Dear scientific editor Mrs. Chap and dear reviewers,

Thank you for considering our study for publication in your journal. The authors declare that they have no competing interests. We also wish to add our thanks to MSc Ullamari Hakulinen in the acknowledgements section, for her assistance with some corrections.

Please find our answers to the reviewers’ comments below. Although most of the questions were suggestions or recommendations only, we have tried to give at least some answer to each one of them. After the modifications, the manuscript underwent a renewed language checking.

Best regards,

Maija Rossi

Major Compulsory Revisions (which the author must respond to before a decision on publication can be reached)

Lars T. Westlye:

1) My main concern is related to how the authors deal with the large number of statistical tests performed. The lack of corrections for multiple comparisons seems remiss and undermines the interpretations and validity of the results. For example, Table 2 shows the results from 52 correlation analyses, but the authors still regard an uncorrected p < .05 as statistically significant. Post-hoc corrections (e.g. Bonferroni) should be performed to control for Type I errors. This should also be performed for all other analyses presented in the paper, including the correlations between lesion volume and clinical test data.

We have now included the post-hoc correction with the number of ROI comparisons as the factor. Accordingly, we do not highlight as many results as
statistically significant anymore. This has led to several modifications in the abstract, the methods and results section, Table 3 and finally, the discussion.

Hans-Jörg Wittsack:

1. In acute phase of stroke CT investigations were performed. Why? Stroke lesions are often not detectable in CT if symptom onset is less than 24h. Diffusion weighted MRI is more reliable in acute stroke imaging to determine the lesion volume. Further, in this context the lesion growth described in this work may be a result of the comparison of acute CT with MRI in chronic phase. Please discuss.

CT investigation was performed as part of the university hospital's clinical routine. An additional study for scientific purposes was permitted by the ethics committee for the chronic stage. We acknowledge that acute lesions may not be detectable on CT and therefore did not exclude patients whose CT finding was negative – indeed, 7 such patients showed a lesion later on MRI. This may affect our results. However, the only comparison to volume change was performed with neurological findings and our correlation with NIHSS conformed to previous literature. We discuss this in the penultimate paragraph of the discussion.

As an ongoing project in the Stroke Unit of our university hospital, we routinely image patients with hyperacute infarctions (i.e., less than 3-4 hours from the onset) using conventional CT, CT angiography, and perfusion CT. The conventional CT is used to rule out haemorrhage, visible acute infarctions, and significant microangiopathy. CT angiography is performed to image acute carotid or vertebral artery dissections or blocks. Perfusion CT is performed to rule out perfusion disturbances in the suspected ischemic area. Our initial experience has been positive regarding the CT diagnosis in the hyperacute stage. We have now explained this more thoroughly in the "Time frame of the study" paragraph of the methods section and also comment on our initial experience in 4th paragraph of the discussion.

2. FA and MD values were measured in a set of standardized ROIs as well as within the lesion. Is there any dependency between localization of the lesion and regional changes of FA and/or MD? In particular: Are the regional alterations larger in regions of afferent fibers? A fiber tracking analysis should be possible based on the measured DTI. Maybe this could deliver additional information.

In our study until now we have used a conventional DTI software (Neuro 3D, Siemens, Erlangen, Germany) for all possible DTI analyses and tractography. However with this software the problem arose with the crossing fibers. For this reason we are currently starting using diffusion spectrum imaging (DSI) and TrackVis software to get more information about the crossing fibers and anatomical regions. We are attaching two examples of 3D tractography images taken from the area of infarction.

In patients with chronic infarctions all possible tracts in the area of lesion disappeared. In addition, in patients with visible Wallerian degeneration the fiber tracts in the area of mesencephalon were distorted. In our study no dependency between location of lesion and regional changes of FA and/or MD was seen.
FA/MD changes were seen in any location which had suffered from the ischemic or infarction changes and associated Wallerian degeneration. This information is now added in the paragraph "DTI results". Greater details of this regional correlation will be studied in our next study.

3. The mean diffusivity MD is very comparable to usually measured trace DWI. The temporal evolution of DWI signal with stroke lesion evolution is well studied in literature. What is the effect of temporal changes of MD in your data of patients in chronic phase?

It is difficult to comment on the temporal changes in the chronic stage because we have studied the chronic-stage DTI images only at one time point. In literature we found detailed information on DWI signal evolution during the first weeks after stroke. At least MD seems to continue changing at least six months from the onset (or longer). However, data on 6-12 months was not found. This was the reason we decided to exclude the comparison of DTI and neurological tests which were not performed simultaneously. We have expanded the last paragraph of the discussion and added there references [6,20,31,32]. If you can suggest us another reference we are unaware of, we are happy to revise the manuscript again.

In our new ongoing study we are currently examining patients at 7 days, 6 months and 12 months.

4. Methods => DTI Analysis:

ROI size is 4 to 16 voxels. The ROI size of 4 voxels is very small. In the usually noisy DTI data a dislocation of a small voxel leads to large changes in the results. How did you handle with this problem and the objectivity? Further, the anatomical quality of DTI (FA, MD) is poor. How were the ROIs placed reliably in the DTI maps (in particular the small ROIs)? Was an image overlay technique with high resolution anatomical data performed?

4 voxels were used in capsula interna due to the small size of the structure. All DTI images were carefully analyzed by the same reader. 3D rotation was utilized to assure correct localization of the ROIs while analyzing the data, along with image overlay. This information is now added in the "DTI analysis" paragraph of the methods section. We are starting to use DSI which will further improve the correct setting of ROIs.

Discretionary Revisions (which are recommendations for improvement but which the author can choose to ignore)

Lars T. Westlye:

1) The paper would be much more compelling if the authors had provided one or more specific hypotheses guiding their methodological and statistical choices.

We have modified the first paragraph of the abstract and the last paragraph of the introduction section and hypothesize that DTI and lesion volume are associated after infarction.
2) The authors present FA and MD data only. Increased MD could be explained by both increased longitudinal and/or increased radial diffusion, which each would imply distinct biological interpretations. The authors might want to include these additional measures in their analyses or (to keep the number of tests at a minimum) at the very least discuss possible limitations with FA and MD measures.

We have now included discussion on the causes of MD alterations as suggested (5th paragraph in the discussion) and added a reference [23]. This discussion partly affects FA measures also, but FA is further complicated by the problem with crossing fibers for which, as mentioned above, we do not currently have the tools.

3) The patients are in several paragraphs described as having “attention deficits”. In neuropsychological and cognitive terms, this is a vague and ambiguous description and should be operationalized or avoided. If the authors are referring to some kind of neglect or inattention, specific neuropsychological test data, screening instruments or clinical information supporting this interpretation/diagnosis should be presented.

Further, the neurological findings presented in the Results section are restricted to aggregate measures across subjects, with no descriptions of the psychometric properties of the tests or the clinical significance of the results. For example, MMSE is a widely used screening instrument for cognitive impairments related to neurodegeneration including Alzheimer’s disease. Ceiling effects in non-demented samples are expected (as probably seen here, with the lower inter-quartile range of 27/30). It would be interesting to know if the correlation between lesion volume and MMSE (if the authors are able to show that this correlation is not coincidental) is entirely driven by (the few?) patients scoring below some cut-off score and therefore might present cognitive impairments. Since the extreme scores on such screening instruments are of particular interest, please also report minimum and maximum scores.

The detailed neuropsychological results, including neglect syndrome, are currently being analyzed by another author and may not be discussed here for copyright purposes. Therefore, we have now removed the term in the abstract. Also, we now only describe that our patients are “suspected for attention deficit”. However, we feel that it may not be completely removed as these patients may have lesions in both MCA and PCA.

Regarding MMSE, we have now changed the interquartile range into whole range. The whole range was not considerably larger than the interquartile range and the correlation does not seem to depend on few extreme patients (see the new Figure 5 we have now added).

4) Both FA and MD are highly sensitive to age. Given the relatively large age-span of the sample (57-75 years), the authors should include age as a covariate to exclude possible confounding effects of age on the relationships between lesion volume and DTI metrics.

We agree with you, and we have started recruiting control patients aged 57-75
years for a future study. However in this current study and in the structures selected here, age was not correlated with lesion volume. MD and age tended to correlate with left-side centrum semiovale but this result was rejected by post hoc correction.

5) Discussion, fifth paragraph: The authors claim that the negative correlation between the chronic lesion volume and the FA value can be “easily understood; in larger lesions, the tissue is destroyed more thoroughly, leading to smaller anisotropy.” This is an unnecessary simplification of a complicated issue, and I suggest it be removed or entirely reformulated. For instance, in the case of selective degeneration of one axonal pathway in areas of crossing fibres, one would expect to see an increase in FA rather than a decrease. FA is an index of the relative deviation between the eigenvalues of the diffusion tensor, and interpretations might therefore vary according to the underlying tissue and the nature of the neurobiological processes. This limitation to the interpretations of the FA index should be incorporated in the Discussion.

We agree with you. In the images analyzed, the tracks were completely destroyed in the lesion. For prospective studies we are hoping to better be able to answer this question using DSI for the calculation of afferent fibers. The sentence and the paragraph are, however, now reformulated as you suggested (6th paragraph in the discussion).

6) Discussion, sixth paragraph: 1) The correlations between DTI indices measured in the corpus callosum and lesion volume might be reiterated in the results section in addition to the mention in Table 3. 2) It is not entirely clear to me why the correlations between the acute lesion volume and MD measured in the splenium of the corpus callosum necessarily points to a Wallerian degeneration. Please clarify.

1) We have added the mention about the correlation in the text under paragraph "Correlation between lesion volume and DTI results" in the results section.

2) We have re-evaluated our results grouping our patients into two groups. We now see no reason that the correlation points to WD. It is also to bear in mind that this correlation was not very strong (Table 3, p = 0.027) and we should not draw strong conclusions. We have completely removed the sentence in the discussion. Our further analysis on WD is currently in press and not discussed here in more detail.

7) The authors relate the lateralization of the DTI measures to the infarcts only. While this could be the case, existing literature have pointed to significant lateralization of DTI metrics in specific white matter tracts, e.g. related to language. Although the lack of a control group in the present study precludes direct comparisons, the authors should discuss the possibility that some of the hemispheric differences might not be exclusively related to the infarct.

We have now addressed this issue in the second paragraph of the discussion as suggested.

8) The section describing the DTI processing and analysis is on the short side.
For example, did the authors correct for possible artefacts related to subject motion or eddy currents in the DTI data? How was FA and MD defined and calculated?

The equations for FA and MD are added in the methods section. Unfortunately we did not have the software for correction of motion artefacts and eddy currents, but the DTI images were of acceptable quality by visual judgement. We will soon have a software for motion artefact correction.

9) Did the DTI protocol vary between individuals? If it did, the possible implications of this should be discussed and protocol should perhaps be included as a covariate or factor in the statistical models.

The only variations in the DTI protocol were very subtle changes due FOV, due to the size of the patient. These changes were so small that we believe that accounting for these would bring unnecessary complications to the article.

10) Is it not entirely clear to me whether any of the ROIs were placed in close proximity to (or even within) the lesion in any of the patients. For example, patient 29 had a lesion in the internal capsule, which is one of the ROIs. How did the authors address issues concerning partial voluming in these cases?

This was the only exception where the ROI of the healthy tissue was allowed to be smaller than the pre-defined volume for that structure in order to minimize these effects.

11) The authors report a negative correlation between lesion volume and MMSE in the chronic phase. In addition to the issues concerning multiple comparisons (see above), the authors should explicitly test if a common effect of age could explain this association.

We searched into this and we see no effect of age in the new Figure 5 that we have added in the context of the last paragraph of the results section.

12) Title: The title might imply that the analysis included longitudinal DTI data, which is not the case. Perhaps replace “changes” with “imaging”?

We agree with your suggestion and have changed the title.

13) It seems that the authors measured diffusion along 12 gradient directions (I guess this is what “number of excitations: 3 and 12 orthogonal axes” means). Although mathematically 6 directions is adequate for the calculation of the diffusion tensor, there is an increasing awareness in the imaging community that number of directions applied influences the reliability of the estimated tensor and its parameters. Indeed, Jones (2004) states that “at least 20 unique sampling orientations are necessary for a robust estimation of anisotropy, whereas at least 30 unique sampling orientations are required for a robust estimation of tensor-orientation and mean diffusivity”. Although I’m not convinced that the results would change a lot with a larger number of sampling directions, I think this important point needs to be addressed in the manuscript.

Since the initiation of this study we have increased the number of gradient encoding directions up to 20. We have clarified the writing in the methods
section. We have added a notice of the issue along with the suggested reference in the 8th paragraph of the discussion. The figures in the reference show the largest variations at FA = 0.9, but smaller variations at lower FA values. Low FA better describes our study, especially in the lesion area, and hence also the variations due to the number of gradient encoding directions are smaller. Nevertheless we agree with the benefit of more directions.

Eduardo M. Castillo:

The question posed by the authors is not well defined from the get go (abstract and introduction). Writing can be improved.

We have modified the first paragraph of the abstract and the last paragraph of the introduction section.

Hans-Jörg Wittsack:

1. MD and usually used DWI are very related. A detailed discussion of DWI findings in literature and MD of this work may improve the discussion.

We have expanded the discussion in the last paragraph of the discussion section and added references up to 6 months after infarction. Unfortunately the MD in this work was measured at only one time point, which limits the interpretation.

Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

Lars T. Westlye:

1) Table 2: All SDs are negative? Typo?
We have corrected the typo, thank you for the observation.

2) In the Methods section (first paragraph), the authors state that they studied 41 consecutive patients. However, in the rest of the manuscript, this number is reduced to 33. If this is not a typo, please clarify.

We have now clarified the exclusion of 8 patients due to lacking signs of lesions in the chronic-phase MRI in the beginning of the results section.

3) Abstract (Methods): For clarity, the authors might want to add “thrombolytic” before “therapy”.
We have now defined in the abstract that it was thrombolytic therapy given to some patients (n = 10). We have also added a reference for the therapy [14].

4) Results (DTI-results): Please report the unity of the MD measures.
We have added the unit (×10-3mm2s-1) in both the methods and results section.

Eduardo M. Castillo:

Title. The title suggests that this finding is specific of Right hemispheric lesions. The term “right” is unnecessary in the title since can lead to wrong assumptions.

We have changed the title as suggested.
Pg2. when they say “lesion volume and abnormalities in diffusion tensor imaging are individually associated with poor prognosis” it probably should say “are independently associated with poor prognosis” since the main aim of the study is to test their relationship.

Thank you for noting, we have corrected the use of words.

Pg. 2 “This study assesses the correlations between large lesion volumes and diffusion tensor imaging”. Correlation is a term that can be used for two variables like volume and MD or area and FA; but not for a variable like area and a procedure DTI.

We have now changed the use of words into "association".

Pg2. Methods. The authors describe: “The effects of therapy were assessed”. At this point they should mention the type of therapy. There should be a reference to the study of therapy effects in the background if it is part of the method.

We have now defined in the abstract that it was thrombolytic therapy given to some patients (n = 10).

Pg 2. Methods. The authors describe: “values were measured at the site of the lesion and selected white matter tracks”. Later (in page 6) the lack of concreteness continues since ROIs were defined in 13 locations but authors only describe 4 structures in their methods.

The 13 ROIs were the lesion, corresponding area, CC genu, truncus and splenium, and bilaterally centrum semiovale, the cerebral peduncle, the thalamus and the knee of the internal capsule. We agree that there is confusion since only some of these were bilateral measures (and thus calculated twice). We have removed the mention of 13 locations as it is confusing and brings no additional information.

Pg 4. Authors describe the exclusion criteria including “previous lesion in the left hemisphere or the right hemisphere found on acute CT”….why not just saying previous lesion on the brain?

We agree and have corrected the sentence accordingly.

Pg 6. Figure 2. The square shaped ROIs in figure 2 seem to have different area from LH to RH. How that can affect the results?

The example ROIs were drawn while writing the manuscript (after the image analysis) using an image processing tool separate from the DTI software. We did not see the exact area in this separate image processing tool. We have corrected the image (2b) where the red ROIs are now equal-sized as they were during actual image analysis. The size of the ROIs changed due to the size of the structure (e.g. blue area is larger than the red in Fig 2b) but for one individual structure, e.g. thalamus, the size of the ROI was equal over all patients. We have also tried to reformulate the last sentences in the "DTI analysis" paragraph of the Methods section.
Hans-Jörg Wittsack:

1. Methods => Imaging Sequence:
Please insert the imaging settings for CT, e.g. kV, mAs, …
These settings were 120 kVp on both scanners; and 430 mAs on Philips and 320 mAs on GE. We have added these under "Imaging sequences" in the methods section.

2. Methods => DTI Analysis:
Please insert the definitions of FA and MD (formulas, citations).
We have added the equations and reference [15].

3. Results:
The initial CT was within 24h after symptom onset in all patients. Please add the concrete time range and mean of the initial investigation.
This was a good point as we realized that we have falsely written the information. The initial CT was performed within 3-4 h after symptom onset. The 24 h was the time between thrombolytic therapy and second CT which was not used in this study. We have now corrected the paragraph "Time frame of the study" in the methods section.

4. Results: "Seven patients did not have any signs...."
Are these the same seven patients that did not show infarctions in MRI? Please clarify.
These were not the same patients. Seven patients were excluded because there were infarctions on neither CT nor MRI; and seven patients were included who had no infarctions in CT but who had infarctions on MRI. We have clarified this in the beginning of the Results section.

5. The found correlations are tabulated in table 3. A diagram of the most relevant findings may improve the readability.
We have added a new Figure 4 to show correlation between acute-stage lesion volume and the MD of both lesion and the corresponding contralateral area.

6. Table 2:
Standard deviations are negative – please correct.
We have corrected the typo, thank you for the observation.