Why does language laterality index vary?

A. B. Waites[1], P. Mannfolk[1], J. Olsrud[1], D. F. Abbott[2], G. D. Jackson[2]
[1]MR Department, Lund University Hospital, Lund, Sweden, [2]Brain Research Institute, Melbourne, Australia

Introduction

A strong motivation for clinical use of functional magnetic resonance imaging (fMRI) is to determine functional language lateralisation prior to neurosurgery. Problems remain however with this approach: Agreement is not complete between fMRI lateralisation of language and the Wada test (1). Further, it has been seen that the laterality index (LI) measured with fMRI can vary with analysis approach (eg threshold), over time, and between centres. It is not known to what extent these differences are due to systematic differences or limitations in approach, and to what extent they depend on variability in the underlying distribution of cognitive activity. This study considers these questions, by studying the behaviour of LI as a function of statistical threshold and analysis approach, both in simulated datasets, where the ground truth is known, and in real language fMRI datasets.

Methods

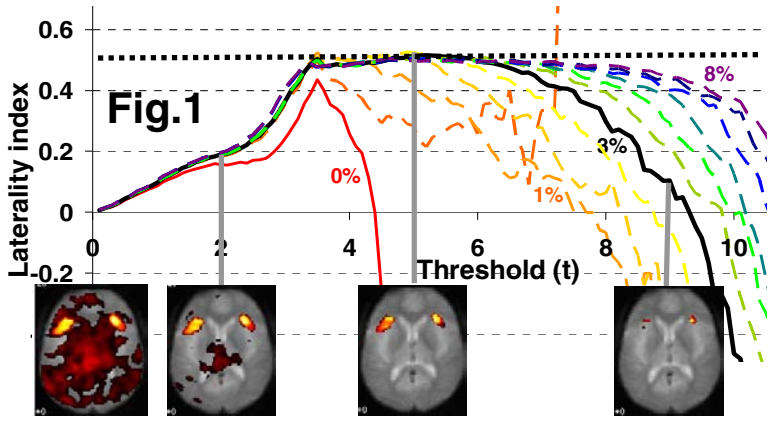
Twelve healthy volunteers were considered in this study. The fMRI studies were performed with a 3 tesla GE Signa LX scanner (GE, Milwaukee, WI). Functional images were acquired as a series of 90 gradient-recalled echo planar imaging (GR-EPI) volumes (TR/TE=3600/40ms, flip angle=60 degrees, 25 oblique slices 4mm thick+1mm gap, 24cm FOV, 128x128 matrix). During the language paradigm, a visual fixation block was interleaved with a block of orthographic lexical retrieval (OLR), a verbal fluency task where the subject generated words beginning with a displayed letter. Analysis was performed using SPM2 (www.fil.ion.ucl.ac.uk/spm). Data were motion corrected, transformed to standard space and smoothed (8mm isotropic Gaussian kernel). In the statistical analysis, motion correction parameters were included as covariates of no interest.

In addition, a series of simulated datasets were constructed using a resting state dataset from one of the subjects (subject instructed to lie still and rest). To this, known amounts of activation were added. The timecourse was constructed as the same series of blocks of task and rest as in the language study, convolved with a haemodynamic response function. Percent signal change (between task and rest conditions) was varied between 0 and 8%.

LI was calculated by comparing activation above threshold in a left and right region of interest (ROI), which included lateral cortical brain voxels excluding the cerebellum. The measure used was varied (number of voxels or average t score), as well as the method of selecting threshold (fixed, or subject dependent based on levels of activation)(2).

Results

Fig 1 shows the behaviour of the measured LI both as a function of the chosen threshold (x-axis) and as a function of the percent signal change (different coloured lines) in the simulated datasets. The true LI of 0.5 is shown by the bold dotted line. The solid black line represents 3% activation. Image slices represent statistical parametric maps calculated for this dataset at various thresholds, as indicated by the vertical grey lines.

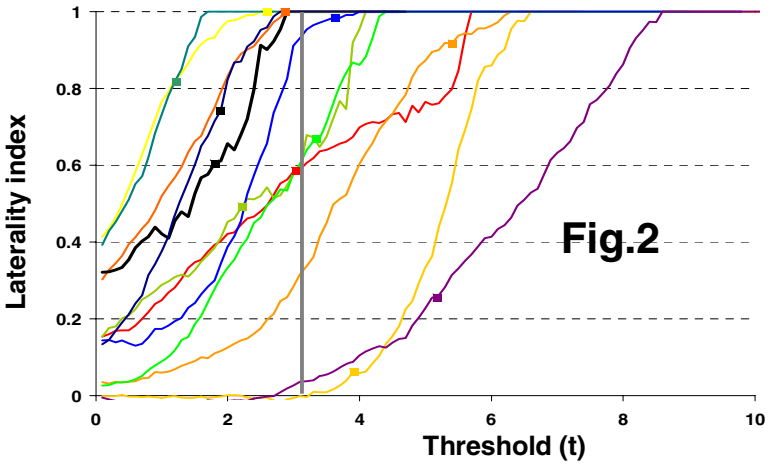


In Fig.2, patient data is shown, again with LI varying as a function of the statistical threshold chosen. For each subject, represented as different coloured lines, the threshold chosen using the method of ref 2 is shown as a point. Also shown is the "standard" threshold of t=3.2.

Discussion

LI varies as a function of the statistical threshold chosen, and this variation is subject specific. This is a result of three competing processes on the calculated LI, as seen in Fig.1. First, increasing from low thresholds, one can observe the continuously changing LI, as the influence of false positives decreases. In the simulated data, this is followed by a region where the true LI can be measured, and then finally, a region where signal to noise ratio (SNR) limits are reached, and LI behaves unpredictably for the final few (most) activated voxels.

In the patient data presented in Fig.2, it appears that the effects of false positives and false negatives overlap, so that there is no region where LI is independent of threshold. This problem is even seen in the simulated data for low percent signal change (1%, Fig.1). It appears that, at least for these subjects, setting a single threshold such as t=3.2 reveals little consistent information about the relative distribution of their language activation. Even an adaptive threshold such as we have considered yields results which are difficult to interpret, given the distribution of LI with threshold seen in Fig.2. As noted in Fig.1, an increase in SNR (reflected in simulated data as increased percentage signal change) leads to an increase in the length of the plateau in LI. Thus, it is possible that improving SNR in patients (eg by increasing the paradigm duration) could improve the stability of the calculated LI. Alternatively, it may be that the calculation of LI is a somewhat unreliable measure of the distribution of language activation. Other approaches for describing the language network may improve robustness, for instance taking consideration of the t-scores of voxels, rather than simply counting them. It may be that the observed unreliability in measured LI only reflects quantitative problems, but that the qualitative result (typical or atypical laterality) remains a robust and clinically useful tool.



References

1). Woermann, F. et al., 2003. Neurology, 61(5) 699-701.
2). Fernandez, et al. 2001. Neuroimage, 14(3), 585-94.