OPEN PEER REVIEW REPORT 1

Name of journal: Neural Regeneration Research
Manuscript NO: NRR-D-20-00129
Title: Protective Effect of Diffusion Tensor Tractography on Surgical Outcome of Cerebral Lesions Involving Arcuate Fasciculus
Reviewer's Name: Maria Marcella Lagana
Reviewer's country: Italy
Date sent for review: 2020-2-26

COMMENTS TO AUTHORS

The Authors virtually reconstructed the left arcuate fascicle of 65 patients with cerebral tumor lesions involving arcuate fasciculus (AF), using diffusion tensor tractography. The AF volume changes (comparing pre- and post-surgery) significantly correlated with the language function changes. The cases with and without language changes after surgery were evaluated relating to the minimum distance between AF and lesion.

The materials of this paper are interesting, i.e. 65 patients with a tumor that involves the left AF, evaluated with DTI and tractography before and after surgery; patients' language functions were evaluated using Western Aphasia Battery. The work is potentially very interesting, but it should be extensively revised and many points should be clarified. Some additional statistics and a more clear structure could help the reader to understand what has been done and the importance of the findings. Hereafter my notes for possible improvements to the work, given in order as the manuscript text is written.

- The generic term "cerebral lesions", used in the title and in the abstract, does not allow to understand the kind of lesion
- The abstract should be extensively revised because it does not allow the reader to appreciate what has been done. The aim of the work is not clear. Reading only the abstract, the reader does not understand if the three last lines are a future perspective or if it is a conclusion of the current work. Lines 33-39: wasn't the Spearman's correlation performed between AF volume change and language function change (line 45 page 8)? Writing "language deficits with arcuate fasciculus injury fraction" seems a single time point evaluation. The sentence "Diffusion tensor tractography played a positive role in minimizing the impairment of language function during the surgery" is misleading: the Authors couldn't know how language function could be impaired without tractography and it is not clear if they think that performing the presurgical tractographic planning itself protects the AF or if the language improvement after surgery is due to the presurgical AF impairment (evaluated with Dmin in this work).
- Globally, the results are difficult to understand because many of them can be found only in the discussion. Almost the whole sections "Possible factors that Affect the Protection of Intraoperative Language Function" (lines 9-44) and a part of "The DTT in the prognosis estimation of the postoperative language function" (lines 43-48).
- It would be interesting to investigate if also the volume of AF occupied by the lesion is related to the language function impairment and if it is predictive of the language improvement. In some stroke studies, the fraction of the involved tract predicts the functional improvements after surgery (i.e. Riley et al, Stroke 2011 for corticospinal tract and motor recovery). The same could be here. Was this tested, besides how much the AF volume changed after surgery?
- If I well understood, the language function tests were performed before surgery and at two time points after surgery. Conversely, tractography was performed only twice (before-after surgery). The results of

these evaluations are sparsely reported, in the result section and in the discussion, and subdivided by the kind of lesion. It would be easier for the reader if the results would also be summarised in a table, reporting all of them at all the time points (language function scores, volumes, Dmin).
- The work investigates "65 patients with cerebral lesions involving AF" (line 28 page 4). However, the minimum distance between the lesion and the reconstructed AF is computed (Dmin, line 28 page 6), so it is different from zero: how was a subject defined as with "a cerebral lesion involving AF", if Dmin is >0? Is there a maximum Dmin in order to be considered as lesion involving AF?
- Regarding the DTI sequence: why wasn't it isotropic? Please comment on this and discuss if and how this could affect the tractographic results. How was the phase encoding direction? AP/PA/RL/LR? DTI pre-processing: how were the distortion corrected? Since the sequence was not acquired using two opposite phase encoding direction, please discuss if this is a limit and how it could affect the shape of the AF tractography.
- Tractography: 1) Why was 0.15 chosen as FA threshold? 2) A Figure with the 3 ROIs and an exemplificative AF reconstruction is warranted. 3) It would be better to show also the Dmin measure for an exemplificative case. How was it manually identified (line 31 page 6)?
- The language was evaluated after 2-4 weeks and after 3-6 months but it is not clear if RAQ was computed for both the time points and how it was used.
- Why the two indices Rv and RAQ were computed using the absolute value of the pre-post difference? In such a way, how was it possible to differentiate an improvement compared to a worsening? If these measures were without the absolute value, the negative sign would indicate an AF volume reduction for RV and positive/negative sign would mean worsening/improvement of the language function. Performing a correlation analysis without keeping the sign of Rv and RAQ could bring to wrong conclusions and small changes could confound the statistics.
- Reading the results, it seems that the RV and RAQ score damage ratios (1-5) were never used. They appear only in figure legend 3 and in the discussion. There, it is not clear what it means that 1 is true negative and 5 is true positive. So, it is not clear how the ROC curve was obtained and how to interpret it.
- Given the last two comments, the results were difficult to interpret. What does it mean that the language function level improved or exacerbated (lines 33-36 page 7)? RAQ is always positive (see my last two comments).
- Results, page 6: were the long term language dysfunction 2 out of the 15 cases that exacerbated after surgery?
- It is not clear which Spearman's rank coefficient is reported in the results of glioma patients (line 31 page 8). Here Figure 1 is referred to, however, this correlation is not explained in the methods and it regards prevalences (number of patients that improved or remained stable or worsened, split according to 3 Dmin ranges. So, Spearman's correlation should not have been used here.
- Results from line 40 page 8: 1) the sentence written in the results at this point "The correlation analysis between RV and RAQ was performed in 24 patients with perioperative language changes and the result showed positive correlation" it is clearer than the one written in the methods section (line 9 page 7) "Linear correlation analysis was used to evaluate the relationship between RV and RAQ". The latter can be joined to the description of the following Spearman's correlation test (line 14 page 7). 2) Differently from what was written in the methods, in the results, it is written that the correlation was tested only among the 24 subjects "with perioperative language changes"(line 42 page 8). Explain in the methods which patients were selected, why and how: what does it mean with "language changes"? Did they use the language change score? Also, in the discussion section there are the same two sentences (end of page 11 and beginning of page 12). 3) The results of the ROC curve are written in this way "ROC was used to evaluate the accuracy of the RV prediction for RAQ (Figure 3)." But the AUC and p-value were not written. Moreover, neither in the method section is clear how the ROC curve was obtained. A sort of description is written in the discussion section (lines 39-48 page 12), but it should be moved to the methods and rephrased. At the same lines, there is also written how much the

AUC was and this should be written in the results.
- Two illustrative cases were described, but it was not written why those two have been selected. The Dmin, RAQ and Rv of these cases were not reported but would be useful.
- Discussion: Please insert some references for supporting this sentence "There is still some controversy about whether the white matter fiber reconstructed by DTT is consistent with the actual anatomy". Then, only a study supporting the agreement between DTT and anatomy was reported and discussed. Please report also a study that shows the opposite, since it was written that there is controversy in the literature.
- The message that the Authors want to give with the whole page 10 (discussion) is not clear: they comment about intraoperative electrical stimulation and brain shift. Since they were neither introduced not evaluated in the current study, one page for them seems excessive. Reading page 10 is difficult, for the following reasons. 1) It is difficult to interpret why the authors reported studies such that of the optic radiation (lines 26-33). 2) The sentence about the study of Leclercq et al. is about the language tract but it is not clear and how it helps in the discussion of the current work. 3) Conversely, it is useful to know that some studies verified the reliability of DTT with subcortical electrical stimulation, but what does it mean that ". The sensitivity of pyramidal tract is 95%, and the sensitivity of the language tract is 97%". Please summarize what is this sensitivity. 4) these sentences seem contradictory: "The combination of intraoperative subcortical electrical stimulation and DTT could reduce the incidence of epilepsy during awake surgery" and "Intraoperative electrical stimulation is currently considered as the "gold standard" for the positioning of the functional cortex and important subcortical functional tracts, but this time-consuming and invasive manipulation has a high incidence of epilepsy in Awake craniotomy". Do the Authors support future studies with the combination of DTT and intraoperative subcortical electrical stimulation or did the Authors write about these studies in order to support the use of DTT only? 5) The use of intraoperative MRI and the reconstruction of the AF after brain shift could help to minimize the AF injury: the Authors mentioned this but they did not contextualize the sentence in the current work: is this a limit of the work? Is this a proposal for future development? 6) The sentence "The incidence of short-term and long-term postoperative language function deficit is similar to that reported in previous studies (Sanai et al., 2008)." is written after the sentence about intraoperative MRI for correcting the brain shift. Why? Are they linked? Is this sentence talking about the postoperative incidence of language deficit found in the current work? If so, it would be useful to report them and those found in the cited work (Sanai 2008), comparing and discussing them: are there methodological differences between the current work and the one of Sanai, so do the Authors want to comment that the incidences are similar anyway?
- More than half of the section "Possible Factors that Affect the Protection of Intraoperative Language Function" of the discussion is occupied by results. Which Spearman's rank correlation is shown here (line 42)? What is the coefficient of this correlation? The last 4 lines of this part are discussion and can be kept here, but the sentences should be deeper justified: 1) why "it is important to reconstruct the AF according to the intraoperative DTI for glioma patients with D min ≤ 5mm"? This sentence could be essential for demonstrating the importance of the current work, but it is pending. In the methods section, a paragraph regarding the surgical approach is missing. Were the AF tractography used during surgery? How? Was a navigator system used? Do the Authors hypothesize that this approach allowed preserving the AF as much as possible? 2) Also, the following sentence is difficult to understand in this context ("Repositioning residual lesions and AF based on the intraoperative images after updating navigation is conducive to retain the AF."), because intraoperative images were not used in the current work. It could be written as potential future development of the work if correctly supported by literature and contextualized.
- The term "accuracy" was used in the section "The DTT in the prognosis estimation of the postoperative language functions" and at the beginning of the "Conclusion". However, the accuracy of the AF DTT was not evaluated in the current study. So, why did the Authors wrote: "According to previous researches and our results, The AF reconstructed by DTT is accurate and reliable"?

- The two sentences written in the discussion at the end of page 11 and beginning of page 12 are the same written in the results (page 8) and it is not clear why they were reported here again.
- The following sentence (lines 4-6 page 12) about the study of Hayashi et al 2012 could be joined to the previous sentence about DTT accuracy studied in the literature.
- Lines 6-9: is reported here another result of the current study or is this a result of Hayashi's study?
- A part of the conclusion was not fully supported by the current study data. As I wrote above, a paragraph describing the surgical approach is missing and might allow understanding why the Authors wrote "The AF functional neuronavigation based on DTT is helpful to minimize the injury of language function of patients and improve the prognosis of patients while maximizing the resection of lesions." Although this is potentially an important sentence, it is not clear which are the results supporting that the neuronavigation based on DTT minimized the language function injury and maximized the resection?

Minor:
- Please refer to at least one of the works published by Catani and Thiebaut de Schotten on AF in the introduction, where DTT for basic researches on the anatomical structure is mentioned.
- Typo: optic radiation (page 4) instead of optic radiation
- At line 28 page 4 it is written "PLA general hospital" without the city, which is written on the next page. Note: the paper was not fully blinded.
- Please substitute "millisec" with "ms".
- Please provide a reference for the common color-coded mode
- Please could the Authors explain what they mean with the following sentence? "The peripheral closed curve formed by all AFs passing through each level were connected and then displayed as a three-dimensional (3D) object."
- Line 56 page 6: the following sentence is the same at line 31 page 6 and it was wrongly pasted at line 56: "The volume change ratio of AF was calculated according to the following formula."
- Verb is missing in the following sentence "The coefficient of correlation $r = 0.895$ ($p < 0.001$)."
- The term "functional" referring to AF DTT in the conclusion is not appropriate. The DTT is a structural technique.