Perelman SCHOOL OF MEDICINE
UNIVERSITY of PENNSYLVANIA

Institute for Biomedical Informatics

Joseph D. Romano, PhD
Postdoctoral Researcher

Department of Biostatistics, Epidemiology, and Informatics

September 28, 2021

RE: Response to Reviewers for accepted PSB 2022 paper

Dear PSB 2022 Organizing Committee,

Please find in this letter the responses to reviewer comments for our recently accepted paper titled *Improving QSAR Modeling for Predictive Toxicology using Publicly Aggregated Semantic Graph Data*. We would like to thank the conference organizers, track organizers, and especially the reviewers for their thoughtful and detailed responses to our paper.

Below, we have reproduced each of the reviewers' comments and concerns along with narrative responses to each of them. Although there are no substantial changes made to the text of the manuscript, we do feel that the tweaks suggested by reviewers have significantly improved the paper by closing some remaining gaps in the analysis and providing more complete background such that it will be more useful for a wider audience of readers. Additionally, please note that some of the edits were made in the Supplementary Materials rather than the main text, which was necessitated by the 12-page limit of the main manuscript.

Reesponses to reviewer comments:

Reviewer 1

> *The labelling on Figure 3 was slightly confusing but perhaps I'm accustomed to different formatting in my discipline.*

We have tweaked the labeling of the violin plot (and the caption) to reduce the overall amount of text and make the important parts of the plot more intuitive and visually obvious.

> *. . . I wonder if additional, more observational, data could be intgegrated into the graph data and ultimately the GNN? The availability of such data may be sparse but I was wondering if additional data from epidemiological or clinical studies would be useful?*

This is indeed an insightful and important comment, and an area we are currently investigating as the subject of a follow-up study. Specifically, we are looking at harmonizing chemicals in the graph database with drug and environmental chemical exposures documented in electronic health record data, and seeing if the EHR data corroborates novel findings from the GNN approach. Although it is not immediately clear how, we are also exploring methods for including observational data directly in the GNN (e.g., each patient will be their own node in the graph, and we will use clinical findings, genetic data, and exposures to link patients to corresponding entities in the graph). We have added some brief allusions to this proposed future work in the Discussion section of the paper.

Reviewer 2

> *I recognize the page limitations, but the paper could be strengthened with a bit more detail about the algorithm used in the body of the work.*

Thank you for this comment. We did indeed hope to include more details on the model/algorithm but ended up having to keep it minimal to fit the 12-page limit. However, we did remove several extraneous sentences that allowed us to provide better detail as requested. Please refer to the Appendix to see the changes.

> *I do think that while the improvement is statistically significant, I'd like to know if the improvement in*

D206 Richards Medical Research Laboratory, 3700 Hamilton Walk • Philadelphia, PA 19104-6116 • Phone: (802) 558 5839
Email: joseph.romano@pennmedicine.upenn.edu

> *algorithm performance is the equivalent of "clinically relevent" to drug development.*

Although this is definitely the long-term goal of our research, we were careful to avoid making claims that the outcomes are clinically significant. However, we have added additional text to the Discussion explaining how our results can be considered *pre-clinically* significant in that they can directly guide the drug discovery industry in prioritizing compounds based on their predicted safety profiles.

> *Recognizing page limitations, the paper would benefit from a brief description of other problems that are being tackled with GCNs...*

Thank you for understanding our limitations in adding much additional text. However, this is an important comment, and we have been able to add a sentence that briefly introduces existing applications of GCNs.

> *There is ample literature on gene networks, but I'd like to see a theoretical discussion of GCNs in the context of gene networks as well, because toxicity is a very challenging problem and a solid foundation in theory would help move the field forward.*

Reviewer 3

> *Presentation of some parts (e.g., the first sentence of the abstract) can be improved more, but it is very minor. Algorithm can be more detailed, but I understand due to the limited pages.*

> *In the experiment, GNN showed higher variance than the others, and it was justified that it is because that neural networks tend to struggle as data become more sparse. I am not quite convinced for that. If data is imbalanced, cost function with class weights can be considered, and weighted F-1 score can be considered.*

> *I am wondering the performance with conventional neural networks because it compared only non-deep learning methods.*

> *I am wondering how to choose threshold to compute F-1 scores.*

Sincerely,

Joseph D. Romano, PhD
Postdoctoral Researcher

D206 Richards Medical Research Laboratory, 3700 Hamilton Walk • Philadelphia, PA 19104-6116 • Phone: (802) 558 5839
Email: joseph.romano@pennmedicine.upenn.edu