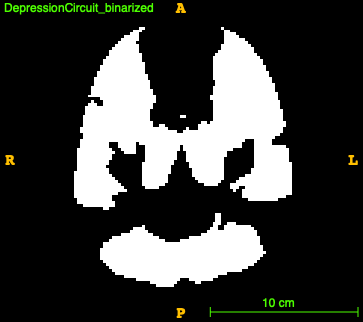
**Reviewer Comments:**  
  
**Reviewer 1:**   
  
*This paper analyzes the structural connectivity disruption of MS lesions in individuals with depression and without depression, and find that people with MS and depression have higher overall burden, overall and specifically to a "depression" network identified previously. The results are straightforward and interesting, but there are a few choices in the analysis that seem tenuous and need to be tested/modified to make sure the original results hold.*

Thank you for positive appraisal of the paper and the opportunity to strengthen the manuscript.

*1) Please justify the use of this threshold "T>3.09 to identify voxels with a statistically significant positive association between depression symptoms and brain disease or stimulation." or quantify how the choice impacts the final results.*

Thank you for allowing us to clarify. The original depression mask is a continuous map, with each voxel representing the association of a dimensional measure of depressive symptoms with functional connectivity at that location (Siddiqi et al., 2021, *Nature Human Behavior*). This creates two challenges. First, the binary mask of all values with positive associations is so large that every fascicle intersects with the mask substantially (see below).



Second, in the original map, there are many voxels where the relationship between depressive symptoms and functional connectivity is not statistically significant. We chose T>3.09 given that this is a standard uncorrected threshold often used in functional connectivity analyses. We agree that this is relatively arbitrary and have repeated all analyses at two additional thresholds, T>2.3 (P<0.01) and T>2.6 (P<0.005). Our results remained statistically significant and with the same directionality in all analyses. For both thresholds, MS lesions preferentially targeted white matter fascicles connecting the depression network, MS+Depression had more disease burden than MS-Depression, and this was driven by disease specifically within the white matter depression network.

T>2.3 (P<0.01):

Main effect of network: (β = 0.048, 95% CI 0.037-0.058, P <0.001)

Main effects of diagnosis: (β = 0.053, 95% CI 0.0085-0.098, P= 0.020)

Network\*Diagnosis interaction (β = 0.016, 95% CI 0.0028-0.030, P=0.019)

T>2.6 (P<0.005):

Main effect of network: (β = 0.084, 95% CI 0.070-0.10, P <0.001)

Main effects of diagnosis: (β = 0.056, 95% CI 0.011-0.10, P= 0.016)

Network\*Diagnosis interaction: (β = 0.021, 95% CI 0.0032 - 0.039, P=0.021)

*2) What is the motivation for adding extra steps to the comparison of the "depression network" and the lesion masks from the MS patients (i.e. fascicles). Why not just overlay the lesion masks directly onto the original depression network and obtain the sum or average of the weights within the depression network? it seems a more straightforward metric that doesnt rely on threshold selection at multiple stages.  
3) Also a simpler measure is the lesion overlap with each of the fascicles, rather than the volume of the voxels with streamlines passing through a lesion. This latter metric doesnt seem to consider how many streamlines' disruption occurs at each voxel. If only one streamline is contained in a "lesioned" voxel is it counted the same as a voxel with 100 lesioned streamlines in it? I would like to see some consideration of the severity of the streamline disruption considered.*

*4) why is sex not included as a factor in the model?*

Thank you for this feedback. We did not model sex given that the samples were matched on sex and men accounted for a small percentage of each group. Thirty-three of 232 (14%) of the participants were male in the MS+Depression group, and 31/148 (21%) in the MS-Depression group. However, there is a possibility that sex differences could influence our results. We have repeated our analyses with sex as a factor in the model and our findings remained statistically significant. We did not find a statistically significant effect of sex.

Main effect of network: (β = 0.090, 95% CI 0.076-0.10, P <0.001)

Main effects of diagnosis: (β = 0.058, 95% CI 0.013-0.10, P= 0.012)

Network\*Diagnosis interaction: (β = 0.022, 95% CI 0.0035 - 0.040, P=0.020)  
  
Main effect of sex: (β = 0.021, 95% CI -0.061-0.020, P = 0.32)

**Reviewer 2:** Comments to authors:  
  
*Excellent work!  
  
On the introduction, there was a lack of epidemiological data of the disease on a global scale. Your data is very interesting and will be aplicable on other countries. I'd recommend including more on this topic.*

We are pleased that you found our manuscript to be of merit and are happy to elaborate on MS disease burden on a global scale. Our introduction now includes the following sentence, and we have also clarified that the rates of depression comorbidity in MS were derived from systematic reviews and meta-analyses from international samples:

“Worldwide, 2.8 million people are estimated to be living with MS, and the prevalence is rising.”

*Are there no data on neurological disability, such as the EDSS, of these patients that could be correlated with the burden of injury in the depression network? A greater number of injuries may be directly related to greater EDSS score in these patients, wich can lead to the development of depression.*

We agree with the Reviewer that testing for relationships between a measures of neurological disability with burden of disease in the depression network could be of interest to the interpretation of this study. A limitation to retrospective studies that utilize electronic medical records for phenotyping is that we are restricted to data that have been collected as part of clinical care and coded in discrete fields. Though the EDSS is often acquired in clinical settings, it is a research measure that was designed to measure the outcome of clinical trials and is not part of the clinical standard of care (Meyer Mooke). Unfortunately, the EDSS is not routinely collected in our MS Center. Additionally, critiques of the EDSS often highlight its subjective nature and poor interrater reliability (Mikael Cohen), and that it does not sufficiently assess mood and cognitive function (Meyer-mooke). Future studies that do rigorous clinical and cognitive phenotyping would significantly enhance the literature.

Do I want to add something about PROMIS? Small sample, no differences between groups in disease differences, so it is not just disability overall?

**Reviewer 3:** Comments to authors:  
  
*This paper presents results from an interesting research concerning the biological correlates of depression in multiple sclerosis (MS). Particularly, Authors investigated the association between white matter lesion load in brain networks and depression in people with MS (PwMS). They highlighted how subjects diagnosed with MS may be prone do the development of depression due to a higher lesion load in brain circuits involved in affective regulation.  
  
Since psychiatric - and particularly mood - symptoms have a high prevalence and a serious clinical impact in PwMS, this paper is of interest to a wide audience. The research is well presented and the results add knowledge to the current literature on the topic The study is interesting and well conducted. The study results further underline how complex, and long underestimated, manifestations of MS can nevertheless be traced back to a dysfunction of brain networks induced by the accumulation of white matter lesions.*

We thank the reviewer for their thoughtful review of the study and its implications.

*However, it must be emphasized that the study has certain limitations that affect the interpretation and generalization of the results (some of which have already been discussed by the Authors, such as the exclusive focus on white matter lesions in MS).*

We agree that there are limitations to the study and are happy to expand on them in the manuscript. *The population of MS patients appears not to be well characterized from a clinical point of view. The impossibility of establishing the clinical phenotype of the disease (progressive, relapsing or both) may represent a bias of the study, also considering the potential psychological effects linked with the presence of higher disability scores in progressive MS. In this respect, it might be useful to include at least the EDSS score and disease duration in the analysis as potential confounding factors.*

Wish we could do EDSS and disease duration. As above, since EDSS is not part of the standard of care, it is not routinely collected in our clinic and is not coded in the electronic medical record. Additionally, establishing duration of disease is challenging due to a number of factors. First, it is well known that people are diagnosed a mean of \*\*\* years after they develop symptoms. Second, because our clinic is a quaternary medical center, people often come after they have been diagnosed or for a presumed diagnosis. We have expanded upon this in the limitations and agree that a future direction would be weaving this important information into clinical care and into the medical record.

*The possible presence of recent clinical relapses and/or the presence of Gd-enhancing lesions in the MRI examination used for the analysis should be specified, given that MS relapses could influence the patient's mood both because of psychological issues and because of the potential influence exerted by inflammatory mediators on brain networks function.*

This is the tough one. Need to clarify that this is lifetime diagnosis, and that there is not a direct correlation between time linking of mood with GD enhancement. It is important to do causality studies, to do mood ratings, etc. Could we address by saying next study is a causality study?

We are happy to clarify and expound on this. In the MS clinic, MRIs are routinely collected yearly on participants, as well as if they have new symptoms. The automated white matter lesion segmentation cannot distinguish between new lesions or scars. Great opportunity for future research to use NLP to review records, label gad enhancing lesions, or develop methods that can specify between the two.

*The potential effect of MS therapies on mood disorders should be better considered. Not only steroids have been associated with mood disorders, but disease modifying therapies such as IFN beta are also linked with worsening of depressive symptoms. If possible, these issues should be considered in the analyses, or at least discussed as study limitations.*

The author is correct in that numerous MS therapies could contribute to psychiatric side effects, including steroids (Schippling).

11 people in total were on interferon, MS+Depression = 6; MS-Depression = 5. Prednisone (2/1), Methylprednisolone(3/3) = Total = 20.

Sensitivity analysis removing people on methylpred/pred/interferon B (lose 20 people, 5% of sample; MS+Dep = 221 (-11, 4.7%); MS-Dep = 139(-9, 6%))

Main effect of network (β = 0.09, 95% CI 0.070-0.10, P <0.001)

Main effects of diagnosis (β = 0.05, 95% CI 0.001-0.095, P= 0.046)

~~Main effect of interaction (β = 0.018, 95% CI 0.00 - 0.030, P=0.060) -> now trend~~

*Authors state that depression is associated with a doubled suicide risk in PwMS when compared to those who do not suffer from MS. Since depression is not the only factor associated with suicide in MS, I believe it would be more useful in a clinical perspective to underline how suicide rates vary between PwMS who suffer from depression and those who do not.*

Happy to do this. Turns out, very understudied!

*It would be useful if Authors could add some statements in the Background section concerning the biological framework in which depression in the course of MS should be placed and quote related literature (see e.g. Menculini et al. JNNP, 2023).*

Happy to!

*In the Methods section, I would suggest to describe why subjects with bipolar depression were excluded from the study.*

Happy to! Bipolar depression excluded because Shan excluded from the nature paper (Siddiqi 2021). Map was convergence of poststroke data (padmanaban), TMS for treatment-resistant depression (Siddiqi American journal of psychiatry), and DBS for depression. 7 of the 14 studies targeted MDD. Additionally, different networks have been implicated in depression versus bipolar disorder (Xiaobo Zhou) even in asymptomatic bipolar disorder (Rai). Not yet any powered studies on TMS for bipolar disorder (Bi et al). Future studies looking at dimensions of psychopathology in this model.

*In the Discussion section, Authors mainly focus on MRI findings, which is coherent with the paper main aim. Since depressed MS subjects also resulted to have lower quality of life and functioning as resulting from the PROMIS administration, I would anyway evaluate adding some comments about this, as well as on the prevalence of depression in the considered sample.*

Happy to add more info on the PROMIS data, including prev of depression in the smaller sample.

Chi sq to see if there was a greater prevalence of depressed patients in the group that got PROMISE

49 PROMIS scores in depressed group (out of 232); 36 PROMIS scores in nondepressed group (out of 148)

X-squared = 0.21, p = 0.65, not significant

*I would avoid citing the prevalence of MS in the Background section of the Abstract, focusing on the prevalence of psychiatric and depressive symptoms and introducing the aims of the study.*  
 Easy peasy!