Reviewer’s report

Title: Reduced serum levels of pro-inflammatory chemokines in Fragile X Syndrome

Version: 0 Date: 12 Nov 2019

Reviewer: Cristan Farmer

Reviewer's report:

Dear Dr. Madsen,

Thank you for the opportunity to review the manuscript, "Reduced serum levels of pro-inflammatory chemokines in Fragile X Syndrome." It is important for me to note that I do not have expertise in immunology or biology more generally, and my remarks are based on my expertise in neurodevelopmental disorders and in statistics and methodology. The authors report the results of a small study of individuals with FXS and non-FXS controls. Two approaches to analyzing the data converged upon several chemokines which were decreased in FXS relative to controls; a clustering method suggested that an additional subset of analytes were associated with diagnosis. The report is clearly written and the methods are generally sound. My primary concern is that the authors focus on the novelty of the analytic approach as a primary contribution, but they do not interpret and synthesize the results extensively enough in the discussion. I describe that and more minor concerns below, in no particular order.

1. The authors imputed zero for <LOD. This is one of many approaches to this inherent limitation; I wonder whether the results were robust to the choice of approach (e.g., consistent results when imputed with the LOD itself or 0.5*LOD or quantile regression imputation for censored data).

2. Extreme values are common in this type of data, so the minimum and maximum may not adequately characterize the spread of the data. I would prefer the quartiles to the exclusion of the minimum and maximum, but perhaps since the table with this information is supplementary these values could simply be added.

3. It would be helpful to also report the raw p-value in addition to the FDR corrected p-value. This type of correction is necessary when considering the results of analyses to be "significant" or "non-significant"; I agree with the current American Statistical Association guidelines that this is not a useful distinction and the raw p-value should be reported and interpreted (if the correction were accounting for the correlation structure of the data, I would perhaps feel differently). I am unsure of the Journal's stance on p-values, but I would encourage at least that the raw p-values are included alongside the FDR values in Table S1.

4. I am unfamiliar with sparse k-means clustering, which I believe will be true of many readers. The authors remark in the discussion that this approach complemented the univariate approach by identifying additional features which were found to contribute to the discrimination between the groups; these features did not differ between groups in the univariate approach. Looking at Table S1, it is interesting to consider the features which were selected by the clustering algorithm which differed from the results of the univariate analysis. If one computes the difference in spread of the data, the features where the spread was much greater in the FXS data tended to be selected by the clustering algorithm (but not univariate analysis). All of this leads me to my point: it would be helpful for the
authors to describe in further detail what drives the apparently discrepant results between the univariate and clustering analyses, especially since the authors remark that their use of the latter approach paves the way for others. At present, there is no exposition on why the clustering analysis identified these features—do extreme values matter (re: my point about spread above)? Does relative concentration matter? Without a more detailed consideration of the methodological "why" of these results, it's difficult to understand the meaning of the results—what does it mean that a handful of features discriminates the clusters but seemed to be drawn from the same distribution when the diagnoses were compared?

5. I believe that it is uncontroversial to remark that this study is underpowered. The authors describe the sample size calculation as relying on a theoretically large effect size, but powered for only one outcome (rather than ~50). The authors should take more care with this fact in the discussion; underpowered studies tend to overestimate effect sizes. Especially because little attention in the results is given to variability in the data (no confidence intervals, explained at least partially by the analytic methods), it is difficult for the reader to interpret the meaning of the results.

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

Yes

**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.

Yes

**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 recommend additional statistical review

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

Acceptable

**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