Robust fMRI for language lateralisation: How do different threshold-independent methods compare?

Irène Brumer[1,2], Marco Borri[1], Enrico De Vita[2], Jonathan Ashmore[1], Jozef Jarosz[1]

[1] *Neuroradiology, King's College Hospital, London,* [2] *School of Biomedical Engineering and Imaging Sciences, King's College London*

Introduction: Language lateralisation, and in particular the determination of the dominant brain hemisphere for language functions, can aid presurgical planning, especially for epilepsy patients [1]. Hemispheric or regional dominance in language functions can be evaluated using functional MRI (fMRI), and is quantified using the laterality index (LI). This value is conventionally determined by counting the number of voxels above a chosen activation threshold within the left and right hemispheres and yields values ranging from +1 (left dominant) to -1 (right dominant). A major limitation of this approach is the strong dependence of the LI on the arbitrarily chosen threshold (see Figure 1). Although a number of threshold-independent methods have been proposed [2], they have not been directly compared. In this work we apply three threshold-independent methods to the same cohort of healthy subjects, who have performed identical tasks, and compare the produced LIs.

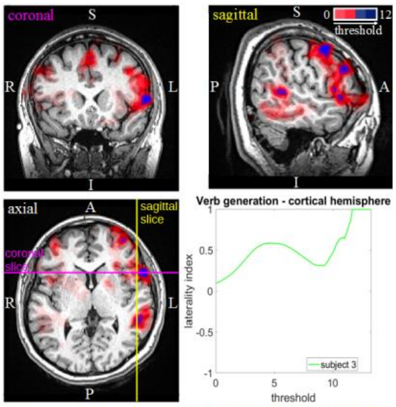

Figure 1 : Subject 3, verb generation. The extent of the activation regions depends on the threshold. At a low threshold (light red), numerous regions are considered active. As the threshold increases, only small spots remain (dark blue). The plot shows the dependence of the LI (evaluated over the cortical hemispheres) on the threshold.

Methods: Twelve healthy volunteers were scanned at 1.5 T (Magnetom Aera, Siemens AG, Erlangen, Germany) using the standard head receive coil. All volunteers gave informed consent under local ethics. The MRI sequence protocol consisted of a 3D T1-weighted MPRAGE anatomical sequence (echo time TE = 3.02 ms, repetition time TR = 2200 ms, flip angle FA = 8°, field of view FOV = 250x2526 mm^2 (transversal), voxel size 1x1x1 mm^3, GRAPPA of 2) and three BOLD contrast fMRI EPI sequences (TE = 40 ms, TR = 3000 ms, FOV = 240x240 mm^2, in plane resolution 2.5x2.5 mm^2 (transversal), slice thickness 3 mm, 30 slices). Each volunteer performed three different tasks: picture naming, verb generation, and word fluency. Each fMRI acquisition consisted of 6 cycles of alternating rest and activation periods of 30 s each (120 measurements). In this work we consider two different regions of interest (ROIs) created using standard brain atlases: 1) the 'hemisphere ROI', which encompasses the entire cortical hemisphere (excluding the cerebellum) and 2) the 'language ROI', which is defined as the combined Broca's (Brodmann areas 44 and 45) and Wernicke's areas (posterior division of the superior temporal gyrus). Three different LI calculation methods were chosen for the comparison. The first method, labelled '*curveLI*', calculates the LI as a function of the total number of activated voxels within the ROIs corresponding to different threshold values and produces a LI curve [3]. The single LI value used for the comparison was calculated at 50% of the total number of activated voxels (mid curve). The second method ('*aveLI*') calculates an average LI by averaging the conventional LI values over a range of thresholds [4]. The third method ('*histoLI*') integrates the weighted histogram of voxel counts against threshold in right and left ROIs to calculate a global LI [5]. For each method, the subjects were ranked according to the obtained LI values. To quantify the agreement between pairs of different rankings, Spearman's correlation coefficients ρ were calculated.

Results and discussion: Figure 2 shows the results obtained with the *curveLI* method for the verb generation task. The curves show smoother variation over the range of activated voxels compared to the conventional LI versus threshold curves (see Subject 3, Figure 1). The curves show higher LI values for the language ROI compared to the hemisphere ROI, indicating a more lateralised activation on a regional scale (language area) than on a larger scale (cortical hemisphere). However, the highest and lowest curves belong to the same subjects for both ROIs. The subject rankings are similar between methods, with Spearman's correlation coefficients ρ ranging from 0.66 to 0.99. Overall, the agreement between methods is higher when the language ROI is considered rather than the hemisphere ROI. However, the localisation of language areas might be uncertain in patients presenting altered brain anatomy and potentially reorganised functionality, and therefore an approach which considers the whole hemisphere might be preferable. The agreement between the *curveLI* and *aveLI* methods is optimal for the verb generation (ρ=0.99 for both ROI) and word fluency tasks (ρ=0.96 and 0.94 for the language and hemisphere ROIs, respectively). The ranking comparisons with the *histoLI* method result in the lowest Spearman's correlation coefficients. Figure 3 shows the ranked LI values for all subjects calculated with the three methods for the language ROI. All three methods agree on ranking the top and bottom subjects. Considering both the variation of the LI with the threshold (conventional LI) and with the total number of voxels (*curveLI*), the verb generation task yields the most stable values, confirming previous findings [6,7]. The picture naming task shows the most variability both on the conventional LI and *curveLI* curves, and yields lower LI values compared to the two other tasks.

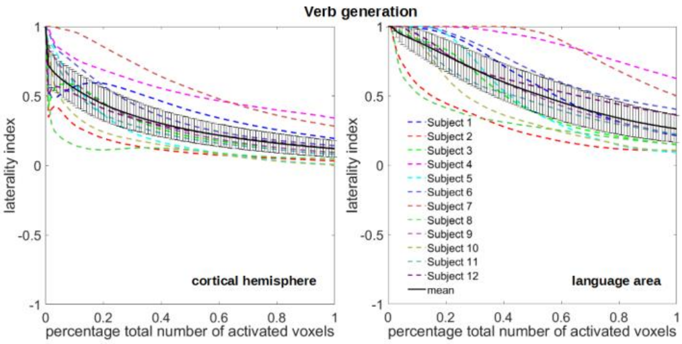

Figure 2: Laterality index versus number of activated voxels plot for the verb generation task for all subjects, including mean laterality index curve and 95% confidence interval.

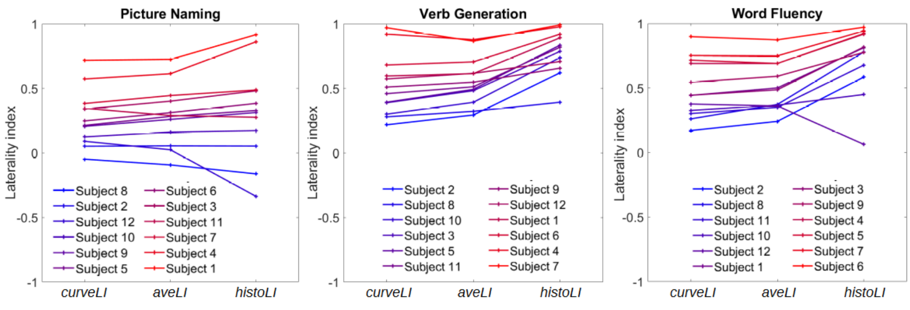

Figure 3: Subject ranking obtained with the three different methods for the language area. The subjects are ordered according to *curveLI* and the colour changes gradually from blue for low LI values to red for high LI values.

Conclusion: In this work we have compared three different threshold-independent LI calculations methods. Our results show that even though the methods are based on different assumptions, they overall agree in differentiating strong from weak lateralisation on both hemispheric and regional scales. These results demonstrate the feasibility of implementing robust and clinically relevant language lateralisation using fMRI.

References: [1] Tailby et al., 2017. *NeuroImage: Clinical*, *14*, pp.141-150. [2] Bradshaw et al., 2017. *PeerJ*, *5*, p.e3557. [3] Abbott et al., 2010. *Neuroimage*, *50*(4), pp.1446-1455. [4] Matsuo et al., 2012. *Journal of neuroscience methods*, *205*(1), pp.119-129. [5] Suarez et al., 2008. *Brain Imaging and Behavior*, *2*(2), pp.59-73. [6] Bizzi, 2009. *Neuroimaging Clinics*, *19*(4), pp.573-596. [7] Black et al., 2017. Recommended fMRI Paradigm Algorithms for Presurgical Language Assessment. *American Journal of Neuroradiology*.