ELSEVIER

NeuroImage

www.elsevier.com/locate/ynimg
NeuroImage 40 (2008) 1429–1435

Comment

Ten simple rules for reporting voxel-based morphometry studies

Gerard R. Ridgway,[a] Susie M.D. Henley,[b] Jonathan D. Rohrer,[b] Rachael I. Scahill,[b] Jason D. Warren,[b] and Nick C. Fox[b,*]

[a]*Centre for Medical Image Computing, University College London, UK*
[b]*Dementia Research Centre, Institute of Neurology, University College London, UK*

Received 14 September 2007; revised 26 December 2007; accepted 3 January 2008
Available online 17 January 2008

Voxel-based morphometry [Ashburner, J. and Friston, K.J., 2000. Voxel-based morphometry—the methods. *NeuroImage* 11(6 Pt 1), 805–821] is a commonly used tool for studying patterns of brain change in development or disease and neuroanatomical correlates of subject characteristics. In performing a VBM study, many methodological options are available; if the study is to be easily interpretable and repeatable, the processing steps and decisions must be clearly described. Similarly, unusual methods and parameter choices should be justified in order to aid readers in judging the importance of such options or in comparing the work with other studies. This editorial suggests core principles that should be followed and information that should be included when reporting a VBM study in order to make it transparent, replicable and useful.
© 2008 Elsevier Inc. All rights reserved.

Keywords: Voxel-based morphometry; computational neuroanatomy; structural MRI; guidelines

Introduction

Voxel-based morphometry (Ashburner and Friston, 2000; Mechelli et al., 2005) is becoming increasingly widely used as a tool to examine patterns of brain change in healthy aging (Good et al., 2001) or neurodegenerative disease (Baron et al., 2001) and neuroanatomical correlates of behavioural or cognitive deficits (Abell et al., 1999). VBM essentially involves voxel-wise statistical analysis of pre-processed structural MR images. Although much of the processing and analysis is automated in software packages such as SPM,[1] many methodological decisions remain, including what template to use for normalisation, what level and type of correction to use and how best to display results. Different approaches, such as VBM using RAVENS maps (Davatzikos et al., 2001), introduce yet more options. It can therefore be difficult to replicate or draw conclusions from VBM studies if the processing steps are not clearly described. Similarly, if unusual methods or parameters are employed without sufficient justification, it can be challenging for readers to judge the potential impact on results or to compare the work with other studies. In light of these issues, this editorial presents a set of recommendations, in the form of ten "rules" accompanied by a checklist, which we hope will be helpful to authors when writing up VBM studies. The rules are intended to outline core principles that should be followed and information that should be included when reporting a VBM study in order to make it transparent, replicable and useful. Since the field is rapidly developing, such rules must not be overly restrictive; therefore in some instances, where a clear protocol cannot be stated, general advice is given in the hope of aiding the reader to follow good practice. As VBM data sets accumulate and alternative procedures and techniques proliferate, we feel that guidelines are crucial for clear scientific communication and further development of the field. Additional motivation for this work came from a related effort in the field of functional brain imaging ("Guidelines for reporting an fMRI study", Poldrack et al., in press).[2]

1. Set out the rationale for your study and describe the data fully

What are the key experimental questions, and why was VBM preferred over other techniques in order to address these questions?

Abbreviations: SPM, statistical parametric mapping; WM, white matter; GM, grey matter; MNI, Montreal Neurological Institute; ICBM, International Consortium for Brain Mapping; FDR, false discovery rate; FWE, family-wise error; FWHM, full-width half-maximum; DCT, discrete cosine transform; SVC, small-volume correction.

* Corresponding author. Dementia Research Centre, Institute of Neurology, University College London, Queen Square, London WC1N 3BG, UK. Fax. +44 20 7676 2066.
E-mail address: N.Fox@dementia.ion.ucl.ac.uk (N.C. Fox).
Available online on ScienceDirect (www.sciencedirect.com).

1053-8119/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2008.01.003

[1] Statistical Parametric Mapping—see: www.fil.ion.ucl.ac.uk/spm/.
[2] See also: http://www.fmrimethods.org.

Prior hypotheses should be stated; either experimental ones or *a priori* anatomical or spatial regions in which effects might be expected (Maguire et al., 2000). This is particularly important if search volumes are restricted when correcting for multiple statistical tests during data analysis (see Rule 5). The study design should be described in enough detail for readers to be confident that subjects have been included appropriately and that important sources of error have been identified, and, where possible, controlled for. Subject inclusion and exclusion criteria should be clearly set out, as well as baseline demographic information (such as age and gender) and any other variables which are relevant to the interpretation of the findings (Scahill et al., 2003). Examples of such variables could include IQ in a study of cognitive function or measures of disease severity or duration in a clinical study. Image acquisition can influence morphometry results (Littmann et al., 2006), and it is therefore essential to report any variations in acquisition such as different scanners, scanner upgrades or pulse sequence changes. The relative timing of data acquisition should be specified, for example, whether MRI and any clinical or behavioural data for each subject were collected on the same day; if not what was the interval? It is also important to specify whether MRI data for different groups were collected in an interleaved fashion or in blocks (which raises the danger that changes in scanner calibration over time could confound effects of interest, Whitwell et al., 2001). Scanner models and locations should be listed for multi-centre studies, and assessment intervals (for MRI and any other data collection) should be made clear for longitudinal studies. If analysing multiple groups (e.g. patients and controls), discuss whether potential confounds, such as age, gender or acquisition differences are balanced between groups. If subjects or scans were excluded from the analysis, this should be stated and justified (see Rule 9).

01 taka Nemoto

2. Explain how the brain segmentations are produced

02 taka Nemoto

The inputs to VBM's statistical analysis are derived from structural MR images using tissue segmentation, spatial normalisation and smoothing. Additional pre-processing is often performed before the main segmentation step, generally using automatic algorithms such as MR bias correction or skull stripping, or manual techniques such as semi-automatic brain segmentation or interactive reorientation. Multiple processes may be combined within unified algorithms, such as that of Ashburner and Friston (2005). The pre-processing steps must be reported in sufficient detail for the methods to be clear and reproducible; as a minimum, this should include the software packages used (with version numbers) and any parameters altered from the default values. For interactive steps, authors should clarify the protocol, for example, whether operators were blind to subject identity. The segmentation method itself should be reported so as to be reproducible, either through clear identification of the software package and description of any defaults modified or via careful description of the algorithm. Some popular segmentation algorithms use registered spatial priors (Ashburner and Friston, 1997), in which case the source of the priors and the means of alignment should be clear. In particular, with SPM2, different methods of iterative segmentation and normalisation have been used, often including iterative regeneration of priors (Good et al., 2001; Douaud et al., 2006); these should be reported in detail—terms such as "optimised VBM using SPM2" are not sufficiently precise. Following segmentation, other image-processing methods can be used to condition the data further. Such techniques include morphological filtering (used in the "clean-up" option of SPM2 and 5), application of Markov Random Field models[3] or interactive editing of segmentations. These approaches tend to be less standardised, so should be reported carefully. The final image-processing step is usually to smooth the segmentations, typically through convolution with a Gaussian kernel, in which case the Full-Width at Half-Maximum (FWHM) should be reported. Since smoothing sensitises the analysis to a particular spatial scale of effect (due to the matched filter theorem; Ashburner and Friston, 2001) some justification of the choice of FWHM would be helpful. Less widely used smoothing techniques, such as anisotropic smoothing (Gerig et al., 1992), should be explained in detail.

03 taka Nemoto

3. Describe the method of inter-subject spatial normalisation

In order to compare different subjects, it is essential to use some kind of registration algorithm to bring the images into at least approximate correspondence. Both the technique used and the reference space to which brains are aligned can impact on the results (Senjem et al., 2005), so clear reporting is crucial. As with the other pre-processing steps (see Rule 2), if a popular software package is used, deviations from the default options should be highlighted. If a non-standard approach is employed, more detail is required, describing the four basic elements of image registration: the spatial transformation model; the objective function, including any regularisation terms or Bayesian priors; the optimisation algorithm; and the interpolation method (Hill et al., 2001). Spatially normalised segmentations may be subsequently "modulated" with the Jacobian determinants from the transformation in order to adjust for the resulting volume changes (Good et al., 2001). This can heavily influence the results and their interpretation (Keller et al., 2004; Mechelli et al., 2005), so authors should state whether or not modulation has been performed and justify this choice. It is important to clearly report the reference space to which brains are being aligned as there are a number of different options available that are defined in quite different ways, ranging from low degree of freedom landmark based reorientation and scaling (Talairach and Tournoux, 1988) to automated registration with greater degrees of freedom, either to a template (e.g. Ashburner and Friston, 1999; Shen and Davatzikos, 2002) or to tissue probability maps (Ashburner and Friston, 2005). Template images or segmentations may be standard, such as the popular MNI or ICBM ones (used in SPM), or may be derived from the subjects themselves (e.g. Good et al., 2001; Kochunov et al., 2001; Joshi et al., 2004; Ashburner, 2007). If a subset of the data are used to generate custom templates or tissue probability maps, then which subjects (e.g. healthy, diseased or a balanced mix), and why, should be clear. Poldrack et al. (in press) further discuss the choice of reference space, with particular focus on the concept of Talairach space and its relation to standard atlases.

4. Make your statistical design transparent

There are two issues here, model specification and contrast testing. When constructing a model, it is important to be clear about which variables are included, and why. In the case of factorial designs, it should be obvious to the reader exactly what the factors were, the levels of each factor, and which interactions between factors were modelled. With estimation methods more

[3] See e.g. http://dbm.neuro.uni-jena.de/vbm/markov-random-fields/.

advanced than ordinary least squares, it may be necessary to report extra information; for example, SPM5 includes non-sphericity options that allow levels of a factor to be dependent or to have different variances. Subject characteristics (Rule 1) should be assessed critically to ensure confounding variables have been included as covariates where appropriate. It is helpful to the reader to indicate why each variable has been modelled, and whether it is a variable of interest (e.g. a psychometric score) or a potentially confounding factor (e.g. age). It may be desirable to adjust for each subject's global brain tissue volume or total intra-cranial volume (Whitwell et al., 2001; Good et al., 2001). Adjustment may be performed either by entering the global values as a covariate, or using them to scale the original voxel values (Kiebel and Holmes, 2007, discuss the differences in the context of PET imaging). Adjusting for global variables can alter findings (Good et al., 2001) and remains a topic of debate in VBM (Mechelli et al., 2005), which motivates both careful planning and thorough reporting. For all covariates, options relating to centring or orthogonalisation should be reported, especially if factor–covariate interactions are modelled. When interrogating the model, the contrasts tested should be described precisely, in terms of the variables involved and their weights. The choice of statistic (*t*-test or *F*-test) should be justified and (for single-tailed *t*-tests) the direction specified. Inclusion of a diagram (e.g. the design matrix) or equation summarising the model and contrasts may be helpful.

04

otaka Nemoto

5. Be clear about the significance of your findings

05

otaka Nemoto

As with other mass-univariate image analysis techniques, a large number of statistical tests are performed in a VBM study. The method used to correct for multiple testing should be both clearly stated and carefully considered—ideally, *a priori*. VBM is often performed on limited numbers of subjects (for example, to investigate rare disorders), when there is a temptation to report uncorrected results due to low statistical power. If this is done, it should be made obvious and it is probably best avoided—alternatives include correction at a less stringent alpha-level, or clear presentation of unthresholded *t*- or effect-maps.[4] Studies have also been published comparing single subjects to larger control groups; the standard parametric statistical framework is poorly suited to such unbalanced designs unless large smoothing kernels are employed (Salmond et al., 2002). Control of the voxel-level family-wise error rate (FWE) using methods based on random field theory requires estimation of the smoothness of the data and depends strongly on the size of the search region. Therefore, interpretation is aided by reporting the estimated FWHM smoothness (not the same as the smoothness applied during pre-processing) and the resel count. In addition, the method used to define the search region (e.g. an explicit mask, or an absolute or relative threshold) should be specified. Cluster-level control of FWE usually assumes stationary smoothness, which is unlikely to be appropriate for VBM, unless special techniques are employed.[5] If it is used, it should be justified, and the cluster-defining threshold must be reported. Permutation-based statistics (Nichols and Holmes, 2002) provide an alternative method to control FWE (based on voxel value, cluster size or cluster mass). These make fewer assumptions but require careful explanation of the statistical design (including any steps for orthogonalising covariates). If sub-volumes of the main search region are analysed (known as small volume correction in SPM), authors should explain how and why these regions of interest were selected. Such regions should ideally be anatomically defined and chosen *a priori* with justification (see also Rule 8). False discovery rate (FDR) correction (Genovese et al., 2002) can follow either parametric or permutation-based statistics, over the whole search region or sub-volumes; these choices mean reporting should be more detailed than a simple statement that FDR was used.

06

otaka Nemoto

6. Present results unambiguously

The type and level of correction should be stated in all figure and table legends, and if the statistical parametric map (SPM) is displayed as orthogonal slices or sections then coordinates should be given. It is helpful to present tables that include statistic values and cluster sizes, as well as coordinates of local maxima. SPMs should be displayed on a template that represents some form of average anatomy, for example, the MNI T1 template often used for normalisation, or ideally, a study-specific mean image. Displaying overlays on a single high-resolution image is misleading: an individual subject is likely to be poorly representative of the group (Devlin and Poldrack, 2007) and implies a higher level of anatomical precision than is possible with smoothed data.[6] A similar caveat applies to the use of anatomical labels. Methods for converting MNI coordinates to Talairach space should be referenced and may be best avoided (Devlin and Poldrack, 2007). Comparison of results can be aided by using the same *t*- or *F*-statistic colour scales across figures. If an SPM is displayed at a threshold lower than that used to locate significant voxels (for example, in order to show small effects or give an impression of the overall distribution of change), this should be made explicit. If single-tailed *t*-tests are focused on (for example, in a study of atrophy where tissue gain would be clinically implausible), it may nevertheless be helpful to report the reverse contrast as it can indicate mis-registration as a potential confound – or even a possible cause – for the main findings.

7. Clarify and justify any non-standard statistical analyses

As a general principle, the less standard the analysis, the more thoroughly it should be explained. Here, we discuss three of the more common examples. Contrast masking may be used to disambiguate multiple possible causes of an effect or to define smaller search regions, in which case authors should clarify not only which contrasts were analysed, which were used for masking and at what threshold, but also the motivation for doing so and their interpretation. If a conjunction of analyses is tested using the minimum of several statistic images, it is crucial to clarify the null hypothesis—global, conjunction, or intermediate (Friston et al., 2005). If data are extracted (e.g. eigenvariates from volumes of interest, peak voxels or cluster summaries) for analysis with other statistical software, this should be explained and justified (see also the rule below).

8. Guard against common pitfalls

Here we discuss a few potential problems with VBM analyses that might be easily overlooked. Firstly, note that while voxel-wise

[4] See http://imaging.mrc-cbu.cam.ac.uk/imaging/UnthresholdedEffect-Maps.

[5] See http://fmri.wfubmc.edu/cms/NS-General.

[6] Consider also the limitations of spatial normalisation discussed in Rule 9.

multiple testing is usually corrected for (see Rule 5), most software packages do nothing to correct for the user's investigation of multiple contrasts—the more conventional multiple-comparison problem (Hochberg and Tamhane, 1987). A simple example of this occurs if two opposite single-tailed *t*-contrasts are analysed: if findings in either contrast could be considered significant but only one is eventually reported, then either this must be noted or the alpha-level or *p*-values should be adjusted. With more complex models, it can be difficult to decide on a suitable correction procedure (Ludbrook, 1991), but if many contrasts have been tested and not presented, this must be noted. A more insidious multiple-comparisons problem can occur if part or all of the VBM analysis is repeated for any reason. The context for this is crucial: for example, different amounts of smoothing (see Rules 2 and 9) may be used to match the filter size to multiple spatial scales of expected effects, whereas it would be misleading to try several FWHM values before reporting only the most appealing results. It is also possible to invalidate correction for voxel-wise multiple tests by extracting sub-regions of the images for further analysis; it is essential that the procedure used to select data is independent of the subsequent analysis (Friston, 1997) and clearly described. Similar caveats apply to the selection of alternative parameters at other pre-processing stages, or the analysis of multiple sub-groups of subjects (e.g. for disease sub-types), unless this is done using independent data sets. It is sometimes necessary to exclude certain subjects or scans (for example, due to artefacts or pre-processing failures); such decisions should ideally be blind to the subjects' identity, and care should be taken to avoid bias or, if this is not possible (e.g. if more severely affected subjects are more likely to be excluded due to poor segmentation), sources of bias should be acknowledged.

9. Recognise the limitations of the technique

Like all image analysis methods, VBM has inherent limitations (Bookstein, 2001). The basic premise of inter-subject spatial normalisation is problematic: different subjects can have different gyral variants with no "true" correspondence between them and information from structural MRI (even manual sulcal labelling) does not necessarily predict underlying cytoarchitectonic borders (Amunts et al., 2007). Normalisation accuracy is also likely to vary between brain regions, for example, highly convoluted cortex will register less well than simpler structures. This suggests that conclusions regarding fine-scale anatomical localisation should be cautious; there is no single "correct" normalisation method. Smoothing can alleviate some of the problems of inter-subject correspondence (in addition to making the data more normally distributed) but brings problems of its own. Variations in smoothing can produce very different results (Jones et al., 2005), and while investigators may have a rough idea of a reasonable kernel size for their study (based on *a priori* beliefs about the likely scale of interest), a degree of arbitrariness remains. All classical statistical tests share the limitation that failure to reject the null hypothesis does not imply that it is true (this is particularly pertinent if tests only just fail to reach arbitrary significance levels, e.g. $p=0.051$). More specifically, with SPM, the absence of a statistically significant effect in a particular region does not prove that the region is unaffected. This is especially true for VBM, where regional variation in normalisation accuracy (Crum et al., 2003) or smoothness (Ashburner and Friston, 2000) is likely to result in statistical sensitivity varying over the brain.

10. Interpret your results cautiously and in context

When implemented rigorously and interpreted carefully, VBM can be a powerful technique. Authors should be forthright in discussing potential sources of bias or imprecision, whether they arise from the study's design or analysis, or from the nature of VBM itself. Particular care should be taken when interpreting results which appear fragile with respect to more arbitrary aspects of the method such as pre-processing options and nuisance variables. A conservative approach based on robust findings, related to *a priori* hypotheses, is preferable to reporting weak effects that may be idiosyncratic to the particular parameters chosen. This approach reflects an awareness of the potential sources of error and bias that can be introduced at the different stages of a VBM study—effects that are likely to be amplified in clinical populations with inherently atypical anatomy. Despite the caveats, our basic message is brief: your VBM study should be conducted and reported in a way that is principled, transparent and replicable. Such studies have potential to become valuable contributions to the literature.

Acknowledgments

We are grateful to John Ashburner for his helpful comments and to Karl Friston for inviting the submission of this paper for review. We thank the authors of "Guidelines for reporting an fMRI study" for sharing an early draft of their work. Chris Frost provided helpful statistical advice. We are grateful to the reviewers, who made numerous detailed suggestions.

GRR is supported by the Engineering and Physical Sciences Research Council and GlaxoSmithKline through an Industrial CASE Studentship. SMDH is supported by the High Q Foundation. JDR is supported by a Wellcome Trust Research Training Fellowship. JDW is supported by a Wellcome Trust Intermediate Clinical Fellowship. RIS and NCF acknowledge support from the UK Medical Research Council. This work was undertaken at UCLH/UCL who received a proportion of funding from the Department of Health's NIHR Biomedical Research Centres funding scheme. The Dementia Research Centre is an Alzheimer's Research Trust Co-ordinating Centre.

Appendix A. VBM Reporting (and reviewing) Checklist

The aim of the following table is to assist authors in checking they have not overlooked important aspects of reporting their VBM studies; it is not intended to be an inflexible list of requirements.

Paper section and topic	Item	Rules	Description
Introduction	1	1, 9	Scientific background and rationale for the study—why use VBM to address the particular question

Appendix A (*continued*)

Paper section and topic	Item	Rules	Description
Methods			
Participants	2	1, 8	Inclusion and exclusion criteria for patients Nature of the control subjects, how they were chosen and how they were matched to the patients Location in which the data was collected and over what period of time
Objectives	3	1, 5	Specific objectives and hypotheses and in particular any a priori anatomical hypothesis
Brain imaging	4	1	Scan acquisition parameters[a] Whether all subjects were scanned on the same scanner and same parameters Timing of imaging in relation to any neurological, behavioural and/or psychometric assessments
Software	5	2, 3, 4	Name and version of package, with version of supporting software if applicable, e.g. if SPM is used: its version and the version of MATLAB Whether any parameters have been altered from the defaults[b]
Manual pre-processing	6	2, 3	Whether manual pre-processing was performed and if so, what procedures[c]; briefly describe why any procedures were thought necessary
External programs	7	2, 3, 7	Whether external programs were used[d] in addition to the main analysis software and if so, why they were thought necessary
"Optimisation"	8	2, 3	The term optimised VBM arose from the paper by Good et al. (2001) and is largely specific to SPM2; SPM5 uses unified normalisation and segmentation (Ashburner and Friston, 2005); other software packages may share some aspects VBM analyses should only be reported as "optimised" if the analysis exactly followed that outlined in Good et al. (2001)[e]; if this is not the case, the following steps should be reported Normalisation—whether this was performed to: (a) standard MNI whole-brain template (not optimised) (b) study-specific whole-brain template (semi-optimised) (c) standard ICBM GM/WM templates (more optimised) (d) study-specific GM/WM templates (i.e. as per "Good et al" optimised) Segmentation—whether this was performed to: (a) standard SPM GM/WM templates (not optimised in the "Good et al" sense) (b) study-specific GM/WM templates (i.e. as per "Good et al" optimised) How these steps were performed, e.g. within SPM[f] Why the various steps were performed in that way[g]
Clean-up procedures	9	2	Whether any clean up procedures or masking to help get rid of non-brain were used
Modulation	10	3	Whether this was performed or not and justification of this
Smoothing kernel	11	2, 8	What size kernel was used and a brief justification of why that was chosen
Statistical design			
Models	12	4, 5, 7	Factors, levels and non-sphericity options used All covariates used should be listed (with brief justification[h]), including 'nuisance' covariates Whether global normalisation was used, and how (see Rule 4) The nature of the model should be explicit and clear[i] Masks: the level of absolute or proportional masking should be specified (with justification); if an explicit mask was used, it should be specified why and how it was created; if contrast-masking is used, it should be specified why and how it was done
Contrasts	13	4, 7	These should be explicit and very clear[j]
Results			
Baseline data	14	1	Baseline demographic characteristics of the patients and controls
Numbers analysed and nature of analysis	15	1, 8	Number of patients and controls initially entered into the analysis Whether any patients or control subjects were excluded at any point during the analysis and why Whether the analysis was rerun for *any* reason and how many analyses were performed[k]
Type of correction	16	5, 7	What type of correction was used with a brief justification: (a) Uncorrected: say why, and refer to this in the interpretation of results (b) FDR (c) FWE (smoothness (FWHM) and resel count should also be reported)[l] (d) Other (e.g. resampling-based, cluster-size with non-stationarity)
Level of correction	17	5, 7	Voxel Cluster[m] SVC[n] Whether an arbitrary extent threshold was used after statistical thresholding
Threshold for all statistical maps displayed[o]	18	5, 6	This should be clear in the figure caption and in the text if the display is referred to
Threshold for reported results (e.g. tables)[o]	19	5, 6	This should be clear in the table caption and in the text where results are reported and discussed
Overlay maps and anatomical localisation	20	6	The type of image and in which space the SPMs are overlaid should be stated.[p] At which mm coordinates the slices/crosshairs are located[q]

(continued on next page)

Appendix A (*continued*)

Paper section and topic	Item	Rules	Description
			How SPM "blobs" are converted from MNI coordinates into anatomy[r] Coordinates of all local maxima should be tabulated
Other features	21	2, 3, 7	Whether any specific variations from the norm were performed in either the methods or statistical design[s]
Discussion			
Interpretation of the results	22	5, 9, 10	Interpretation of the results, taking into account study hypotheses, sources of potential bias or imprecision and the dangers associated with multiplicity of analyses and outcomes If uncorrected data are reported, this should be clear in the discussion and interpreted in light of the fact the results are uncorrected for multiple comparisons; similar considerations apply to thresholds used

[a]In particular, voxel size should be noted as larger voxels will mean that data are relatively smoother to start with.

[b]For example, DCT cut-off or smoothness of templates used in SPM5.

[c]For example, reorienting or brain masking.

[d]For example, external programs used for bias correcting, skull stripping or affine registration.

[e]Good CD, Johnsrude IS, Ashburner J, Henson RN, Friston KJ, and Frackowiak RS., 2001, A voxel-based morphometric study of ageing in 465 normal adult human brains. *NeuroImage* 14 (1 Pt 1), 21–36.

[f]For example, referencing any third-party scripts or toolboxes that have been used.

[g]For example, why study-specific whole-brain template rather than GM to GM normalisation was chosen.

[h]For example, why it was thought necessary or not to control for total intracranial volume (TIV), age, etc.

[i]Is the model a simple linear regression or are there interaction terms? The model can be shown either as an equation or in text, but the end result must be that all the regressors and their interactions are clear to the reader.

[j]For example, if there are interaction terms in the model are these really tested or not? This can be done either by listing the parameters tested (e.g. $\beta_1 > \beta_2$; $\beta_3 < 0$), or in text, but again the reader should be in no doubt as to which regressors were tested, either against each other or against zero. If single-tailed *t*-contrasts are tested the direction should be clear, and a comment on whether the "reverse contrast" was tested would be helpful (note that in most cases testing the reverse contrast will be appropriate as a 'quality assurance').

[k]For example, for technical problems, with different sub-populations, etc.

[l]Ashburner J, and Friston KJ., 2000, Voxel-based morphometry—the methods. *NeuroImage* 11(6 Pt 1), 805–821.

[m]Note that standard cluster correction as implemented in SPM should not be used for VBM because of the non-stationary smoothness of the residuals (although a suitable alternative is available: http://www.fmri.wfubmc.edu/cms/software#NS).

[n]This should only be used with *a priori* hypotheses and the centre and shape of the SVC should be made clear—ideally, an anatomical image should be used (as provided in recent versions of SPM), and the manner in which the anatomical SVC image was generated should be specified.

[o]It is not uncommon for SPMs to be displayed at a low threshold (e.g. uncorrected) but for results at a higher threshold (some form of correction) to be reported in the text and discussed; this needs to be made explicit to avoid readers mistakenly thinking that the two are equivalent.

[p]Note that ideally the SPMs should be overlaid on smooth images since they are the result of smoothed data. Overlaying an SPM on a single brain gives an impression of more precise anatomical localisation than is actually possible.

[q]Without this it can be very hard to replicate precisely an SPM, or to compare results across studies.

[r]For example, visually, or with an MNI-Talairach atlas.

[s]Examples from published papers are the use of high-pass filtering of images, or taking maximum voxel values and analysing them in non-imaging statistic software. Any extra processing or statistics outside the main software package needs to be made very clear (replicable) and should ideally be justified. VBM-like pre-processing using software not typically employed for such steps (e.g. registration or segmentation) should be carefully described.

References

Abell, F., Krams, M., Ashburner, J., Passingham, R., Friston, K., Frackowiak, R., Happe, F., Frith, C., Frith, U., 1999. The neuroanatomy of autism: a voxel-based whole brain analysis of structural scans. NeuroReport 10 (8), 1647–1651.

Amunts, K., Schleicher, A., Zilles, K., 2007. Cytoarchitecture of the cerebral cortex—more than localization. NeuroImage 37 (4), 1061–1065.

Ashburner, J., 2007. A fast diffeomorphic image registration algorithm. NeuroImage 38 (1), 95–113.

Ashburner, J., Friston, K.J., 2005. Unified segmentation. NeuroImage 26 (3), 839–851.

Ashburner, J., Friston, K.J., 2001. Why voxel-based morphometry should be used. NeuroImage 14 (6), 1238–1243.

Ashburner, J., Friston, K.J., 2000. Voxel-based morphometry—the methods. NeuroImage 11 (6 Pt 1), 805–821.

Ashburner, J., Friston, K.J., 1999. Nonlinear spatial normalization using basis functions. Hum. Brain Mapp. 7 (4), 254–266.

Ashburner, J., Friston, K.J., 1997. Multimodal image coregistration and partitioning—a unified framework. NeuroImage 6 (3), 209–217.

Baron, J.C., Chételat, G., Desgranges, B., Perchey, G., Landeau, B., de la Sayette, V., Eustache, F., 2001. In vivo mapping of gray matter loss with voxel-based morphometry in mild Alzheimer's disease. NeuroImage 14 (2), 298–309.

Bookstein, F.L., 2001. "Voxel-based morphometry" should not be used with imperfectly registered images. NeuroImage 14 (6), 1454–1462.

Crum, W.R., Griffin, L.D., Hill, D.L.G., Hawkes, D.J., 2003. Zen and the art of medical image registration: correspondence, homology, and quality. NeuroImage 20 (3), 1425–1437.

Devlin, J.T., Poldrack, R.A., 2007. In praise of tedious anatomy. NeuroImage 37 (4), 1033–1041.

Davatzikos, C., Genc, A., Xu, D., Resnick, S.M., 2001. Voxel-based morphometry using the RAVENS maps: methods and validation using simulated longitudinal atrophy. NeuroImage 14 (6), 1361–1369.

Douaud, G., Gaura, V., Ribeiro, M., Lethimonnier, F., Maroy, R., Verny, C., Krystkowiak, P., Damier, P., Bachoud-Levi, A., Hantraye, P., Remy, P., 2006. Distribution of grey matter atrophy in Huntington's disease patients: a combined ROI-based and voxel-based morphometric study. NeuroImage 32 (4), 1562–1575.

Friston, K.J., 1997. Testing for anatomically specified regional effects. Hum. Brain Mapp. 5 (2), 133–136.

Friston, K.J., Penny, W.D., Glaser, D.E., 2005. Conjunction revisited. NeuroImage 25 (3), 661–667.

Genovese, C.R., Lazar, N.A., Nichols, T., 2002. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. NeuroImage 15 (4), 870–878.

Gerig, G., Kubler, O., Kikinis, R., Jolesz, F., 1992. Nonlinear anisotropic filtering of MRI data. IEEE Trans. Med. Imag. 11 (2), 221–232.

Good, C.D., Johnsrude, I.S., Ashburner, J., Henson, R.N., Friston, K.J., Frackowiak, R.S., 2001. A voxel-based morphometric study of ageing in 465 normal adult human brains. NeuroImage 14 (1 Pt 1), 21–36.

Hill, D.L.G., Batchelor, P.G., Holden, M., Hawkes, D.J., 2001. Medical image registration. Phys. Med. Biol. 46 (1), 1–45.

Hochberg, Y., Tamhane, A., 1987. Multiple Comparison Procedures. John Wiley & Sons.

Jones, D.K., Symms, M.R., Cercignani, M., Howard, R.J., 2005. The effect of filter size on VBM analyses of DT-MRI data. NeuroImage 26 (2), 546–554.

Joshi, S., Davis, B., Jomier, M., Gerig, G., 2004. Unbiased diffeomorphic atlas construction for computational anatomy. NeuroImage 23 (Suppl 1), S151–S160.

Keller, S.S., Wilke, M., Wieshmann, U.C., Sluming, V.A., Roberts, N., 2004. Comparison of standard and optimized voxel-based morphometry for analysis of brain changes associated with temporal lobe epilepsy. NeuroImage 23 (3), 860–868.

Kiebel, S.J., Holmes, A.P., 2007. The General Linear Model. Ch. 8 in Friston. In: Ashburner, K.J., Kiebel, J.T., Nichols, S.J., T.E., Penny, W.D. (Eds.), Statistical Parametric Mapping—The Analysis of Functional Brain Images. Academic Press, pp. 109–110.

Kochunov, P., Lancaster, J.L., Thompson, P., Woods, R., Mazziotta, J., Hardies, J., Fox, P., 2001. Regional spatial normalization: toward an optimal target. J. Comput. Assist. Tomogr. 25 (5), 805–816.

Littmann, A., Guehring, J., Buechel, C., Stiehl, H., 2006. Acquisition-related morphological variability in structural MRI. Acad. Radiol. 13 (9), 1055–1061.

Ludbrook, J., 1991. On making multiple comparisons in clinical and experimental pharmacology and physiology. Clin. Exp. Pharmacol. Physiol. 18 (6), 379–392.

Maguire, E.A., Gadian, D.G., Johnsrude, I.S., Good, C.D., Ashburner, J., Frackowiak, R.S., Frith, C.D., 2000. Navigation-related structural change in the hippocampi of taxi drivers. Proc. Natl. Acad. Sci. 97 (8), 4398–4403.

Mechelli, A., Price, C.J., Friston, K.J., Ashburner, J., 2005. Voxel-based morphometry of the human brain: methods and applications. Curr. Med. Imaging Rev. 1 (1), 1–9.

Nichols, T.E., Holmes, A.P., 2002. Nonparametric permutation tests for functional neuroimaging: a primer with examples. Hum. Brain Mapp. 15 (1), 1–25.

Poldrack, R.A., Fletcher, P.C., Henson, R.N., Worsley, K.J., Brett, M., Nichols, T. E., in press. Guidelines for reporting an fMRI study. NeuroImage. doi:10.1016/j.neuroimage.2007.11.048.

Salmond, C.H., Ashburner, J., Vargha-Khadem, F., Connelly, A., Gadian, D.G., Friston, K.J., 2002. Distributional assumptions in voxel-based morphometry. NeuroImage 17 (2), 1027–1030.

Scahill, R.I., Frost, C., Jenkins, R., Whitwell, J.L., Rossor, M.N., Fox, N.C., 2003. A longitudinal study of brain volume changes in normal aging using serial registered magnetic resonance imaging. Arch. Neurol. 60 (7), 989–994.

Senjem, M.L., Gunter, J.L., Shiung, M.M., Petersen, R.C., Jack, C.R., 2005. Comparison of different methodological implementations of voxel-based morphometry in neurodegenerative disease. NeuroImage 26 (2), 600–608.

Shen, D., Davatzikos, C., 2002. HAMMER: hierarchical attribute matching mechanism for elastic registration. IEEE Trans. Med. Imaging 21 (11), 1421–1439.

Talairach, J., Tournoux, P., 1988. Co-planar stereotaxic atlas of the human brain: 3-dimensional Proportional System: an Approach to Cerebral Imaging. Thieme, Stuttgart.

Whitwell, J.L., Crum, W.R., Watt, H.C., Fox, N.C., 2001. Normalization of cerebral volumes by use of intracranial volume: implications for longitudinal quantitative MR imaging. Am. J. Neuroradiol. 22 (8), 1483–1489.

Ten simple rules for reporting voxel-based morphometry studies.

Ridgway, Gerard R; Henley, Susie M D; Rohrer, Jonathan D; Scahill, Rachael I; Warren, Jason D; Fox, Nick C

01 Kiyotaka Nemoto Page 2
24/11/2019 3:08

02 Kiyotaka Nemoto Page 2
24/11/2019 3:08

03 Kiyotaka Nemoto Page 2
24/11/2019 3:08

04 Kiyotaka Nemoto Page 3
24/11/2019 3:08

05 Kiyotaka Nemoto Page 3
24/11/2019 3:08

06 Kiyotaka Nemoto Page 3
24/11/2019 3:08