CORTEX 56 (2014) 51–63

Available online at www.sciencedirect.com

ScienceDirect

Journal homepage: www.elsevier.com/locate/cortex

Special issue: Research report

A new method for automated high-dimensional lesion segmentation evaluated in vascular injury and applied to the human occipital lobe

Yee-Haur Mah [a], *Rolf Jager* [a], *Christopher Kennard* [c], *Masud Husain* [a,b,c] *and Parashkev Nachev* [a,b,*]

[a] *Institute of Neurology, UCL, London, UK*
[b] *Institute of Cognitive Neuroscience, UCL, London, UK*
[c] *Department of Clinical Neurology, University of Oxford, Oxford, UK*

ARTICLE INFO

Article history:
Received 8 December 2011
Reviewed 01 February 2012
Revised 25 May 2012
Accepted 6 December 2012
Published online 25 December 2012

Keywords:
Lesion segmentation
Diffusion-weighted imaging
Stroke
Zeta score
Magnetic resonance imaging
Occipital lobe
Lesion-mapping

ABSTRACT

Making robust inferences about the functional neuroanatomy of the brain is critically dependent on experimental techniques that examine the consequences of focal loss of brain function. Unfortunately, the use of the most comprehensive such technique—lesion-function mapping—is complicated by the need for time-consuming and subjective manual delineation of the lesions, greatly limiting the practicability of the approach. Here we exploit a recently-described general measure of statistical anomaly, zeta, to devise a fully-automated, high-dimensional algorithm for identifying the parameters of lesions within a brain image given a reference set of normal brain images. We proceed to evaluate such an algorithm in the context of diffusion-weighted imaging of the commonest type of lesion used in neuroanatomical research: ischaemic damage. Summary performance metrics exceed those previously published for diffusion-weighted imaging and approach the current gold standard—manual segmentation—sufficiently closely for fully-automated lesion-mapping studies to become a possibility. We apply the new method to 435 unselected images of patients with ischaemic stroke to derive a probabilistic map of the pattern of damage in lesions involving the occipital lobe, demonstrating the variation of anatomical resolvability of occipital areas so as to guide future lesion-function studies of the region.

1. Introduction

1.1. Overview

To do functional neuroanatomy in the brain is to relate a discrete area, or network of areas, to a specific function, or set of functions. The strongest evidence for such a relation is the observation of disruption of a function following disruption of its putative anatomical substrate. Unfortunately, such evidence is difficult to obtain in the human brain because disrupting its activity can be done experimentally only transiently, and only for accessible regions of cortex. We must

* *Corresponding author*. Institute of Neurology, UCL, 33 Queen Square, London WC1N 3BG, UK.
E-mail address: pnachev@ion.ucl.ac.uk (P. Nachev).
http://dx.doi.org/10.1016/j.cortex.2012.12.008
0010-9452/

therefore rely on data derived from patients with focal brain lesions of natural, or incidental surgical, causes.

Now human lesion data is difficult to use for the purposes of functional neuroanatomy for two conflicting reasons. First, delineating the precise extent of the lesion—at a resolution commensurate with the underlying anatomical architecture—has hitherto been done manually, by trained operators, making the process extremely time-consuming and susceptible to operator bias. Most studies therefore rely on relatively small numbers of meticulously characterised cases. Second, the resolving power of a lesion study is generally limited not by the physical resolution of the images but by the resolution of the lesion sampling of the brain: effectively the anatomical scale at which the effects of the absence or presence of damage can be reliably determined. For example, if within a set of patients under study whenever a given voxel is hit a cluster of other voxels are always *also* hit the resolution of the resultant lesion map is not limited by the dimensions of the voxel but by the size of the cluster of invariantly affected other voxels. This limit depends not only on variations in the frequency of damage to locations across the brain but also on the multivariate *pattern* of damage in the population of lesions, a factor that is hard to quantify owing to the likely complexity of what is a very high-dimensional multivariate distribution (Nachev and Husain, 2007). Most studies may therefore require much larger numbers of cases than they actually use.

A further complication of lesion studies is the dynamic nature of the consequences of focal injury on the operation of what is inevitably a distributed, plastic network. Acutely, areas remote from the site of injury may be transiently affected in ways that do not necessarily reflect the functional contribution of the target. Chronically, remote reorganisation may abnormally compensate for a deficit, camouflaging the target's true role in the normal state. To obtain a synoptic picture of the role of a given area we therefore need *both* acute and chronic lesion studies, with image processing methodology optimised for each.

To realize in practice the power lesion-mapping has in theory we thus need methodology that permits much larger datasets to be generated; inevitably, in a fully-automated manner. This requires the development of unsupervised algorithms for the two critical steps in the processing of lesion images: distinguishing damaged from normal brain (lesion segmentation) and determining the anatomical labels of the damaged areas (lesion registration). Although a number of satisfactory algorithms exist for the latter (Crinion et al., 2007; Andersen et al., 2010; Nachev et al., 2008) no comprehensive solution exists for the former.

Here we seek to do three things. First, we show how a simple recently-described general measure of anomaly can be used to perform lesion segmentation theoretically in any imaging modality where inter-subject registration to a set of reference normal images is possible. Second, we describe and proceed to evaluate an algorithm based on this approach that is optimised for the characterisation of acute ischaemic lesions as shown by diffusion-weighted magnetic resonance imaging (MRI) (DWI). Third, we apply the new method to acute ischaemic lesions involving the occipital lobe so as to generate a map of the patterns of damage to the region, facilitating predictions about the resolvability with lesion-mapping of specific subareas within this region. In describing the general approach and our specific application we need to make a set of general points about lesion segmentation agnostically of the lesion type and imaging modality, and a set of specific points pertinent to acute stroke.

1.2. *Zeta lesion segmentation*

Taking the general points first, any comprehensive method for lesion segmentation has to deal with five fundamental problems. First, for any given imaging modality, the signal at any specific point in the brain will usually vary from one normal individual to another in a way that is difficult to parameterise: the population distribution is often not only not Gaussian, but multimodal. Our method is therefore non-parametric (Lao et al., 2008). Second, deciding whether or not a region is abnormal often depends on the signal not just in one imaging sequence but several different ones: where an abnormality is not replicated across more than one type of sequence it may merely reflect noise or artefact. Our method is therefore potentially multispectral (Prastawa et al., 2004). Third, although the signal properties of normal tissue may be definable, they are often not for lesions, simply because it is in the very nature of pathology to be heterogeneous in signal. Our method is therefore agnostic of the specific properties of the lesion signal: it identifies everything that is *anomalous* in relation to the normal reference (Prastawa et al., 2004; Shen et al., 2010). Fourth, whether or not the signal at any given locus is interpreted as normal or damaged often depends on the signal in its immediate anatomical vicinity. Our method therefore incorporates local information, in a high-dimensional manner, when determining the anomaly of each point in the brain. Fifth, the optimal properties of an image on which to perform lesion segmentation are opposite to those of an image on which to perform lesion registration: this is so because in the former normal tissue contrast interferes with the lesion contrast one needs to distinguish normal from damaged brain, whereas in the latter lesion contrast interferes with the normal tissue contrast one needs to determine the anatomical labels of the lesion. Our method therefore uses different imaging sequences for each task: one optimised for lesion segmentation and another optimised for lesion registration. Since clinical scans invariably use at least two different sequences this does not limit the application.

The core of our method is a simple measure of the anomaly of an unknown test datum in relation to a reference set of data already known to be normal. To determine the anomaly of a single datum one may simply compare it to the k instances within the reference set that resemble it most closely: its k nearest neighbours (Cover and Hart, 1967). Where the datum is a single value, a scalar, this is simply a matter of finding the k points that are closest to it on a linear scale. Where the datum has n variables describing it, a vector, this is some distance measure in n dimensional space—most simply the Euclidean—of the datum to its k nearest neighbours. To derive a unitary measure, one can take the mean of the distances of the test datum to each of the k nearest neighbours, a measure known as gamma (γ) (Harmeling et al., 2006). The attraction of gamma is that it is indifferent to the shape and number of modes of the reference population distribution, and it is