Brain Network Measures and Spatial Lesion Distribution in a Sample of 47 Patients with Systemic Lupus Erythematosus (SLE)

Authors: Maria del C. Valdés Hernández, Keith Smith, Mark E. Bastin, E. Nicole Amft, Stuart H. Ralston, Joanna M. Wardlaw, Stewart J. Wiseman

ABSTRACT

This dataset corresponds to a research on network organisation of brain structural connectivity in systemic lupus erythematosus (SLE) and its putative association with lesion distribution and disease indicators.

White matter hyperintensity (WMH) segmentation and connectomics were performed in 47 patients with SLE from structural and diffusion magnetic resonance imaging (MRI) data. Network nodes were divided into hierarchical tiers based on numbers of connections. From each tier, and globally, brain network metrics were extracted for each patient. Voxel-based analyses of the spatial distribution of WMH in relation to disease indicators and network measures were also conducted.

The data comprise all lesion distribution maps from the sample, the raw results (in nifti-1 format) of the voxel-based analyses of the WMH distribution in relation to disease indicators and network parameters, the MATLAB code to process the data, and the excel spreadsheets with the brain network metrics.

Key words: Connectome, SLE, network analysis

Abbreviations: EP = echo planar; FA = fractional anisotropy; ROI = region of interest; SLE = systemic lupus erythematosus; SVD = small vessel disease; VWF = von Willebrand Factor; WMH = white matter hyperintensities;
MATERIALS AND METHODS

Subjects

The data correspond to 47 patients with Systemic Lupus Erythematosus (SLE) recruited by advertisement from staff working at the University of Edinburgh, the Western General Hospital and the Royal Infirmary, Edinburgh, United Kingdom, who underwent brain MRI and had available connectome data. All patients were examined by a consultant rheumatologist at a specialist SLE clinic between April and December 2014. The South East Scotland Research Ethics Committee gave approval (01, 14/SS/0003) to the primary study for which they were recruited. All participants gave written consent (Wiseman et al., 2016b).

Disease indicators

Figure 1 shows the clinical data analysed. Current SLE disease activity was assessed using the Systemic Lupus Erythematosus Disease Activity Index 2000 (Gladman et al., 2002). Increasing level of anti-double-stranded DNA and lower levels of complement C3 and C4 proteins from blood samples were also considered indicators of disease activity. Accumulated permanent damage from SLE was assessed with the Systemic Lupus International Collaborating Clinics (SLICC) (Gladman et al., 1996; Gladman et al., 2000) damage index and disease duration. We used indicators of endothelial dysfunction and inflammation extracted from SLE patients’ blood samples. These were the von Willebrand Factor (VWF) antigen, F8c, homocysteine, and the cytokine interleukin 6 (IL6). We also used the following vascular risk factors: presence vs. absence of hypertension and smoking status from the patients’ medical history, and measures of total cholesterol, homocysteine and anticardiolipin IgG and IgM obtained from the analyses of the blood samples (Wiseman et al., 2016b). Fatigue was assessed using the Fatigue Severity Scale (Wiseman et al., 2016a). Depression and anxiety were assessed with the Hospital Anxiety and Depression Scale (HADS) and fibrinolysis was assessed through D-dimer presence in blood.
From 47 patients with Systemic Lupus Erithematosus (SLE)

<table>
<thead>
<tr>
<th>Disease activity</th>
<th>Accumulated damage</th>
<th>Endothelial dysfunction and inflammation</th>
<th>Vascular Risk Factors</th>
<th>Fatigue, Depression and Anxiety</th>
<th>Fibrinolysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>SLEDAI</td>
<td>SLICC</td>
<td>Von Willebrand factor</td>
<td>Hypertension (Y/N)</td>
<td>FSS</td>
<td>D-dimer</td>
</tr>
<tr>
<td>Ds-DNA</td>
<td>Disease duration</td>
<td>Homocysteine</td>
<td>Smoking</td>
<td>HADS Depression Score</td>
<td></td>
</tr>
<tr>
<td>C3</td>
<td></td>
<td>F8c</td>
<td>Cholesterol (Total, LDL, HDL total/HDL ratio)</td>
<td>HADS Anxiety Score</td>
<td></td>
</tr>
<tr>
<td>C4</td>
<td></td>
<td>IL6</td>
<td>Homocysteine</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 1. Data analysed from the sample.

**MRI acquisition**

All MRI data were acquired using a GE Signa Horizon HDxt 1.5 T scanner (General Electric, Milwaukee, WI, USA) using a self-shielding gradient set with maximum gradient strength of 33 mT m\(^{-1}\) and an 8-channel phased-array head coil. The scan protocols included axial T\(_2\)-, gradient-recalled echo-, fluid-attenuated inversion recovery (FLAIR)-, sagittal T2- and high-resolution coronal 3D T1-weighted volume sequences, and a whole brain DT-MRI acquisition. The DT-MRI protocol from both studies consisted of three T2-weighted and 32 diffusion-weighted (\(b=1000\) s mm\(^{-2}\)) axial single-shot spin-echo echo-planar (EP) imaging volumes (field of view 240 × 240 mm, matrix 128 × 128, TR 13.75 s, TE 78.4 ms) (Wiseman et al., 2016a,b). Scanning protocol parameters are detailed in Table 1.
<table>
<thead>
<tr>
<th>Sequence</th>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1-weighted (T1W)</td>
<td>Acquisition/Orientation</td>
<td>IR/Sag 3D</td>
</tr>
<tr>
<td></td>
<td>TR/TE/TI (ms)</td>
<td>9.7/4/500</td>
</tr>
<tr>
<td></td>
<td>Voxel size (mm³)</td>
<td>0.94 x 1.3 x 0.94</td>
</tr>
<tr>
<td></td>
<td>Flip angle (°)</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Bandwith (KHz)</td>
<td>15.63</td>
</tr>
<tr>
<td>T2-weighted (T2W)</td>
<td>Acquisition/Orientation</td>
<td>Propeller FSE/Ax 2D</td>
</tr>
<tr>
<td></td>
<td>TR/TE/TI (ms)</td>
<td>9060/100</td>
</tr>
<tr>
<td></td>
<td>Voxel size (mm³)</td>
<td>0.47 x 0.47 x 2.5</td>
</tr>
<tr>
<td></td>
<td>Flip angle (°)</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Bandwith (KHz)</td>
<td>25</td>
</tr>
<tr>
<td>FLAIR</td>
<td>Acquisition/Orientation</td>
<td>Ax (2D)</td>
</tr>
<tr>
<td></td>
<td>TR/TE/TI (ms)</td>
<td>9402/145/2350</td>
</tr>
<tr>
<td></td>
<td>Voxel size (mm³)</td>
<td>0.47 x 0.47 x 5</td>
</tr>
<tr>
<td></td>
<td>Bandwith (KHz)</td>
<td>15.63</td>
</tr>
<tr>
<td>SWI/GRE</td>
<td>Acquisition/Orientation</td>
<td>GRE/Ax 2D</td>
</tr>
<tr>
<td></td>
<td>TR/TE/TI (ms)</td>
<td>1460/14</td>
</tr>
<tr>
<td></td>
<td>Voxel size (mm³)</td>
<td>0.47 x 0.47 x 2.5</td>
</tr>
<tr>
<td></td>
<td>Bandwith (KHz)</td>
<td>48.8</td>
</tr>
<tr>
<td>DTI</td>
<td>Acquisition/Orientation</td>
<td>Single-shot spin-echo echo-planar/Ax</td>
</tr>
<tr>
<td></td>
<td>TR/TE/TI (ms)</td>
<td>13750/78.4</td>
</tr>
<tr>
<td></td>
<td>Voxel size (mm³)</td>
<td>2 x 2 x 2</td>
</tr>
<tr>
<td></td>
<td>Directions</td>
<td>72</td>
</tr>
<tr>
<td></td>
<td>FOV (mm)</td>
<td>240 x 240</td>
</tr>
<tr>
<td></td>
<td>Matrix</td>
<td>128 x 128</td>
</tr>
</tbody>
</table>

Legend: Ax = axial. Cor = coronal. DTI = diffusion tensor imaging. FOV = field of view. FSE = fast spin echo. GRE = gradient-recalled echo. IR = inversion recovery. ms = milliseconds. Sag = sagittal. TR = time to repetition. SWI = susceptibility-weighted imaging. TE = time to echo.

**Structural image processing**

Each 3D T1-weighted volume was parcellated into 85 regions-of-interest (ROI), consisting of 68 cortical (34 per hemisphere) and 16 sub-cortical (eight per hemisphere) regions, plus the brainstem, using the
Desikan-Killiany atlas in FreeSurfer (http://surfer.nmr.mgh.harvard.edu). The results of the segmentation procedure were then used to construct the tissue and ROI masks for use in the connectome network construction and to constrain the tractography output. To extract intracranial and WMH volumes, all structural sequences were rigidly linearly aligned to T2-weighted space using the FMRIB’s Linear Image Registration Tool (FLIRT) from FSL (http://fsl.fmrib.ox.ac.uk/fsl). The intracranial volume (ICV) was extracted semi-automatically using the T2*-weighted GRE sequence with the Object Extraction Tool in Analyze™ 11.0 followed by manual editing to exclude any erroneously included extracranial tissues. The WMH were extracted semi-automatically using the MCMxxxVI Lesion Extraction tool available from www.sourceforge.net/projects/bric1936. This tool uses co-registered images, which are mapped and fused in the red-green-blue colour space. In this analysis the FLAIR and T2*-weighted images were mapped in the green/red space respectively. After the colour-fusion, a minimum variance quantisation was applied to reduce the bit resolution and facilitate the thresholding selection of each colour band to segment the tissue of interest and allow manual editing of the results for improved accuracy.

**Spatial maps of WMH**

All patients’ structural brain images were co-registered to a study template using 12-degrees affine registration (as per Dickie et al. 2015 and Valdés Hernández et al. 2015, specifically for the case of inter-subject co-alignment of periventricular and deep brain lesions). We used NiftyReg (Modat et al., 2010) (http://sourceforge.net/projects/niftyreg/) through TractoR (http://www.tractor-mri.org.uk/diffusion-processing) for co-registering the images. The study template was the 55 years old brain template from https://datashare.is.ed.ac.uk/handle/10283/1957 (Dickie et al. 2016), as it reflects the approximate mean age of this sample (i.e. 48.5 +/- 13.7 years). Then, we applied the space transformation to the WMH binary masks (obtained as per above). Once the WMH from all patients were all in the study template space, we generated a) spatial probability maps of WMH for each patient subgroup (e.g. hypertensive/normotensive patients, patients with high/low cholesterol, etc.) and b) a 4D volume of all WMH maps concatenated. Patient subgroups were determined by dichotomising and separating into quartiles the disease indicators listed in the subsection “Disease indicators” above. The threshold used to dichotomise the continuous variables was the median value in the sample.

**Analyses of the spatial WMH maps**

To explore putative regions where the WMH burden could be associated with each disease indicator we used the Wilcoxon’s rank sum test (to perform a voxel-wise comparison of the WMH maps...
between two opposite patient groups: e.g. normotensive vs. hypertensive patients, patients with high cholesterol vs. those with low cholesterol, etc.) or the Kruskal-Wallis test (for comparing the WMH maps from more than two groups, e.g. patients falling in each quartile of the vWF antigen). Voxel-wise false discovery rate was used to correct for multiple comparisons. This analysis was implemented through two MATLAB functions that are part of this dataset, namely ‘robust4Dranksum’ and ‘robust4Dkruskalwallis’, which use the MATLAB functions ‘ranksum’ and ‘kruskalwallis’ respectively. Figure 2 schematically represents this process.

Group comparison (of WMH spatial distribution)

Figure 2. Schematic representation of the steps for comparing the WMH spatial distribution of the SLE patients grouped by levels of each disease indicator

To determine whether the spatial patterns observed reflected true associations after correcting for age and biological sex differences between patients, we conducted a spatial voxel-based analysis of the WMH distribution of the sample in relation to each disease indicator, using the 4D WMH volume constructed as previously explained and a machine learning approach (Figure 3). This used the MATLAB function “fitrlinear” to fit a regularised Support Vector Machine regression model with a ridge penalty type optimised through a stochastic gradient descent approach for accuracy. This model was selected due to the high-dimensionality and sparsity of the predictor data. In these regression models our predictor was the probability distribution map of WMH in the sample, the covariates were age and biological sex and the outcome was the disease indicator. Also, to reduce sparsity, each of the 3D WMH arrays (i.e. these 3D arrays form the 4D array used in the models) were resized to the 3D space
limited by the bounding box of the intracranial volume of the study brain template. The regularisation term strength was set at 1/47.

Regression models  
(Associations between WMH and disease indicators accounting for age and biological sex)

Figure 3. Schematic representation of the steps for calculating the association - at a voxel level - of the WMH spatial distribution in the sample and the disease indicators

Diffusion MRI data processing

We pre-processed the DT-MRI data using tools provided by the FMRIB Diffusion Toolbox (FDT) package in FSL, to reduce systematic imaging distortions and bulk subject motion artefacts (Jenkinson and Smith, 2001). After skull stripping and brain extraction of the diffusion-weighted EP volumes the fractional anisotropy (FA) volume was calculated by DTIFIT in each subject (Basser and Pierpaoli, 1996). The neuroanatomical ROIs determined by FreeSurfer were then aligned to the diffusion space using a cross-modal nonlinear registration method as explained in Wiseman et al. 2018.

Tractography
Whole-brain probabilistic tractography was performed using FSL’s BedpostX/ProbTrackX algorithm (Behrens et al., 2007). Probability density functions, which describe the uncertainty in the principal directions of water molecule diffusion, were computed with a two-fibre model per voxel (Behrens et al., 2007). Streamlines were then constructed by sampling from these distributions during tracking using 100 Markov Chain Monte Carlo iterations with a fixed step size of 0.5 mm between successive points. Tracking was initiated from all white matter voxels and streamlines were constructed in two collinear directions until terminated by the following stopping criteria designed to minimize the amount of anatomically implausible streamlines: (i) exceeding a curvature threshold of 70 degrees; (ii) entering a voxel with FA below 0.1; (iii) entering an extra-cerebral voxel; (iv) exceeding 200 mm in length; and (v) exceeding a distance ratio metric of 10. The distance ratio metric (Bullitt et al., 2008), excludes implausibly tortuous streamlines.

**Network construction**

FA-weighted networks were constructed by recording the mean FA value along streamlines connecting all ROI (network node) pairs. The endpoint of a streamline was the first grey matter ROI encountered when tracking from the seed location. Self-connections were removed, and if no streamlines were found between a pair of nodes, the corresponding matrix entry was set to zero. All networks had the same number of nodes, however the number of links varied across participants and cohorts. As many network metrics are dependent on number of nodes and links (Dormann et al., 2009), we binarized all networks to have the same number of links. The minimum link density (proportion of existing links to total possible) of any participant was 29.72%. We rounded this down the nearest 5%, taking the links corresponding to the largest 25% of weights as our binary networks, ensuring maintenance of hierarchical structure (Smith, Abasolo and Escudero, 2017). The left and right ventral diencephalon were not considered as nodes, hence network construction was based on 83 ROIs.

**Network analysis**

Two levels of analysis (firstly global then hierarchical tier-based) were implemented to understand characteristics of the connectomes and possible relationships between these two levels. For each level, all network metrics described in Smith et al. 2019 were calculated. From them, three global network metrics were chosen for more detailed analyses, based on their suitability and known relevance to the human structural connectome:
1) **network density / average degree:** refer to the number of links relative to the number of nodes in the network is a widely studied property (Newman, 2003). Globally we compute network density for the unthresholded networks, since network density is fixed by the threshold and so would not differ from participant to participant.

2) **global clustering coefficient:** assesses the tendency of neighbouring nodes to connect to the same other neighbours. Based on modelling which incorporated a variety of different topological characteristics, this concept of homophily has been shown to be a particularly evident trait of structural connectomes (Betzel *et al.*, 2016).

3) **hierarchical complexity:** measures the extent of topological/functional diversity across the degree hierarchy of the connectomes (Smith and Escudero, 2017; Smith *et al.*, 2019).

Following Smith *et al.* (Smith *et al.*, 2019), each structural connectome was split into four tiers based on quartiles of the maximum degree. Tier 1 consisted of all nodes with degree greater than 75% of the maximum degree, Tier 2 of all nodes with degree greater than 50% and up to 75% of the maximum degree, and so on.

**DATA FILES**

**MATLAB code files**

*Example_voxelwise_comparison_between_group_maps.m* – Example MATLAB script to determine the voxel-wise statistical differences between groups of probability maps. It uses the functions robust4Dranksum and robust4Dkruskalwallis, which are part of this dataset, and others for manipulating the image files, from BRIClib ([www.sourceforge.net/projects/bric1936](http://www.sourceforge.net/projects/bric1936)).

*robust4Dkruskalwallis.m* – MATLAB function that takes a 4D array formed by sequentially concatenating 3D arrays (i.e. volumes), and a numeric vector with values corresponding to the groups to which each volume belongs to (see Figure 2 and example script provided), and performs the Kruskal-Wallis test between the groups of 3D volumes. Then, corrects the results against type 1 error and gives the 3D volumes of the p-values from 1) the Kruskal-Wallis test and 2) false discovery rate correction.

*robust4Dranksum.m* – MATLAB function that takes a 4D array formed by sequentially concatenating 3D arrays (i.e. volumes), and a logic vector with values corresponding to the group to which each volume belongs to (see Figure 2 and example script provided), and performs the Wilcoxon rank sum test between two groups of 3D volumes. Then, corrects the results against type 1 error, giving as
outputs the 3D volumes of the p-values from 1) the Wilcoxon's ranksum test and 2) false discovery rate correction.

*Example_script_4Dvolume_WMH.m* – Example MATLAB script to generate the 4D array of WMH by concatenating all WMH masks from all patients, once they are in the study template space. The output from this script is used as input in *Example_voxelwise_comparison_between_group_maps.m*

**Brain lesion maps**

*p-value maps* – These are nifti-1-formatted files of the results from the voxel-wise comparison of the WMH spatial distribution of patients grouped by disease indicators. However, for visualisation purposes, the values of each voxel are equal to 1 – p-value. Hence, the voxels with value equal or closer to 1 correspond to p-values close to 0. Those files with filenames *p_*_maps_thr01.nii.gz* correspond to p-values less than 0.01. Files with filenames *p_*_maps_thr05.nii.gz* correspond to p-values less than 0.05.

*FDR maps* - These are nifti-1-formatted files of the results from the voxel-wise comparison of the WMH spatial distribution of patients grouped by disease indicators after false discovery rate (i.e. Type 1 error correction). However, for visualisation purposes, the values of each voxel are equal to 1 – p-value. Hence, the voxels with value equal or closer to 1 correspond to p-values close to 0 (as is for the p-value maps).

*B-value maps* - These are nifti-1-formatted files of the B values resulting from calculating the voxel-wise association between the WMH spatial distribution and the disease indicators or network parameters. These were generated using Support Vector Machine through the MATLAB function ‘fitrlinear’ as described above (i.e. in the Methods section of this document).

*Lesion distribution maps* – These are nifti-1-formatted files of the sum of the binary WMH masks (in the template space) of the patients grouped by disease indicator values (e.g. patients with normal-low cholesterol, patients with high cholesterol, etc.)

*WMH_SLE.nii.gz* – White matter hyperintensities probability distribution map of the sample in nifti-1 format, compressed.

*WMH_4D.nii.gz* – 4D volume of the white matter hyperintensity masks of the patients that conform the sample, in nifti-1 format, compressed.
Excel file

*Graph_Analyses_metrics_for_datashare.xlsx* – Microsoft Excel Worksheet (.xlsx) with the values of the network parameters (general and by tiers) for each patient.

PDF files

*Brain_Network_in_SLE_Data_Summary.pdf* – this file

*Overview_SLE_data_processing.pdf* – Graphical overview of this dataset and the processing methods used to derive it.

REFERENCES


Acknowledgements:

We acknowledge support from the Scottish Lupus Exchange registry.

Funding:

MCVH is funded by the Row Fogo Charitable Trust (grant No. BROD.FID3668413) and received funds from the European Union Horizon 2020 [PHC-03-15, project No 666881, ‘SVDs@Target’]. SJW is funded by the Stroke Association (grant No. SA PDF 18\100026). Recruitment, image acquisition and data processing for the SLE patients was funded by Lupus UK and the University of Edinburgh.

Competing interests:

The authors report no competing interests.