**Riley McDonnel, Will Snyder**

**Project 2**

**Nature Research Reporting Summary**

All deficient items in the nr-reporting summary are listed in its associated .pdf document. Additionally, for items where explanation for what should have been included did not have a text prompt, we list what should have been included below:

* A statement should have been included on whether measures were from distinct or the same sample, especially as some sample sizes were especially low (e.g. n = 5)/
* While the analysis did not leave out explaining covariates tested, it would have been beneficial to include female rats and use sex as a covariate, and age of rats was left out as a covariate.
* Assumptions for statistical tests were left out and should have been included. This is especially pertinent to the proteasome activity comparisons, which used t-tests with data of only n = 5 in each group. It is unknown whether the n = 5 could have been from a normal distribution.
* Degrees of freedom and effect sizes should have been reported for each test, but were not included for each test.

**Most important analysis: Proteasome activity changes in response to epoxomicin and PSI treatment**

While others have been able to replicate the analyses (constructing new data, as this data was not provided, leading to results that do not match this study) while only considering technical equipment differences as contributing to results, we were able to identify some information left out by the authors that did not make it clear as to how the methods were carried out.

This analysis could not be replicated for several reasons. The first is that the article describes that brains could be used for both biochemical and histological analysis, leading to using one brain hemisphere for one analysis and using one brain hemisphere for another. The choice of the hemisphere for which analysis is important, because it is unknown whether there could be hemisphere lateralization of potential proteasome activity changes. It should have been noted which hemisphere was chosen, or, better yet, more consistent use of both hemispheres or hemisphere-specific analyses should have been conducted. Additionally, it was not described how brain-region specific analyses were conducted with homogenized brain tissue. It is unclear how brain-region specific biochemical analysis of proteasome activity could have been dissociated from homogenized brain tissue. Therefore there was likely information left out by the authors.

**Questionable Research Practices**

1. Not reporting studies or variables that failed to reach statistical significance (e.g. p 0.05) or some other desired statistical threshold.

In this study all calculated p values were recorded, even when they failed statistical significance. This can be seen in the discussion when a student’s t test gave a p value of 0.19, which is much larger than the desired p value. They followed up by explaining why this value was important for supporting their data that body mass was not impacted by the treatments. It is clear throughout the paper that all p values are clearly reported.

1. Not reporting covariates that failed to reach statistical significance (e.g. p 0.05) or some other desired statistical threshold.

Covariates were not listed in this study. It seems unlikely that it was unnecessary to leave out covariate information as pertinent variables were not accounted for in the models. First, it would have been beneficial to include female rats and use sex as a covariate. Second, the age of the rats was not described, and should have been reported and used in models as a covariate. It is possible that age could have affected how the Parkinson’s Disease model progressed in individual rats. Either the authors left it out because it was insignificant, or it was never measured in the first place.

1. Reporting an unexpected finding or a result from exploratory analysis as having been predicted from the start.

The introduction does not make any specific hypotheses other than that there will be various behavioral, pathological, and neurochemical alterations in response to proteasome inhibitors that will align with a Parkinson’s Disease model. Instead, these levels of investigation are all carried out and compared to present knowledge about Parkinson’s Disease to evaluate whether the disease model they present aligns with other models and human manifestation of Parkinson’s disease.

1. Reporting a set of statistical models as the complete tested set when other candidate models were also tested.

The only models described were the ones that were tested, and it is unlikely that other models were tested and left out of the study intentionally. The reason this is unlikely is that all analyses were group comparisons that were used to identify increases or decreases in variables of interest (e.g. proteasome activity, cell counts in brain regions, etc). This is typical for comparing a control group to a disease model, and more complex hypotheses for disease progression were not outlined, so it is likely that the reported statistical models was actually the complete tested set.

1. Rounding-off a p value or other quantity to meet a pre-specified threshold (e.g., reporting p = 0.054 as p = 0.05 or p = 0.013 as p = 0.01).

No, this did occur in the paper unless the value was below the threshold, so p < 0.01. In the figures and discussion the p values are reported without any rounding.

1. Deciding to exclude data points after first checking the impact on statistical significance (e.g. p 0.05) or some other desired statistical threshold.

In this paper there are no missing data points based on the statistical significance. They reported all statistical values, including values larger than the desired p value. This can be seen in the results, “(334.9 13.7gm and 361 20.6; n 6 and 7, p = 0.19, Student’s t-test).” They also explain any significance, or insignificance, in their data. In each figure caption they explain what tests are being conducted and what the significance level is.

1. Collecting more data for a study after first inspecting whether the results are statistically significant (e.g. p 0.05).

In this study the rats were killed at 2 and 6 weeks depending on their treatment group, so no additional data could be generated after the rats were killed. The data that was collected while the rats were alive was based around the very small treatment groups and all tests were outlined. In this study they were very clear about the p values and the levels of significance they were using for each test. This can be seen in the discussion when looking at different CNS regions after a week of treatments, “One week after initiating treatment with PSI, there was an increase in proteasomal activity compared with controls in most CNS regions examined, although only some changes were statistically significant ( p < 0.01; see Fig 4A)”. They are almost always very clear about the data and its significance.

1. Changing to another type of statistical analysis after the analysis initially chosen failed to reach statistical significance (e.g. p 0.05) or some other desired statistical threshold.

It doesn’t appear that the researchers used any additional tests after previous tests failed, just that multiple tests were used to draw conclusions about the data. This can be seen in the caption for Figure 2, were they used Freidman test with Dunnett’s post test. \*p 0.01 compared with baseline and then using two-way ANOVA with Dunnett’s post-test. \*p 0.01 compared with baseline.

1. Not disclosing known problems in the method and analysis, or problems with the data quality, that potentially impact conclusions.

In the paper the figures and data often don’t have an explanation for where the values came from and some numbers don’t match with the figures generated. For instance Figure 4 has larger values recorded in the caption then the values on the graph. This is confusing as a reader to interpret because I’m not quite sure if the data was excluded or transformed in some way. The bar graph has a max value of of 350 for proteasome activity percent relative to control, but the caption lists values such as: “striatum (817.5 ± 25.2), spinal cord (440.5 ± 35.5), lower brainstem containing the LC and DMN (643.9 ± 132), cerebral cortex (281.7 ± 42.7), and cerebellum (283.6 ± 34.6).” Understanding where these values come from is important for understanding the significance of the figure. Other than that there appear to be no undisclosed errors in any part of the paper, just unclear information for where the data comes from and why it is included.

1. Filling in missing data points without identifying those data as simulated.

There does not seem to be missing data points throughout this paper, just data that is not completely clear as mentioned before. The data that was collected during this study is not publicly accessible, so there is no way to confirm that incomplete data was used, but based on the results and discussion the data was complete.

**Preregistration for replication of most important analysis**

Please refer to the separate preregistration R-markdown file.

**Acknowledgements**

All work was divided evenly. The contributions to parts of the preregistration are shown in each section of the preregistration. However, each section was thoroughly discussed between Riley and Will prior to any division of work so that the work would be consistent and cooperative. All other sections were completed cooperatively.