This is a well conducted study that aims to compare standard DTI metrics against biophysical models. Considering the growing popularity of NODDI in normal ageing and other literatures, this report is timely. NODDI metrics also appeared to correlate more strongly with cognition in the younger age group. However, I have some concerns about the image processing methods which appear to differ between the main data and those in the supplemental marterial.

Otherwise, this is a promising first draft. The Discussion is well written

Methods

1. It is not clear to me whether an opposite phased encode image was acquired. The correction for EPI is particularly crucial at higher B vals and it is common practice is use the TOPUP functionality in FSL.
2. Why did the authors use ANTS for motion correction correction when the eddy correct toolbox in FSL is the gold standard? How are outliers / signal drop outs from motion / slice to volume motion handled in this study? Correction for eddy current distortions were performed on the dataset in the Supplementary Material. More details of those analyses should be reported as main results.
3. Readers will also benefit from having more information about the NODDI modeling. From practical experience, there are several parameters prior to the model fitting (I.e. brain mask etc) and all of this should ideally be reported for transparency.
4. Please report motion parameters and check for potential group differences between the young and older groups. Again, this is automatically produced from the eddy toolbox in FSL.
5. I wonder whether the stronger NODDI associations seen in the Young group could have resulted from atrophy contamination in the older group. The results would be much more interesting if the authors are able to demonstrate that the superiority of NODDI metrics go over and beyond GM volumes.
6. Have the authors looked at the performance of FA and MD that are fitting using only the b =1000 data?

Finally, I wish to offer two suggestions that will significantly raise the impact of this study:

1. In the original NODDI paper (Zhang 2016), the data suggested that Orientation Dispersion Index could be reliably estimated from single-shell sequence at B=1000. Very few studies have actually verified this and this would be a good opportunity since it is relevant to the overarching aim of comparing single tensor vs multi shell data.
2. It would be highly interesting to see how the hippocampal NODDI associations compare against grey matter volumes. The analyses should be straight forward since the masks have already been generated.