Reviewer’s report

Title: Predictors of early-onset post-ischemic stroke depression: a cross-sectional study

Version: 0 Date: 29 Sep 2017

Reviewer: Evan Thacker

Reviewer's report:

In this article, the authors present an analysis comparing 36 stroke patients who developed depression within a few days after the stroke, with 47 stroke patients who did not develop depression. The various measures the authors used are appropriate. The statistical analysis could be improved - see Major Comment #1 below.

Major comments

1. Abstract: In the conclusions section of the abstract, the authors state that their findings "might be useful for clinicians in recognizing and treating depression." I doubt this conclusion can be realized; in my view it stretches beyond what this particular paper has to offer. I don't think there are any immediate clinical implications of this work. A more reasonable conclusion might be something like "These results can direct researchers towards possible biological mechanisms that may explain post-stroke depression, leading to further efforts to elucidate and intervene on those mechanisms to prevent post-stroke depression." This suggestion is also relevant for the Introduction section and the Discussion section, final paragraph and conclusions paragraph.

2. Introduction: As part of their rationale for the study, the authors state "predicting the occurrence of PSD after primary treatment is important not only for counseling the patients about the disease prognosis but also for applying additional treatment." I disagree. While I think this work is interesting for elucidating possible biological mechanisms of PSD, I do not think this work is important for recognizing or treating PSD. We can recognize PSD by screening for it directly with a depressive symptom screener, and we can treat PSD appropriately even without this new knowledge about correlates of PSD. So, I suggest the authors consider more carefully how to express the rationale for their study.

3. Methods: The statistical analysis approach could be improved. By using a P value threshold to screen variables in unadjusted analyses for inclusion into multivariable analyses, the authors miss the potential for variables to be non-significant in the unadjusted analyses but then BECOME significant in the multivariable analysis. A better method would be to include ALL variables in the multivariable model, and compare the fully adjusted odds ratios with the unadjusted odds ratios. In justifying their approach, the authors stated on
page 9 line 157 that the multiple logistic regression model was meant to "minimize confounding" and to examine "independent contributions" of the variables. However, to minimize confounding the authors should consider including ALL variables in the model, not just those that were P<0.1 in unadjusted analyses; and the authors may need to note that when they say an association is "independent" what they mean is that it is "independent of what we adjusted for." All that said, the study has a fairly small sample size, which would make us uncomfortable with building very complicated statistical models. With only 36 cases of depression, it might not really make sense to build models with a dozen or more variables.

4. Tables 3 & 4: If the authors choose to act on my suggestion above about which variables to include in the multivariable model, then they should update Tables 3 & 4 with additional ORs and 95% CIs.

Minor comments

5. Title: I think "Correlates of …" would be more appropriate than "Predictors of …" because in this cross-sectional study it is hard to know whether any of the factors "predict" the occurrence of depression; the factors are present around the same time the depression occurs, but the time sequence of events, and the causal pathways, are not clear.

6. Abstract: Including much more numerical data in the results section of the abstract would be helpful for readers. To accommodate this, consider shortening the methods section of the abstract.

7. Abstract: In the conclusions section of the abstract, there is no need to repeat the results. Instead, immediately state the most important implications of those results.

8. Methods: Inclusion criterion #3 was that patients were admitted to hospital for less than 24 hours after stroke onset. Do the authors mean to say that they only included patients who were admitted within 24 hours of stroke onset? Or do the mean to say that they only included patients who were admitted and then discharged within 24 hours of stroke onset? Either way, it would also be helpful for readers to know why the authors used this inclusion criterion. What was the rationale?

9. Methods: In page 8 line 135 the authors state CT or MRI was done within 24-72 hours of admission, however in page 8 line 147 the authors state CT was routinely done in the emergency room. These statements seem to conflict on the timing of CT scanning. Which statement is correct?
10. Results: the heading "univariate logistic regression analysis" should probably read just "logistic regression analysis," as the section includes results of multiple logistic regression also.

11. Results: It is not clear to me why the authors built models separately for demographic/clinical characteristics and for biochemical indicators. Why not consider all these variables together?

**Are the methods appropriate and well described?**
If not, please specify what is required in your comments to the authors.

No

**Does the work include the necessary controls?**
If not, please specify which controls are required in your comments to the authors.

Yes

**Are the conclusions drawn adequately supported by the data shown?**
If not, please explain in your comments to the authors.

No

**Are you able to assess any statistics in the manuscript or would you recommend an additional statistical review?**
If an additional statistical review is recommended, please specify what aspects require further assessment in your comments to the editors.

I am able to assess the statistics

**Quality of written English**
Please indicate the quality of language in the manuscript:

Needs some language corrections before being published

**Declaration of competing interests**
Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?
3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

I declare that I have no competing interests

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal