fMRI connectivity, meaning and empiricism Comments on: Roebroeck et al. The identification of interacting networks in the brain using fMRI: model selection, causality and deconvolution.

Olivier David

fMRI connectivity, meaning and empiricism

Comments on: Roebroeck at al. The identification of interacting networks in the brain using fMRI: model selection, causality and deconvolution

OLIVIER DAVID

Inserm, U836, Grenoble Institut des Neurosciences, Grenoble, France

Université Joseph Fourier, Grenoble, France

Olivier David
Grenoble Institut des Neurosciences – Inserm U836
CHU Grenoble – Bât EJ Safra – BP 170
38042 Grenoble Cedex 9, France
Tel: +33 456520586
Fax: +33 456520598
Email odavid@ujf-grenoble.fr

September 2009

Roebroeck et al. (this issue) raise fundamental questions regarding the identification of functional networks using fMRI. Their nice methodological review has been triggered by our recent work (David et al., 2008) and its primer (Friston, 2009). I am delighted to have the possibility here to answer to some specific comments made by Roebroeck et al. on (David et al., 2008).

General thoughts after reading Roebroeck et al.

The concept of causality, and in particular the elucidation of cause-effect relationships among variables or events, has been a matter of concerns for many centuries. Two fundamental questions must be considered: (i) how to gather empirical evidence of causal laws; (ii) how to make inference on causal information about an observed phenomenon. Scientific revolutions regarding those aspects have been made during the Renaissance through the development of systematic experiments, and in the late 20th century through the mathematisation of slippery causal concepts (see (Pearl, 2000) for a complete overview). In integrative neuroscience, empirical evidence of causal laws can be obtained from lesion studies, for instance, and is now well accepted. We are facing much more profound difficulties regarding the development of robust inference of causality from neuroimaging data. Roebroeck et al. (this issue) discuss this issue for fMRI. Arguing that temporal resolution is excellent in electrophysiology, intracerebral local field potentials are sometimes perceived as ground truth to speculate on fMRI connectivity results. However, causality inference on electrophysiological data may also be difficult to achieve and, as in fMRI, there is no consensus on optimal approaches.

As perfectly reviewed in Roebroeck et al. (this issue), causal inferences in fMRI can be obtained using different models, from which have emerged Granger causality analysis (GCA) (Goebel et al., 2003) and dynamic causal modelling (DCM) (Friston et al., 2003). From

Roebroeck et al. (this issue) and Friston's comment (this issue), one will understand the main differences. Here I just want to select one technical point:

"[…] in DCM the state variables are given a definite physical interpretation within a generative model of the data. For every selected region a single state variable represents the neuronal or synaptic activity of a local population of neurons and (in DCM for BOLD fMRI) four or five more (Stephan et al., 2007) represent hemodynamic quantities such as capillary blood volume, blood flow and deoxy-hemoglobin content. […] In contrast, in LSM/GCM the state variables may or may not have a definite physical interpretation, depending on the particular representation chosen." (Roebroeck et al., this issue)

In other words, despite the fact that both approaches rely on statistical inferences, it is much easier to interpret causality results from DCM in the standard framework of neuronal networks, despite obvious current limitations on the neuronal and hemodynamic models. This is because DCM connectivity parameters grossly resemble synaptic efficacies of conductance-based neural models. In contrast, GCA of fMRI time series is not sensitive to lack of biophysical knowledge, as DCM can be, but is then entirely dependent on the acquisition modality and associated biases.

To my opinion, this point is very important because it has profound conceptual implications. As a postulate, any cognitive neuroscientist is basically interested in dynamic brain processes that might code for some kind of information. In a neurocentric interpretation of brain function, this means that meaningful inferences should be operated on variables that approximate as best as possible the "neuronal activity", a nebulous concept empirically approached through experimental techniques that give quite different interpretations of what a brain process is. It is evident that models of brain function have been influenced by available experimental techniques, and will evolve along with technological advances. Only biophysical modelling, such as the one proposed in DCM or other generative frameworks, that tries to correct for experimental biases will ensure to stick to the core of biological processes that are the true events of interest.

The question remains as to whether standard BOLD-fMRI is an experimental technique that will allow to estimate neural processes with sufficient accuracy for robust causal inferences, whatever the model used for such purpose. In other words, do fMRI data contain sufficient information on mechanisms of interaction of neural populations? This is far from being clear in the current status of research. In addition to neurally-triggered metabolic events that result in BOLD effect, as mentioned by Roebroeck et al. (this issue), fMRI is indeed potentially very sensitive to unspecific vascular effects, such as differences in vascular tonus, in heart rate variability or in other regulation mechanisms (Gray et al., 2009). These exotic phenomena are not usually well modelled in standard fMRI analyses and may act as dramatic confounds. Estimating causality from standard fMRI cognitive experiments may then be an ill-posed problem because of lack of sufficient experimental control. Acquiring datasets for such a purpose may however be achievable.

Elaborating on important conceptual differences between DCM and GCA, such as determinism and stochasticity, (Roebroeck et al., this issue) and (Friston, this issue) provide very interesting insights on how inferences on causation depend on underlying mathematical assumptions. However it may remain to be evaluated whether such mathematical distinctions are fundamental in relation to the experimental limitations of fMRI regarding the investigation of neural interactions. Modellers are facing here an important difficulty: the lack of available fMRI datasets clearly designed to induce very specific (in time and space) changes of brain connectivity in a highly controlled fashion. Given the complexity of brain processes, achieving, during fMRI scanning, reproducible and robust modulation of specific pathways that solicits the same mechanisms as those involved in human cognition is certainly not an easy task, but is an urgent need. As experimentalist and methodologist in fMRI and electrophysiology, I perceive a gap between what the more advanced models of functional connectivity claim and what experimentalists take for granted. To strengthen findings obtained from connectivity analyses in fMRI, it is evident that causal inferences derived from

neuroimaging data must go beyond modality specific findings. Multimodal data and biophysical models have an essential role to play in that context (see for instance (Valdes-Sosa et al., 2009)). We have now to truly assess whether fMRI alone is a technique that will allow us to achieve such goal with sufficient robustness and adequacy to fundamental questioning on brain function.

But finally, do we really need to reconstruct hidden neuronal states to assess meaningful inferences on brain processes? It depends on what is considered as ontological. Ontology in neuroscience is usually formalised by the neural code, which was well summarised in (Friston, 1997) as:

“A neural code is used here to mean a measurement or metric of neuronal activity that could participate in teleologically meaningful transactions among different parts of the brain. […] a code or measure must necessarily show some dependency when assessed in two interacting neuronal populations or brain areas. The problem of identifying possible codes then reduces to establishing which sorts of measures are mutually predictive or statistically dependent when measured in two parts of the brain”.

So the neural code is implicitly embedded in functional connectivity measures. At the extreme, given some data modality, as long as connectivity inferences are internally consistent over sessions and subjects, ontology is respected and models of how information is coded by the brain can be developed. Now, implicitly relying entirely on how data have been measured is certainly not very appealing to develop concepts on brain function. As in quantum physics, all is about the definition of reality and how we interfere with it. Good integration of modelling and experimental works is the future of integrative neuroscience.

Specific responses to Roebroeck et al.

The experiment developed in (David et al., 2008) was motivated to guide the interpretation of BOLD activations obtained in EEG/fMRI studies in patients suffering from epilepsy that we also perform in our lab. It thus started as a standard EEG/fMRI study on a genetic rat model of absence epilepsy to validate what can be done in EEG/fMRI using known circuitry and invasive recordings. During the data analyses, we found that the hemodynamics in the presumed epileptic focus (Polack et al., 2007; Polack et al., 2008) was very different from the one in other regions. We thus decided to see whether DCM was able to handle that, took GCA to further validate fMRI connectivity, and simply reported our findings. As I was strongly involved during more than two years in every part of the study (including MR sequence programming, animal experimentation, data acquisition and modelling), I know well what can be said, and what cannot, on this dataset. Thus, although I largely agree with Roebroeck et al. (this issue) on many points, I need to comment on some specific statements made by these authors to justify again some of the options taken in (David et al., 2008) and to correct some misinterpretations. Although data presented in (David et al., 2008) are not perfect, they may constitute one of the best available dataset to challenge causal inferences in fMRI. I'll be happy to share them on request.

Introduction

"These authors and a related commentary (Friston, 2009) concluded that: i) The concepts of temporal precedence and G-causality should not be used in fMRI connectivity analysis, and ii) Explicit biophysically motivated models, such as DCM, model true causality in fMRI data, because they account for the hemodynamic processes that intervene between neural activity and fMRI signals."

The first point made by Roebroeck et al. is *not* the conclusion of (David et al., 2008). Only the second point is valid. Note that biophysically motivated models can be connected to GCA as

nicely presented by Roebroeck et al. (this issue) using the notion of *observation equation*. Our actual conclusion was: *"one must minimise spurious interactions due to hemodynamic variability between brain regions using explicit or implicit (such as in DCM) deconvolution of hemodynamic effects in fMRI time series. Otherwise, directed functional connectivity results should be taken cautiously, particularly if one cannot demonstrate that hemodynamic properties are the same in every region analysed."* This does not exclude the use of G-causality.

Model selection

"Only 3 of these structures were selected by the authors as the crucial nodes in the network that generates and sustains seizure activity, and thus worthy of further analysis: S1BF, thalamus and striatum. One cannot stop but wonder whether such a greatly simplified structural model is a justifiable decision given both the rich data-set at hand and indications in the existing literature that generation and maintenance of seizure activity in the employed rat model involves other brain regions, such as frontoparietal cortex (e.g. Danober et al., 1998). It would be interesting to see whether a preliminary structural model selection step using a technique like GCM (on a small part of the data - not to be reused) would lead to a better justified set of selected regions."

The choice of a limited number of regions included in the models was perfectly justified according to the objective of the study, which was to assess clearly whether causality measures in fMRI would be able to minimise the effects of hemodynamic variability on inferred oriented connections. In any case we did not want to provide *the* network of the rat model under study (Genetic Absence Epilepsy Rat from Strasbourg – GAERS (Danober et al., 1998)), simply because there was no point in doing that here. The main reason was that we were not in a position to validate connectivity results that would have been obtained, using GCA as an exploratory step or not. Simply because: (i) There was a certain degree of discrepancy between known electrophysiological recordings and fMRI activations. For instance, some cortical regions, such as the primary motor cortex just anterior to the

somatosensory cortex, showing spike-and-wave discharges (Polack et al., 2007) were not activated in fMRI at the same location. Also some fMRI activated structures, such as the cerebellum, have never been characterised electrophysiologically in the GAERS. (ii) All activated structures at the group level did not show robust activations at the animal level. Because we wanted to minimise the effect of noise on connectivity assessment, we decided to focus on three regions (S1, thalamus, striatum) only that showed excellent signal to noise ratio, that were the best characterised in electrophysiology (Polack et al., 2007; Slaght et al., 2004) and from which the principal anatomical connections were best known. So model selection procedures can be perceived as a philosophical choice. Either it can be systematically driven by exploratory techniques, which rely on specific statistical assumptions as nicely review by Roebroeck et al. (this issue), or it can be more biologically grounded or designed according to specific assumptions to be tested, as what we have done in our study. Both approaches have strengths and shortcomings and I think the choice on how to proceed should be left to the investigators.

Result interpretation

"The GCA analysis after deconvolution in particular is very convincingly in accordance with the gold-standard iEEG analyses (David et al., 2008, their Figure 4 lower right and Figure 7), strongly supporting the value of stochastic dynamical models and G-causality in brain connectivity analysis. In contrast, the final result of DCM analysis of the same data shows less correspondence with the gold standard, not identifying the direct influence of S1BF on the thalamus (David et al., 2008, their Figure 5D and Figure 7). The differences in successful capture of the direct and indirect influence, after deconvolution, are likely due to the difference between a deterministic and stochastic dynamical model, since the observation model was effectively equated. […] In short, however, the results in David et al. show that the stochastic dynamics model of GCM potentially outperforms the deterministic dynamics model in DCM in a confirmatory analysis when both are given the same observation model."

The interpretation of our results made by Roebroeck et al. to prove hypothetic empirical evidence that stochastic GCA outperforms deterministic DCM goes against what we discussed in the manuscript. There is no reason for us to change our point of view. First, because a detailed analysis of our results does not allow us to agree with Roebroeck et al. Indeed, the GCA result (Fig 4) they comment upon was obtained only at the group level, but GCA results reported in (David) were not extremely consistent between animals (maybe because we used the simplest linear version of GCA). In contrast, DCM results were much more reproducible between animals (Fig 5B). In fact, the model shown in Fig 5D emerged at the group level because of Rat 2 which had quite strong evidence for this model. If one removes this "outlier", then models 1, 4 and 5 are almost equivalent, and basically this pool of models compares exactly to the electrophysiological validation shown in Fig 7. Second, a connection which is not kept in a DCM does not necessarily means that it does not exist, but that it is not necessary to explain the data, in contrast to the pairwise GCA which essentially concludes whether there are directed transfers of information.

Hemodynamic deconvolution

"We have suggested that, at the very least, modulation of G-causality between fMRI time-series by experimental context (e.g. a higher level of G-causality during attention to colour than during attention to motion) should be sought to give credibility to these analyses (Roebroeck et al., 2005). It is unfortunate that were David et al. applied GCA to original fMRI time series, they did not take note of these recommendations and did not investigate the modulation of G-causality between ictal and inter-ictal states."

We were perfectly aware of these recommendations. However, they were not applicable in our case because the hemodynamics of ictal and interictal states almost never reached a stable state in the cortical focus S1BF (Figure). Although these states were electrophysiologically well separated, they were not hemodynamically because of strong temporal autocorrelation.

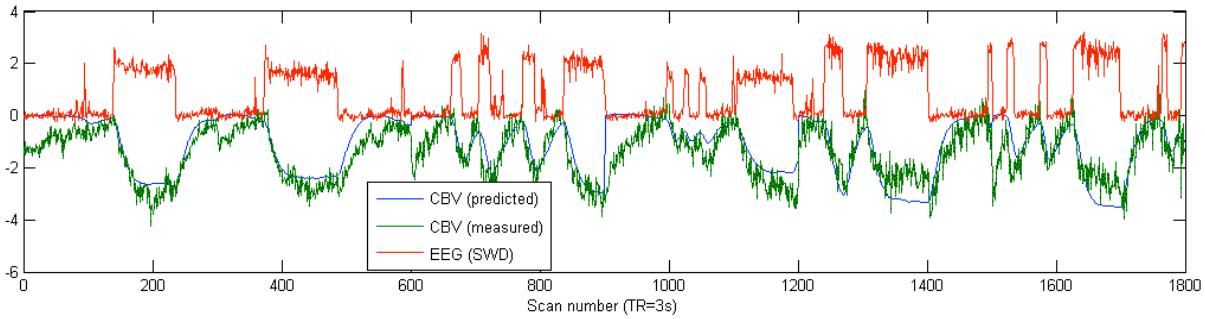

Figure: Concatenated time series of six fMRI sessions of 15 min each, in a rat. Green curve shows CBV-weighted signals in S1BF. Red curve is EEG power and defines seizures. Blue curve is the output of the convolution of EEG power with a hemodynamic response function.

"The decision to use 4-20Hz EEG power as input might well be justified. It characterizes the overall EEG signal power increase that accompanies seizure episodes and might be a good measure of increased neuronal firing and synaptic activity that demands increased metabolism. However, the precise coupling of hemodynamics and local metabolism with neuronal and synaptic activity is complex and partially unknown (Logothetis et al., 2001; Niessing et al., 2005)."

So far, the presence of multi-level dynamic processes has never been detected in recordings of GAERS activity. Our choice to use 4-20 Hz to detect spike-and-wave discharges was thus perfectly adapted to capture their principal mode using the three first components of these 5-7 Hz nonlinear oscillations. In addition, one should not forget that these were EEG signals recorded during EPI scanning, thus contaminated by high frequency noise, which precluded using high frequency EEG components.

However, as well recalled by Roebroeck et al. (this issue), the precise coupling of hemodynamics and local metabolism with neuronal and synaptic activity is partially unknown. I would go further and say that the coupling between EEG and neuronal activity is also unknown. To illustrate this point, I refer to (Polack et al., 2007; Slaght et al., 2004) where intracellular recordings were obtained in the striatum and in the different layers of the S1 during experiments performed in the very same model as the one used in our EEG/fMRI study. From these studies and ours, it is clear that an increase of EEG power was not

systematically correlated to an increase of neuronal firing and synaptic activity, and that the explanation of the sign of cerebral brain volume changes still remains a mystery.

References

Danober, L., Deransart, C., Depaulis, A., Vergnes, M., Marescaux, C., 1998. Pathophysiological mechanisms of genetic absence epilepsy in the rat. Prog.Neurobiol. 55, 27-57.

David, O., Guillemain, I., Saillet, S., Reyt, S., Deransart, C., Segebarth, C., Depaulis, A., 2008. Identifying neural drivers with functional MRI: an electrophysiological validation. PLoS Biol 6, 2683-2697.

Friston, K., 2009. Causal modelling and brain connectivity in functional magnetic resonance imaging. PLoS Biol 7, e33.

Friston, K., this issue. Dynamic causal modeling and Granger causality. Neuroimage.

Friston, K.J., 1997. Another neural code? Neuroimage 5, 213-220.

Friston, K.J., Harrison, L., Penny, W., 2003. Dynamic causal modelling. Neuroimage. 19, 1273-1302.

Goebel, R., Roebroeck, A., Kim, D.S., Formisano, E., 2003. Investigating directed cortical interactions in time-resolved fMRI data using vector autoregressive modeling and Granger causality mapping. Magn Reson.Imaging 21, 1251-1261.

Gray, M.A., Minati, L., Harrison, N.A., Gianaros, P.J., Napadow, V., Critchley, H.D., 2009. Physiological recordings: basic concepts and implementation during functional magnetic resonance imaging. Neuroimage 47, 1105-1115.

Pearl, J., 2000. Causality: Models, reasoning, and inference. Cambridge University Press, New York, NY.

Polack, P.O., Guillemain, I., Hu, E., Deransart, C., Depaulis, A., Charpier, S., 2007. Deep layer somatosensory cortical neurons initiate spike-and-wave discharges in a genetic model of absence seizures. J.Neurosci. 27, 6590-6599.

Polack, P.O., Mahon, S., Chavez, M., Charpier, S., 2008. Inactivation of the somatosensory cortex prevents paroxysmal oscillations in cortical and related thalamic neurones in a genetic model of absence epilepsy. Cereb.Cortex.

Roebroeck, A., Formisano, E., Goebel, R., this issue. The identification of interacting networks in the brain using fMRI: model selection, causality and deconvolution. Neuroimage.

Slaght, S.J., Paz, T., Chavez, M., Deniau, J.M., Mahon, S., Charpier, S., 2004. On the activity of the corticostriatal networks during spike-and-wave discharges in a genetic model of absence epilepsy. J Neurosci 24, 6816-6825.

Valdes-Sosa, P.A., Sanchez-Bornot, J.M., Sotero, R.C., Iturria-Medina, Y., Aleman-Gomez, Y., Bosch-Bayard, J., Carbonell, F., Ozaki, T., 2009. Model driven EEG/fMRI fusion of brain oscillations. Hum Brain Mapp 30, 2701-2721.