimating

Luciano Fadiga

University of Ferrara, Italy, and Italian Institute of Technology, Genova, Italy

Received 24 November 2006; revised 26 February 2007; accepted 28 February 2007
Available online 6 March 2007

Brain imaging techniques largely spread in neuroscience literature. Due to initial technical limitations such as the very low signal-to-noise ratio, group experiments became the rule. This fact, together with the wide use of standard brains to localize the activations, lead several experimenters to the wrong idea that the brain can be described by a Cartesian coordinate system, neglecting at the same time the importance of individual neuroanatomy. My commentary on the paper by Devlin and Poldrack reinforces their reminder that it is necessary to deal with anatomy. Moreover, it adds some considerations on the relevance of single subjects studies and on the importance of the BOLD intensity signal, which should be used to describe brain activity together with the most used statistical tools.

This contribution is voluntarily provocative. Its goal is to take inspiration from the paper by Devlin and Poldrack to underline some aspects of brain imaging technique that, in my view, should be revised in depth. Indeed, some of the reasons that lead neuroscientists to analyze brain imaging data by using the statistical approach to group analysis are now obsolete. The power of the new generation machines allow us today to change the mental habitus and to consider single subjects study as a unique possibility to shed light on the intimate brain mechanisms. The reason for which this approach is only sporadic and concerns almost only high-quality studies is, in my view, because it is much easier to localize brain functions by using the brain cartography of standard atlases in place of localizing them in terms of individual neuroanatomy. I am aware that inter-individual variability is a big challenge for neuroscientists. However, I am convinced that it is time to afford this issue frontally, in place of hiding the problem by spatially smoothing individual activations with Gaussian windows. These issues become particularly relevant if one is aiming at investigating brain physiology in individual patients.

I think that the issue raised in the paper by Devlin and Poldrack should be shared by everyone working with brain imaging: when analyzing fMRI data is mandatory to deal with neuroanatomy. Unfortunately, in several studies, the happiness of the researchers seems more related to the amount of active spots than to their functional relevance (the 'Harlequin' syndrome). Sometimes, hazardous associations between single activation loci and highly cognitive and complex functions (e.g. consciousness) are jauntily suggested. Moreover, in some cases, brain physiology is modeled only on the basis of fMRI results, without taking into account anatomical properties, cytoarchitecture, hodology, comparative studies.

Some funding programs are aware of that and in this line are, for example, the guidelines of the James McDonnell Foundation: "*Proposals proposing to use functional imaging to identify the "neural correlates" of cognitive or behavioral tasks (for example, mapping the parts of the brain that "light up" when different groups of subjects play chess, solve physics problems, or choose apples over oranges) are not funded through this program. In general, JSMF and its expert advisors have taken an unfavorable view of projects attempting too wide a leap in a single bound. Functional imaging studies using poorly characterized tasks as proxies for complex behavioral issues involving empathy, moral judgments, or social decision-making are generally not appropriate responses to this call for proposals*".

We all are aware, however, that brain imaging is causing a true revolution in Neuroscience. When properly used, it allows the investigation of scientific problems that until a decade ago were of exclusive pertinence of disciplines like linguistics or even philosophy. In addition, brain imaging is becoming more and more important in clinical applications too. In this commentary I would like to stimulate a scientific discussion focused on two main points: (i) the necessity to deal with single subjects and (ii) the intensity vs. probabilistic analysis issue.

The single subject approach

At their origins, brain imaging techniques had a very poor signal-to-noise ratio. Thus, the study of single subjects was conceivable mainly for clinical purposes. Moreover, the positron emission tomography was, at its beginning, quite invasive because of the significant amount of radioactive tracer injected at every run. It was thus inapplicable to replicate the same tasks in the same subjects to give the analysis enough statistical power. In addition, scientific results are in general aiming at describing phenomena valid for populations and not for single individuals. Thus, group experiment became the standard. For these exigencies, some statistical analysis tools rapidly reached an almost completely monopolistic position. Among them, Statistic Parametric Mapping

E-mail address: fdl@unife.it.
Available online on ScienceDirect (www.sciencedirect.com).

1053-8119/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2007.02.038

(SPM, see for fMRI, Friston et al., 1994) was the most used one. The diffusion of SPM, together with its undisputable merits, progressively habituated neuroscientists to forget individual anatomy. Subjects were averaged together, strongly smoothed and adapted, as average brains, to the Talairach coordinate system (Talairach and Tournoux, 1988). After the analysis, numbers describing *x*,*y*,*z* coordinates of the activated clusters were generated and clusters were labeled according to Brodmann areas (BA), by looking at the pages of the atlas. This despite the frequent incongruence between the various section plans: I remember the long discussions in our group to attribute Brodmann labels to areas which were differently identified in horizontal and coronal slices.

Now the situation has changed and fMRI scanners of the last generation are really sensitive instruments. Moreover, with respect to PET, fMRI is not invasive at all. It becomes therefore more and more possible to organize studies by dealing with single subjects. In my view, whenever is possible we should shift from groups to individuals. Several papers already investigated the reliability of fMRI over time by scanning the same subjects after hours (Havel et al., 2006, Harrington et al., 2006) or even months (Aron et al., 2006). Their conclusion is that areas primarily expected are constantly activated. The main source of variability comes indeed from the surroundings of the activated foci, or from high-order cortical centers which are typically much more variable in their pattern of activation. We should be aware, however, that the repetition of the experiment in the same subject is not suitable in one-shot experiments, where the study of learning effects is the main goal of the investigation.

To deal with single subjects, however, we should take into account three main points. *The first one* is that individual anatomical variability is so high that it is impossible to ignore neuroanatomical landmarks. The capability to identify the main cortical sulci becomes therefore essential to interpret the functional data. It is possible to argue that the sulcal pattern and the cytoarchitectonical one are distinct things. In fact, the idea that the sulci mark also the borders of cytoarchitectonical territories, is just a rough approximation (see Amunts et al., 1999). In some cases, however, this argument is used to hide a profound ignorance by saying: "considering that sulcal anatomy is not fully reliable, let's ignore individual anatomy at all and prefer standard brains which, when interrogated, always give an answer in terms of Brodmann labels". Conversely, Devlin and Poldrack perfectly underline the necessity to refer to the brain regions by using the name of its 'streets' and 'squares' in place of using GPS-like coordinates. Also in this case, however, a problem arises. Cortical gyri, and sometimes even Brodmann areas, are too wide and include in some cases several cytoarchitectonically distinct areas. A comparison between the human and the monkey brains strongly stresses this aspect. For example, almost everybody agrees today on the fact that the premotor cortex of the macaque monkey (BA6) is indeed formed by a mosaic of areas, each of them being characterized by its own somatotopy and, more importantly, by peculiar sensorimotor properties. We don't see any reason to believe that the human brain is less parceled than the monkey one. However, neuroimaging studies still continue to refer to precentral gyrus as BA6 (in some cases we just add the prefixes 'dorsal' or 'ventral'). The same consideration applies to the parietal lobe, which in humans was subdivided by Brodmann in four areas (apart from SI, BA5, BA7, BA39 and BA40) while in the macaque, it is formed by at least the double of cytoarchitectonically distinct areas, each of them dramatically different from the others in terms of functional properties. How could we afford (and possibly solve) this problem? One possibility could be to intensify the neuroanatomical study of the human brain also by taking into account new cytoarchitectonical and receptors data. Despite the fact that this kind of approach requires hard work and big enthusiasm, and produces few papers comparing to other neuroscience fields, some groups are already at work (consider, e.g., the seminal work by Zilles and Amunts groups or see Mazziotta et al., 2001). Some important results have already been achieved: For example it is now clear that BA 4, the primary motor cortex, has lost its monolithic structure being indeed subdivided into, at least, two different areas (Geyer et al., 1996).

The second point is related to the necessity (in my view) to formulate *a priori* hypotheses before designing the experiments. In other terms, in place of just looking at which areas become active after a given contrast, I consider preferable to investigate whether a given area (or a set of areas) that *a priori* we postulate as possibly involved in a given task, become active and at which extent. In other words, FMRI can be a very powerful tool only if it is used as a microscope. On the contrary, when it is used as a telescope, some very bizarre results arise. Indeed, we all aim at finding a possible explanation for all the activation foci arising from the analysis. Thus, driven by the constellation of foci that sometimes appear, we direct our interpretative efforts also to relatively unknown areas. To do this we consult databases and repositories and very often we forget that each activation strongly depends upon both the experimental paradigm and the explored contrast. Unfortunately, we are more and more forced to neglect the experimental details because of the impressive amount of data published everyday and we associate a given Talairach or MNI (Montreal Neurological Institute) coordinate to a specific brain function (and not to a specific experimental task). This is dangerous and should be avoided as much as possible by focusing the discussion on those parts of the brain considered by the *a priori* hypotheses. This is exactly what one does in monkey electrophysiology, forced by the technique which is 'microscopic' in nature.

The third point concerns the statistical analysis of single subject data. If we accept the idea that it would be better to functionally localize the activation in each subject by using neuroanatomical landmarks and then to look at how many subjects share a similar activation pattern, we need first to know what is significant and what is random at the single subject level. SPM or other statistical packages can obviously do this. I believe, however, that the delayed replication of the experiment in the same set of subjects, would better inform us about the reliability of the achieved results. This could be, in my view, a shared standard which should not influence too much the duration and the complexity of the experiment (particularly if one considers that the amount of time we dedicate to the scanning is almost nothing when compared to that used for the analysis). If, for specific purposes (such as the necessity to detect very weak signals) the group analysis is preferable, again I would keep in mind the anatomical constraints by selecting subjects which share a similar sulcal pattern in the region target for the study and to align (and warp) brains not on the standard commissural plane, but on the region of the brain we are interested in. For example, it is known that Broca's region has at least three main morphological variants (see Ono et al., 1990). It would then be possible to group subjects sharing a similar morphology driven by the idea that, very often, morphology and physiology go together. Some attempts in this direction have already been successfully proven (see Stark and Okado, 2003). It is evident that, particularly in this case, the necessity to make a priori hypotheses on a given area is mandatory.

The analysis of the intensity of the BOLD signal

SPM and other statistical tools provide results in terms of probability. Sometimes I have the impression that this evidence is neglected, particularly when one considers the likelihood as equivalent to the strength of activation. One cannot compare the results of two independent contrasts (or even of different areas within the same contrast) by looking at how much the activation A is more probable than B, also because factors different from the activation strength strongly influence the significance of the results (e.g. the spatial concentration of the activation). Therefore, statistical results arising from the average of multiple subjects cannot discriminate between wide hills and tall peaks. Moreover, a very constant but weak activation could be in principle more significant than a more variable, but always stronger one. From this consideration it derives, in my view, that while looking at fMRI results we cannot ignore the intensity of the response. Again, as for the argument raised by Devlin and Poldrack on this issue, this should be considered a trivial sentence. Evidence coming from several papers (including those from myself) and from data presentations at scientific meetings, says that this is not the case and that this argument would require some additional attention. In the more optimistic situations the authors analyze the data in terms of regions of interest or provide the signal intensity data just for the peak of those activations they consider more relevant for their purposes. By this way, a lot of data and information are lost. Why don't we associate to every statistical map an intensity map depicting the amount of signal change around the brain? I would like to take this opportunity to encourage people to work on this issue, also by solving problems like normalization, because I believe that without displaying data on the intensity of the response, neuroimaging paper could be seen just like those upsetting works, sometimes appearing in literature, showing results from analysis of variance without displaying means and standard deviations.

In conclusion, brain imaging is a powerful and useful technique that, if correctly used, could dramatically increase our knowledge on the brain at work. I fully agree with Devlin and Poldrack conclusions about the necessity to deal with neuroanatomy in a serious way. By this commentary I would like to suggest that it is time now to change our mental habit and see fMRI and PET more as measuring techniques than estimating ones.

Acknowledgments

This work is supported by EC and Italian MIUR funds to LF.

References

Amunts, K., Schleicher, A., Burgel, U., Mohlberg, H., Uylings, H.B.M., Zilles, K., 1999. Broca's region revisited: cytoarchitecture and intersubject variability. J. Comp. Neurol. 412, 319–341.

Aron, A.R., Gluck, M.A., Poldrack, R.A., 2006. Long-term test–retest reliability of functional MRI in a classification learning task. NeuroImage 29, 1000–1006.

Friston, K.J., Jezzard, P., Turner, R., 1994. Analysis of functional MRI time series. Hum. Brain Mapp. 1, 153–171.

Geyer, S., Ledberg, A., Schleicher, A., Kinomura, S., Schormann, T., Burgel, U., Klingberg, T., Larsson, J., Zilles, K., Roland, P.E., 1996. Two different areas within the primary motor cortex of man. Nature 382, 805–807.

Harrington, G.S., Buonocore, M.H., Farias, S.T., 2006. Intrasubject reproducibility of functional MR imaging activation in language tasks. Am. J. Neuroradiol. 27, 938–944.

Havel, P., Braun, B., Rau, S., et al., 2006. Reproducibility of activation in four motor paradigms—An fMRI study. J. Neurol. 253, 471–476.

Mazziotta, J., Toga, A., Evans, A., Fox, P., Lancaster, J., Zilles, K., Woods, R., Paus, T., Simpson, G., Pike, B., Holmes, C., Collins, L., Thompson, P., MacDonald, D., Iacoboni, M., Schormann, T., Amunts, K., Palomero-Gallagher, N., Geyer, S., Parsons, L., Narr, K., Kabani, N., Le Goualher, G., Boomsma, D., Cannon, T., Kawashima, R., Mazoyer, B., 2001. A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM). Philos. Trans. R. Soc. Lond., B Biol. Sci. 356 (1412), 1293–1322 Aug 29.

Ono, M., Kubick, S., Abernathey, C.D., 1990. Atlas of the Cerebral Sulci. Thieme, New York.

Stark, C.E., Okado, Y., 2003. Making memories without trying: medial temporal lobe activity associated with incidental memory formation during recognition. J. Neurosci. 23, 6748–6753.

Talairach, J., Tournoux, P., 1988. Co-Planar Stereotaxic Atlas of the Human Brain. Thieme, New York.