Author’s response to reviews

Title: Diffusion Tensor Imaging in Neuropsychiatric Systemic Lupus Erythematosus

Authors:

Rex E Jung (rjung@mrn.org)
Arvind Caprihan (acaprihan@mrn.org)
Robert S Chavez (rchavez@mrn.org)
Ranee A Flores (rbarrow@mrn.org)
Janeen Sharrar (jsharrar@salud.unm.edu)
Clifford R Qualls (cqualls@salud.unm.edu)
Wilmer Sibbitt (wsibbitt@salud.unm.edu)
Carlos A Roldan (croldan@salud.unm.edu)

Version: 2 Date: 25 February 2010

Author’s response to reviews:

Response to Reviewers:

We thank the reviewers for their detailed and helpful comments. We have addressed each of the suggested revisions below, and believe that these revisions have improved the manuscript significantly.

The specific revisions are addressed as follows:

REVIEWER 1:

The authors report mean diffusivity and discuss this finding with regard to potential neuropsychological mechanisms. A more specific analysis of the directionality of this increased diffusivity in NPSLE (i.e. regarding potential alterations in radial and axial diffusivity) would provide additional information on the mechanisms underlying these alterations (i.e. increased axial diffusion is often interpreted in light of axonal alterations whereas increased radial diffusion is assumed to indicate decreased myelination). The authors might want to perform these additional analyses which can be easily done with TBSS.

>>We have conducted these additional analyses with AD and RD to determine whether the effects are more “axonal” or “myelin” driven. These post hoc analyses are now added to the results section, and figures and tables have been added accordingly.

- As indicated in Tab. 2 multiple white matter alterations were detectable in the high resolution scans in the SLE group. At first sight, this is surprising given no
differences between SLE and controls in the TBSS analyses. This might have methodological reasons which, however, should at least be addressed in the discussion.

>>We agree that this was a surprising finding. There are at least two possibilities: 1) that the disease load in SLE did not reach a level of significance that was detectable with TBSS, given the degrees of freedom/sample size, or 2) that the lesions did not overlap significantly in SLE while they accumulated in a systematic way in NPSLE. We have addressed this briefly in the discussion as follows: “This could be due to the lower disease burden of this cohort, which did not reach a “threshold” detectable by our imaging and statistical methodology, or due to the non-overlapping lesion burden of the SLE cohort as compared to the NPSLE cohort. It will be important to determine in future studies, with larger samples, whether there are systematic effects of NPSLE that accumulate in the brain in such a way that they affect white matter in a systematic (as opposed to sporadic) manner, as this preliminary study suggests.”

- No information on duration of illness or number of NPSLE episodes is given. In light of the fact that old infarcts were obviously not decisive regarding the white matter alterations detected in the NPSLE patient group it would be interesting to know if duration of illness had an influence of the results (with e.g. a longer duration of illness in the NPSLE group which might have been associated with stronger (accumulative) degenerative processes across time).

>>Unfortunately, we do not possess accurate data on disease duration. We have some self report, but there are numerous problems with issues of accurate measure of disease onset in SLE, first symptom presentation, and the impact of acute flares. While looking at acute flares etc. would be an interesting contrast to assess within the NPSLE cohort, we do not possess adequate power to determine this in such a small sample. We continue to collect data and expect to be able to address systematic relationships between disease characteristics such as duration estimates, acute flares, types of neuropsychiatric illness (e.g., seizure vs. TIA etc.) as impacting DTI in a future report.

- The diagnostic differentiation between SLE and NPSLE is known to be quite difficult in some cases. The authors declare the presence of acute stroke or transient ischemic attack (TIA), acute confusional state, moderate cognitive dysfunction, seizures, or psychosis as diagnostic criteria for NPSLE. However, no information on the screenings and methods used to determine the existence of these criteria are provided (e.g. how have potential cognitive dysfunctions been assessed, which screening procedure has been applied to determine presence of psychosis etc.). And, maybe even more important, have other potential causes for manifestation of neuropsychiatric symptoms (like e.g.
infection or medication side-effects) been excluded? More information regarding
diagnosis of NPSLE should be given.

>>We have added a section entitled “Clinical Measures” which details the
assessment of patients by Dr. Wilmer Sibbitt, M.D., an expert clinician in the field
of systemic lupus erythematosus. No patients were entered into the study who
were suspected of infection or medication side effects as cause of
neuropsychiatric symptoms. This methodology is considered to be standard in
the field in diagnosis of NPSLE. The paragraph reads as follows: “The Systemic
Lupus Erythematosus Disease Activity Index (SLEDAI) [1] and Systemic Lupus
International Collaborating Clinics/America College of Rheumatology Damage
Index (SLICC/ACR DI) [2] were administered by an experienced rheumatologist
(WLS) to each patient. Clinical diagnosis of NPSLE was defined by presence of
past or current: stroke, transient ischemic accident, psychosis, seizure disorder,
confusional state, and/or moderate or severe cognitive dysfunction. No SLE
patients had any past or current evidence of any of these clinical diagnoses.”

- Minor essential revisions:
- The threshold used for the mean FA skeleton should be provided as results are
influenced by this threshold (as the authors probably know a default of FA>0.2 is
recommended which can be adapted, however).

>>We did use a threshold of FA>0.2 and this is now spelled out in the methods
section.

- The technique of DTI is explained on p. 3 (first page of the introduction).
Therefore, the explanation on p. 4 “(…) a technique that yields quantitative
measures and directionality of water mobility [15]” is redundant and should be
removed.

>>This has been removed as suggested.

- Finally, the authors might want to stress the timeliness of their method by
mentioning that TBSS is increasingly being used for investigation of white matter
alterations in various diseases like schizophrenia (here, the authors might want
to refer to a recent study by Koch et al., [3].) or
depression (here, the authors might want to refer to a recent study by Kieseppä T
et al., [4]).

>>We have added both references and a short sentence to this effect as
suggested in the introduction as follows: “TBSS is a relatively new analysis
technique that has been applied to a broad range of neurological and psychiatric
disorders to assess the microstructural integrity of white matter [17-19].”

REVIEWER 2:
- Major Compulsory Revisions
1. The DTI acquisition and other parameters needs to be clarified: What sequence was used for DTI? Single-shot EPI? Fat or water suppression What was the matrix size of acquisition? What was the b value? How many non-diffusion weighted images were acquired?

>>We have completely rewritten the DTI acquisition and analyses section to address these questions, and those below. For example, the parameter section now reads: “We employed a single shot EPI sequence. The DTI data was collected along the anterior commissure/posterior commissure line, with FOV = 256 x 256 mm, 128 x 128 matrix, slice thickness of 2 mm (isotropic 2 mm resolution), NEX = 2, TE = 92 ms, TR = 10000 ms. We used 12 gradient directions with b = 1000 s/mm². The total acquisition time was 4.32 minutes. The DTI experiment was repeated twice to increase signal-to-noise ratio.”

2. DTI post processing
a. It is not clear how the two sets of DTI images were averaged. Any alignment was prescribed for correcting motion artifacts before averaging? What algorithm was used for tensor fitting? linear least square, non-linear or something else?

>> In the rewritten methods section, these questions are addressed as follows: “Conversion to nifti: The dicom files were converted to nifti using the dicom2nii program (www.sph.sc.edu/comd/rorden/dicom.html). This program also outputs the gradient direction tables after correction for image slice orientation and a b-value table. The two DTI experiments were concatenated into one 4D nifti file and a concatenated table of corresponding b-value and gradient direction tables were also concatenated. Eddy current correction: Eddy current correction consists of registering all the images to a b = 0 s/mm² diffusion image. We used FLIRT (FSL) with a mutual-information cost function for this step. The algorithm registers images of both the DTI measurements to a common image. The data is not averaged for the next step. Calculation of diffusion tensor: The diffusion tensor, scalar diffusion parameters (MD, AD, RD, and FA) were calculated by DTIFIT (FSL).”

b. TBSS method: Did the authors use the TBSS module in FSL package for the analysis? If so, what version of TBSS and FSL was used?

>>We used TBSS version 1.2 in the FSL package. This is now stated in the methods section.

c. For the FA TBSS analysis in FSL, each FA image was “non-linearly” warped to the template or the most representative FA image (target image) in the study population, and the transformation matrices were then applied to MD and other non-FA images. The text on page 6 stated “All subjects’ fractional anisotropy (FA) and mean diffusivity (MD) images are aligned to a target image, then the
entire dataset is aligned into 1x1x1mm3 MNI152 space” did not reflect correctly
the methods used in the TBSS module of FSL, unless the authors had modified
the method. If that is the case, the methods need to be stated clearly.

We appreciate the inconsistency that the reviewer has noted. We have
adapted the methods in the text accordingly to reflect our approach: “Image
registration for group analysis: “The fractional anisotropy (FA) image of each
subject was normalized to a 1x1x1mm3 FA template (FMRIB58_FA_1mm) in the
Montreal Neurological Institute (MNI) space using the non-linear registration
algorithm FNIRT (FSL). The spatial normalization transformation obtained by
registering FA was then applied to other diffusion images (MD, AD, RD).”

d. Similarly to “c”, only the mean FA image was “skeletonized” in the TBSS
module of FSL, and all the MD images are projected onto the original mean FA
skeleton images, i.e. using the original FA, but not the MD, for finding the
projection vectors,. If the authors used the FSL package without modifying the
methods of FSL toolboxes, the statement on page 6, “The mean FA or MD image
is “skeletonized” to reflect common tracts across all subjects, and individual
subject data are projected onto the skeleton.” would be incorrect and need to be
rewritten.

We have also adapted the methods here to further clarify our approach:
“Image skeletons for group analysis: A mean FA image was calculated from the
mean FA images of individual subjects. The white matter regions for this mean
image were skeletonized using TBSS Version 1.2 (FSL) [5]. A threshold of FA >
0.2 defined the white matter regions. Values of FA of each subject were then
projected onto the common skeleton (TBSS). The standard TBSS algorithm in
FSL was used for this purpose. It consists of doing a search in the direction
perpendicular to the skeleton and assigning the maximum value of FA to the
skeleton. The spatial coordinate of this maximum FA value is noted and
skeletons of MD, RD, and AD images are calculated by assigning the
corresponding diffusivity values to the skeleton. At the end of this step we have
skeleton images corresponding to FA, MD, RD, and AD for each subject. The
spatial map of the skeleton is the same for each subject but the values it takes is
subject dependent. All processing is done by standard TBSS algorithms and
further explanation of TBSS is described in [5].”

e. What was thresholding value chosen for skeletonized FA?

A threshold of FA > 0.2 defined the white matter regions.

3. Statistical analysis
a. What were the age ranges of each group, i.e. NPSLE, SLE, and healthy
controls? The wide age range (18-60) with only 17-20 subjects/groups risks
introducing artifacts. Even though there was no significantly different in age
between three groups, the confounding effect introduced by age (which is a known confounder in DTI measures), and possible the gender, would needs to be considered in GLM model.

Controls, SLE, and NPSLE subjects were not significantly different with regard to age. In spite of this, and since (as the reviewer notes) age is a confounding factor in age effects on DTI measures, we added age and gender as nuisance variables to the model.

b. It is not clear in the statement “Our approach was to first compare NPSLE and SLE patients to controls, then acute NPSLE patients to controls, and finally acute NPSLE patients to SLE patients.” on page 6 whether 3 levels in the “group” predictor in a GLM analysis with post-hoc group comparisons was used, or a 2 levels (2 sample groups) GLM modeling, repeated in 3 separate models were applied for evaluating the three group differences. In addition, the degrees of freedom in GLM for FA or MD group comparisons need to be reported in the result session in order to justify the correct use of statistical models.

We appreciate the reviewer’s confusion regarding the models. We conducted three independent contrasts: Controls vs. SLE, Controls vs. NPSLE, and SLE vs. NPSLE. We now report degrees of freedom (df) for each of these contrasts in our results section.

c. It is unclear whether the final statistical significance was referred from the parametric or non-parametric statistical analyses. First, the statements of three P values on page 6 are redundant. The statement of “voxelwise cross-subject statistical analyses are performed using the general linear model, with (FWE) corrections being made for multiple comparisons (p<.05). Additionally, we conducted 5000 Monte Carlo simulations to determine whether our results could have arisen by chance (p <.005)” indicates both parametric and non-parametric tests were used. However in the later sentence, “5000 Monte Carlo simulations to determine whether our results could have arisen by chance (p <.005) “ and “Only data that passed both FWE and Monte Carlo simulations and was significant at p < .01”, indicating a non-parametric permutation test using 5000 Monte Carlo simulations and only FWE-corrected P value less than 0.01 are accepted, or maybe something else. Furthermore, the “p <.005” is confusing and irrelevant. Secondly, it is not clear whether “voxel-based” or “cluster-based” thresholding was used for correcting for multiple comparisons in the permutation test. In the text, it states “Only data that passed both FWE and Monte Carlo
simulations and was significant at p < .01”, which refers to the null distribution of the max voxel-wised permutation tests, but the legends in Table states “Voxels=significant continuous voxel cluster”, which implies cluster-based thresholding. If clustering was used after permutation test, then the clustering criteria need to be mentioned.

>>We thank the reviewer for clarifying this point. We used cluster-based thresholding for this analysis. The rewritten methods section should leave this analyses better clarified: “Statistical group analysis: We assessed group FA differences using FSL’s General Linear Model (GLM) tool. Age and sex were entered into the model as nuisance variables. We thresholded the t-statistic images at t > 3.0 as recommended by FSL. The group mean differences (two-tailed) were tested using permutation methods with FSL’s Randomise. We ran 5000 two-tailed Monte Carlo permutation tests for each of the group differences. All presented results are corrected p-values at p<.05 after controlling for family wise error rate. Next, we created a mask image for significant FA clusters by binarizing the FA image for results that were significant at p<.05. Our approach was to first compare NPSLE and SLE patients to controls, then acute NPSLE patients to controls, and finally acute NPSLE patients to SLE patients.”

4. Results
a. Has the other tail of t-tests been evaluated as well, e.g. lower FA and higher MD in healthy controls compared to SLE patients and so on? If so, any significance?

>>All analyses were conducted 2-tailed, and there were no significant results in the “other” direction.

b. The “anterior” corona radiata in Fig. 1 could be “superior” corona radiata, which can not be differentiated without performing tractography and fiber tracking.

>>Your are correct, although, this is how this structure was classified in the ICBM template provided in FSLView. We did not perform post hoc tractography; therefore, we are unable to infer fine grained tract identifications as would otherwise be the case.

c. Apparently, the significant clusters in Fig. 1 have been dilated, but it was not stated in the text..

>>We now state at the very end of the methods: “Significant clusters were dilated for figure presentation.

5. Discussion
a. “Interestingly, there were no significant FA or MD differences observed
between the 16 SLE patients without NPSLE and the twenty matched controls, many of whom had subcortical white matter and periventricular lesions” on page 8. Should the “many of whom had subcortical white matter and periventricular lesions” refer to healthy controls, but not SLE patients or vice versa? If so, the sample needs better description and justification.

>>We have adapted the wording to reflect the sample better. The sentence now reads: “Interestingly, there were no significant FA or MD differences observed between the 16 SLE patients without NPSLE and the twenty matched controls, many of the SLE patients of whom had subcortical white matter and periventricular lesions.”

b. The authors argue that the “none of these regions were identified on radiological scan to be regions of old infarct in NPSLE” by showing the examples of some proton density images from NPSLE. It has been shown that T2-weighted FLAIR imaging is one of the most sensitive image modality, in detecting the white matter hyperintensities (WMHs), and is superior to proton density imaging, in which WMHs might not be demonstrated. In fact, the WMHs shown in the supplement file are likely to cover the voxels of the body of corpus callosum and corona radiata, where DTI measures were different between NPSLE vs HC and SLE group. The easiest way to verify would be averaging all the transformed PDI images in the MNI template space and check if WMHs voxles match the clusters where DTI measures are different between groups.

>>This is an interesting suggestion that the reviewer makes regarding potential overlap. We are currently unable to perform this analysis and I am unaware of a technology that would allow for quantitative analyses of these lesion/FA intersections; however, this is a great idea for a subsequent paper.

c. The authors conclude that low FA and high MD in NPSLE is likely the manifestation of edema in acute NPSLE stage. In fact, it has been shown that cytotoxic edema in acute stage would restrict free water movement, and thus resulting in a decreased MD, and possibly an increased FA.

>>We have taken out the following sentence which should leave the “many” alternate possibilities intact and not infer undue importance to edema: “Indeed, there is an emerging appreciation for subtle edema being a predominant characteristic of NPSLE, which appears to correspond well to the increased MD findings [6].”

- Minor Essential Revisions
1. Font size mistakes: 1x1x1mm3 on page 6.
2. State the MNI coordinates for each cluster, i.e. MNI (x,y,z) = (-8, 5, 25), (-12, 6, 27) separately, on page 8.
3. “relatively young (<=60) on page 8 in Discussion”, where (<=60) should be changed to “less than 60 years old”

>>These minor corrections have been made in the text.