ELSEVIER

NeuroImage

www.elsevier.com/locate/ynimg
NeuroImage 23 (2004) 21–24

Discussion

Generative and recognition models for neuroanatomy☆

K.J. Friston* and J. Ashburner

The Wellcome Department of Imaging Neuroscience, University of London, Institute of Neurology, London, UK WC1N 3BG

Received 12 February 2004; revised 13 April 2004; accepted 21 April 2004
Available online 26 June 2004

The issue

As with previous critiques of voxel-based morphometry (e.g., see Bookstein, 2001, Mehta et al., 2003), Christos Davatzikos (2004) reprises issues that have engaged the functional neuroimaging community for many years. In this instance, the issue is the distinction between multivariate and mass-univariate analyses of imaging data. Put simply, Davatzikos is pointing out that pathology can be expressed, anatomically, in a distributed and complicated fashion over the brain. Critically, its expression in one part of the brain may depend on its expression elsewhere. Characterizing these interregional dependencies requires a multivariate model (e.g., see Ashburner et al., 1998; Bookstein, 1984) of how pathology causes anatomical changes. In the former example, canonical variates analysis (CVA[1]) was used to assess gender differences using deformation-based morphometry. Deformation-based morphometry represents an analysis of the deformation fields that spatially normalize images. However, differences in brain anatomy may not be completely encoded by these deformations; local structural differences may persist after spatial normalization. Voxel-based morphometry (VBM) was introduced to characterize these differences.

Voxel-based morphometry

Voxel-based morphometry is the application of statistical parametric mapping to (spatially normalized, scalar) images that index some aspect of local brain structure, for example, grey matter density following segmentation or compression maps based on the Jacobian of deformation fields. As such, VBM uses a mass-univariate approach. VBM was proposed and designed for the analysis of regionally specific differences in structural indices. The nature of these indices and ensuing interpretation depends upon the preprocessing of images (e.g., segmentation) before VBM.

By definition, regionally specific effects do not depend on changes elsewhere. This is a fundament of mass-univariate approaches. Consequently, VBM is not used to characterize interregional dependencies. VBM is a simple procedure that enables classical inferences about the regionally specific effects of experimental factors on some structural measure. These effects are tested for after discounting gross anatomical differences that are removed by spatial normalization.

Because gross differences have been removed, VBM is not a surrogate for classical volumetric analysis of well-defined anatomical structures or lesions (see Mehta et al., 2003). Furthermore, because VBM is voxel-based, differences in the shape or form of large structures are discounted by the requisite normalization. This means VBM will never replace careful applications of shape analysis and modeling to computational neuroanatomy (e.g., see Davies et al., 2002; Pitiot et al., 2002). In short, VBM represents an established and effective complement to shape and volumetric analyses (and deformation-based morphometry) that is now used widely in many basic and clinical contexts (see Fig. 1).

Distributed versus dependent

It is important to distinguish between the presence of distributed effects over the brain and dependencies among these effects. It is perfectly possible for VBM to detect a distributed pattern of regionally specific effects. However, VBM is not appropriate for the analysis of statistical dependencies among measures from different regions. The example shown in Fig. 1 (left-hand panel) of Davatzikos's paper is, in fact, a rather bad example to use from his point of view. This is a linear treatment effect that is distributed over voxels 1 and 2. With sufficient sensitivity, VBM would properly characterize the distributed nature of this difference as an effect in both voxels (just because the marginal distributions overlap does not mean there is no significant difference in the distributions).

However, note that within both the normal and patient groups, there is a negative correlation between the measures in voxels 1

☆ A response to "Why voxel-based morphometric analysis should be used with great caution when characterizing group differences," by Christos Davatzikos (doi:10.1016/j.neuroimage.2004.05.010).

* Corresponding author. The Wellcome Department of Imaging Neuroscience, University of London, Institute of Neurology, Queen Square, London, UK WC1N 3BG. Fax: +44-207-813-1445.

E-mail address: k.friston@fil.ion.ucl.ac.uk (K.J. Friston).

[1] CVA is the statistical formulation of partial least squares mentioned by Christos Davatzikos.

Available online on ScienceDirect (www.sciencedirect.com.)

doi:10.1016/j.neuroimage.2004.04.021

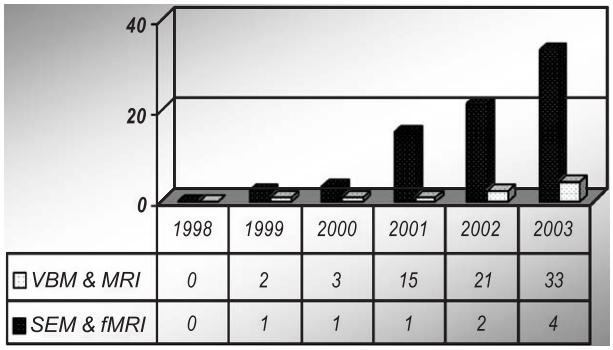


Fig. 1. Instances of PubMed© results for "Voxel-based Morphometry" and "MRI." For comparison with another important method, equivalent search results are provided for "Structural Equation Modeling" and fMRI.

and 2. This means that the probability distribution of the voxel 1 measure depends upon the measure in voxel 2 (see Fig. 2a). These negative correlations could only be assessed with a multivariate (2-voxel) model, for example, CVA. These dependencies are expressed as covariances among the error terms as indicated in Fig. 2b (cf. left and middle panels: to simplify things, only one group is shown, which could also represent case-controlled differences). Identification of distributed spatial modes that discriminate among pathological cohorts using CVA has a long history in neuroimaging (e.g., see Friston et al., 1992). In summary, the central issue is the distinction between multivariate and mass-univariate approaches in accommodating interregional dependencies in structural and functional data. We now focus on this distinction.

Multivariate versus mass-univariate

In functional anatomy, this dichotomy is closely related to the distinction between functional specialization and functional integration in the brain. Most analyses of neuroimaging data are predicated on the specialization or segregation model and are happy to limit their inferences to regionally specific effects using univariate approaches. The alternative is to characterize treatment-related responses in one brain area, in relation to responses elsewhere. This calls for multivariate models that are usually framed in terms of functional or effective connectivity. Multivariate analyses of neuroimaging data can range from the very simple, such as eigenimage (PCA) analysis to complicated nonlinear (i.e., bilinear) models such as dynamic causal modeling. One important distinction between these multivariate approaches is whether they are linear or nonlinear. This speaks to the interesting situation depicted in Fig. 3 of Davatzikos's paper where the separatrix is nonlinear. The distinction between linear and nonlinear models is also illustrated in Fig. 2b (middle and right panels). In the past decade, many nonlinear multivariate models have been explored in the context of functional neuroimaging, including neural networks (Lautrup et al., 1994),

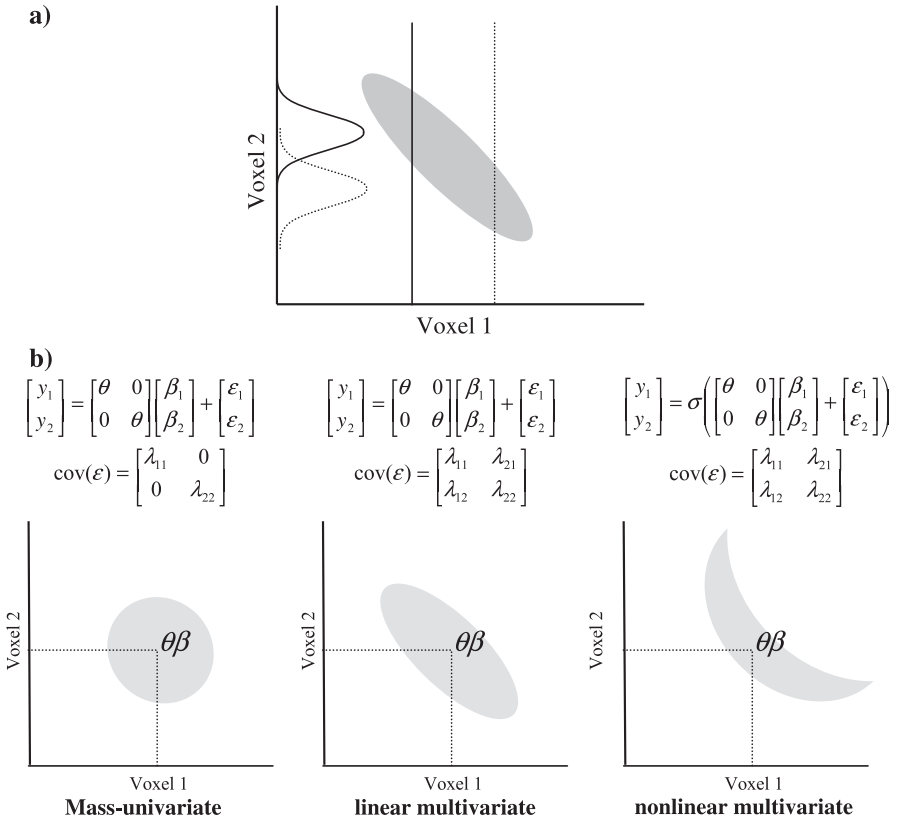
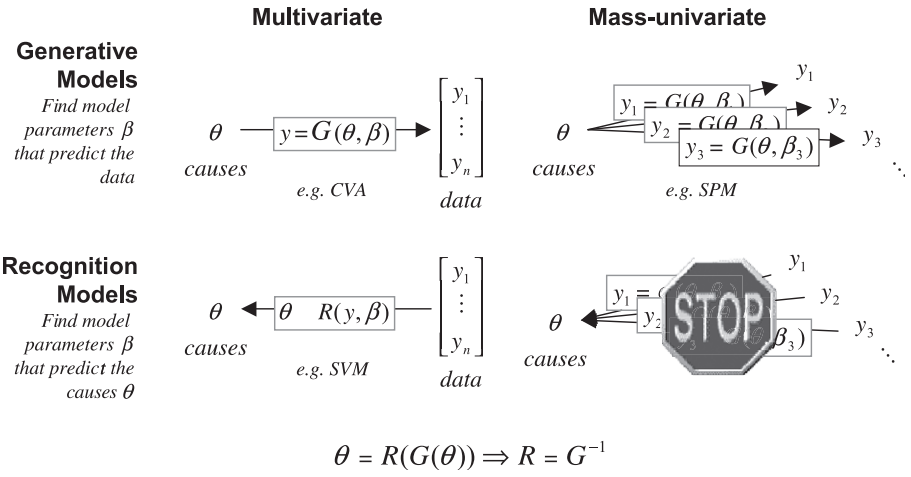


Fig. 2. (a) Schematic highlighting how the distribution of values in voxel 2 depends on the value in voxel 1. (b) Different classes of model.



$$\theta = R(G(\theta)) \Rightarrow R = G^{-1}$$

Fig. 3. Schematic illustrating the relation between generative/recognition models and multivariate/mass-univariate formulations.

nonlinear PCA (Friston et al., 1999), and support vector machines (Cox and Savoy, 2003). The notion of applying these techniques to structural data is a compelling one.

Using VBM to assess interregional dependencies

Usually, interregional dependencies are accommodated by treating the region as a factor in multivariate statistical models and looking for region × treatment interactions. However, there are ways of using univariate VBM to assess these sorts of effects. This entails using the response variable in one region as an explanatory variable for other regions. For example, to test for the negative correlations in Fig. 2b (middle panel), one would simply construct an SPM using the measures in voxel 1 over subjects as a repressor or explanatory variable. To assess significant differences in the regression slopes of voxel 2 on voxel 1 between two groups, one would test for a voxel 1 × treatment interaction. (cf., psychophysiological interactions). Nonlinearities are easily accommodated using polynomial or other expansions of structural measures in the index region. These sorts of analyses are standard practice in functional imaging, which suggests that a careful modeling of structural data with standard univariate techniques may be a useful first step.

Characterization versus classification

Finally, Davatzikos makes an important point that procedures for making statistical inferences about regionally specific effects (i.e., VBM) are not necessarily the best for classifying patients. This is absolutely right. VBM is a research tool that enables people to ask specific questions of their data. It is not a diagnostic or classification device. Although there is a close mathematical relationship between CVA and discriminant function analysis, the objectives of classification procedures based on discriminant function analysis, support vector machines, and related Bayesian decision procedures are different from hypothesis testing. VBM is used to ask well-defined questions about how developmental or disease processes affect anatomy: It facilitates a mechanistic understanding of these processes. It is not concerned with how anatomical differences can be used to predict diagnosis.

Although distinct, the multivariate and classification issues are related. When analyzing data, one is implicitly trying to link the observed data to their causes. From a machine-learning perspective, this link rests upon a model of how causes generate data (generative models) or how data disclose their causes (recognition models). As indicated in Fig. 3, generative and recognition models are the inverse of each other. In short, one can make inferences about the causes of data θ (e.g., embodied in a design matrix) by estimating the parameters β of a generative model, which best predict the data. Alternatively, one can estimate the parameters of a recognition model, which best predict the causes. The former approach is generally used to characterize and make inferences about causes in the context of a model that specifies exactly how those causes are manifested in data. The second is concerned with recognizing the underlying causes, given some new data (the test data), using model parameters β that are learned during training on some old data (the training data).

The key point here is that generative models can be multivariate or mass-univariate (i.e., a model of how pathology is expressed at each brain location). Conversely, recognition generally requires multivariate models. This is because interactions among regional measures (nonlinearities in the recognition model) are precluded in mass-univariate approaches (see Fig. 3). Having said this, in the special case of no spatial dependencies (e.g., the left panel of Fig. 2b), the mass-univariate models can be considered jointly in a recognition model because information from one part of the brain adds linearly with information from another. However, generally, one would use a multivariate model.

Given that both generative models and recognition models are used to make inferences about causes, and one is simply the inverse of the other, why are they fundamentally different? They are different because one cannot necessarily invert the generative model, which we know exists, to obtain a unique recognition model. SVM and related approaches rest on the assumption that the recognition model exists. However, in some cases, it does not. To make this clear, consider the following contrived example. A number of neuropsychiatric disorders are associated with expanded trinucleotide repeats. For example, Fragile X Syndrome is thought to occur when there are more than 200 cytosine–guanine–guanine repeats. Suppose we wanted to test the hypothesis that parahippocampal grey matter density showed an inverted "U" dependency on the number of

repeats (after mean correction) in a cohort of premutation carriers. Here, the generative model would be $y = G(\theta,\beta) = \theta^2\beta + \varepsilon$, and we could perform a simple VBM analysis to test the alternate hypothesis that $\beta < 0$. However, this generative model is not invertible, which means there is no recognition model that could infer the number of repeats, given parahippocampal grey matter density y; i.e., $\theta = R(y,\beta) = \pm\sqrt{(y-\epsilon)/\beta}$. This example highlights the role of generative models in testing mechanistic hypotheses in terms of quantities that control the expression of causes. In contradistinction, recognition models try to infer the causes (e.g., classify) given the data. In these models, the parameters have no physical meaning unless the corresponding generative model is invertible. This distinction is one reason why recognition models are seldom encountered in the scientific literature. In radiology, the converse is true. Indeed, the divide between causal and recognition models is evident at the level of neuroscientific groups (note the complete absence of overlap between the program committees of MICCA 2004 Medical Image Computing and Computer Assisted Intervention (http://miccai.irisa.fr/index2.php) and the Organization of Human Brain Mapping 2004 (http://www.conferences.hu/hbm2004/). In summary, if one wants to classify or recognize underlying causes of data, univariate approaches like VBM are not appropriate. Conversely, if one wants to make inferences about how data are caused, VBM is entirely sufficient.

Summary

In conclusion, there is a clear distinction between multivariate and mass-univariate characterizations of brain imaging data. The application of standard univariate approaches to structural data, namely VBM, has proved extremely successful. Part of this success can be explained by the fact that VBM allows researchers to frame and report their analyses in terms of regionally specific effects that refer directly to structure–function relationships. The main point made by Davatzikos is that complementary multivariate techniques from functional brain imaging and machine learning may also find a useful role and, when used as recognition models, are necessary.

On balance, there is more agreement than disagreement between the commentary and our response. In summary, both papers agree that VBM, on its own, is not a tool for characterizing the spatial interdependencies of subtle distributed effects, and both discuss multivariate tools in this context. Furthermore, both agree that consideration of nonlinear models of dependencies is important. What we have brought to the argument is a clear conceptual distinction between various aims of data analysis that we have framed in terms of generative and recognition models.

References

Ashburner, J., Hutton, C., Frackowiak, R., Johnsrude, I., Price, C., Friston, K., 1998. Identifying global anatomical differences: deformation-based morphometry. Hum. Brain Mapp. 6, 348–357.

Bookstein, F.L., 1984. Tensor biometrics for changes in cranial shape. Ann. Hum. Biol. 11, 413–437.

Bookstein, F.L., 2001. "Voxel-based morphometry" should not be used with imperfectly registered images. NeuroImage 14, 1454–1462.

Cox, D.D., Savoy, R.L., 2003. Functional magnetic resonance imaging (fMRI) "brain reading": detecting and classifying distributed patterns of fMRI activity in human visual cortex. NeuroImage 19, 261–270.

Davatzikos, C., 2004. Why voxel-based morphometric analysis should be used with great caution when characterizing group differences. NeuroImage 23, 17–20.

Davies, R.H., Twining, C.J., Cootes, T.F., Waterton, J.C., Taylor, C.J., 2002. A minimum description length approach to statistical shape modeling. IEEE Trans. Med. Imag. 21, 525–537.

Friston, K.J., Liddle, P.F., Frith, C.D., Hirsch, S.R., Frackowiak, R.S., 1992. The left medial temporal region and schizophrenia. A PET study. Brain 115, 367–382.

Friston, K., Phillips, J., Chawla, D., Büchel, C., 1999. Revealing interactions among brain systems with nonlinear PCA. Hum. Brain Mapp. 8, 92–97.

Lautrup, B., Hanse, L.K., Law, I., Morch, N., Svarer, C., Strother, S.C., 1994. Massive weight sharing: a cure for extremely ill-posed problems. In: Herrmann, H.J., Wolf, D.E., Poppel, E. (Eds.), Supercomputing in Brain Research: From Tomography to Neural Networks. World Scientific, London, pp. 137–148.

Mehta, S., Grabowski, T.J., Trivedi, Y., Damasio, H., 2003. Evaluation of voxel-based morphometry for focal lesion detection in individuals. NeuroImage 20, 1438–1454.

Pitiot, A., Toga, A.W., Thompson, P.M., 2002. Adaptive elastic segmentation of brain MRI via shape-model-guided evolutionary programming. IEEE Trans. Med. Imag. 2, 910–923.