Comment. I commend the authors’s aim to look at the quality and specificity of a TSA compared to TBSS in MDD, and for trying to view DTI and behavioral measures together. This is the right direction/approach to take to better understand the functional implications of structural alterations. I have, however, a few methodological comments and suggestions that I would like the authors to take into account in the revisions.

1. Please expand on the advantages of TBSS vs TSA. In your conclusion please go further in your discussion of what current findings tell you vs TSA findings. The difference between your previous findings and the current paper does not come across and seems to me to be the most relevant part of your work. How does water diffusion affect these measurements for instance?
2. What about TSA and TBSS findings in other psychiatric populations? Does mood state affect these findings?
3. in your abstract please make sure to mention that it was an adult population, along with current mood state, and medication status. Also please labels which brain areas were found to be altered, and implications for future clinical research (given the scope of the journal this is important).
4. Table 1. Please mention number of mood episodes, estimates of global functioning or daily activities, and comorbidities. Also mention min/max duration of illness and range of HAMD scores. For medication status: please state how many participants took more than one medication. Range of scores for all the neuropsychological measures, and exactly which outcome measures were selected. E.g. list learning: is this a composite score, immediate, delayed or recognition score? TOL: which measure did you pick.
5. Please more exact in your description of the neuropsych measures in your methods. Mention outcome measures, normative references, psychometric validation, and provide references from the literature.
6. Figures 1 and 2. Please provide a caption that enables the reader to understand what they are looking without referring to the text (golden rule of thumb). Please provide a table with coordinates, cluster size, p value and labels of major areas of differences. This would apply to all your skeletonized figures.
7. Statistical analysis. Please provide a thorough description of your GLM(s?). for instance did you add all your covariates at the same time or one by one and compared GLMs? How many regressors did you enter? Did you use spss to analyse behavioral data? I hope that these questions induce the authors to be more exact and thorough overall.
8. Results. Please provide statistical parameters as per guidelines. For instance, F values, p values and eta square/effect size for significant findings. When you say you found that neuropsych scores significantly differed between the groups” please provide the direction of this difference. MDD>HC?
9. I may have missed it but I cannot find a table with all the labels of regions/tracts found to be altered. Please address this.
10. Could you please comment on the effect of mood on your DTI and neuropsych findings. Also please explain if you took into account of reaction times or accuracy in your GLMs? And did you exclude participants with very low accuracy for instance (e.g. <50)?
11. Conclusions. Based on the changes suggested here I would highlight even more the differences between TSA and TBSS, and provide more insight into the clinical utility of these analytical approaches in BD and psychiatry in general. What is the future direction of research in this field?

Minor comments

-typo in your highlights: it is reduced FA not “reduce”

-Make sure to “render” your images or highlight sites of relevance for readers who have black/white versions of your manuscript.