

Response to the comment on: Exploring the impact of design criteria for reference sets on performance evaluation of signal detection algorithms: The case of drug-drug interactions.

Elpida Kontsioti^{1,*}; Simon Maskell¹; Munir Pirmohamed²

¹ Department of Electrical Engineering and Electronics, University of Liverpool, Liverpool, UK

² The Wolfson Center for Personalized Medicine, Center for Drug Safety Science, Department of Pharmacology and Therapeutics, Institute of Systems, Molecular and Integrative Biology, University of Liverpool, Liverpool, UK

(* Correspondence: Elpida Kontsioti, Department of Electrical Engineering and Electronics, University of Liverpool, Liverpool, UK.

Email: elpida.kontsioti@gmail.com (E.Kontsioti@liverpool.ac.uk)

We thank Noguchi and Yoshimura for the considered response to our research [1], and for the interest in the CRESCENDDI reference set [2]. Our work has explored the impact of design criteria for reference sets on the relative performance evaluation of signal detection algorithms for the purpose of drug-drug interaction (DDI) surveillance.

We share the view that the analysis presented in [1] is inherently constrained by the incomplete knowledge of all safety issues associated with both single and, in our case, combinations of drugs. This inevitably leads to an inability to be certain that the negative controls in a reference set are not being polluted by any positive controls that have yet to be identified. This issue seems particularly apt for DDIs – indeed, the challenge of classifying a potential DDI as a false positive has been previously discussed [3]. As noted by Noguchi and Yoshimura, this uncertainty can indeed impact the sensitivity values calculated using our reference set. Nonetheless, we are optimistic that the conclusions drawn in [1], which are based on analysing the extent to which altering the design of reference sets impacts the sensitivity (and specificity), are likely to be more robust to changes in the negative controls than conclusions that could be drawn on the basis of absolute values of sensitivity.

Based on the above, it would be interesting to revisit the analysis performed in [1] in the future, aiming to identify potential discrepancies, given that our knowledge of known DDIs is likely to increase, at least for drugs that have been on the market for a more extended period. Additionally, for single-drug adverse effects, which offer a less complicated landscape, it would be interesting to investigate the relative impact of the design criteria considered in our analysis: While there are discussions in the literature about the misclassification of negative controls in the case of single-drug adverse effects [4], we believe that the single-drug negative controls may be more trustworthy than those related to DDIs.

The insights presented by Noguchi and Yoshimura have stimulated us in thinking about developing algorithms for reference sets (and signal detection algorithms) tailored to

situations where we cannot be confident in the identification of negative controls. Specifically, it would be interesting to incorporate uncertainty into the classification of negative controls by assigning confidence scores based on the evidence drawn from different information sources, including clinical information resources, the academic literature, and product labels. This approach would give a probabilistic dimension to the classification task of constructing a reference set, and avoid the imposition of rigid definitions of absolute positive and negative controls. For single drugs, previous research used evidence from the LAERTES platform [5] from the Observational Health Data Sciences and Informatics (OHDSI) collaborative network to assign prediction scores to drug-adverse event (AE) pairs for the classification of positive and negative controls [6]. This serves as a good framework for future work in the context of DDIs.

To conclude, we emphasise the importance of developing reproducible frameworks that facilitate the maintenance of reference sets which can be kept up to date with the latest safety information, while harnessing multiple sources of information and the most up-to-date best practice that exists in the design of such reference sets. Without automated tools, the need for extensive manual effort will limit the community's ability to maintain and adapt such reference sets. Conversely, with automated tools, such as the ones developed for building the CRESCENDDI reference set, the community has the capacity to assess the relative importance of different design criteria and to revisit that assessment in light of comments such as those sensibly raised by Noguchi and Yoshimura.

REFERENCES

- [1] Kontsioti, E., Maskell, S., & Pirmohamed, M. (2023). Exploring the impact of design criteria for reference sets on performance evaluation of signal detection algorithms: The case of drug-drug interactions. *Pharmacoepidemiology and drug safety*, 32(8), 832–844. <https://doi.org/10.1002/pds.5609>
- [2] Kontsioti, E., Maskell, S., Dutta, B., & Pirmohamed, M. (2022). A reference set of clinically relevant adverse drug-drug interactions. *Scientific data*, 9(1), 72. <https://doi.org/10.1038/s41597-022-01159-y>
- [3] Quinney S. K. (2019). Opportunities and Challenges of Using Big Data to Detect Drug-Drug Interaction Risk. *Clinical pharmacology and therapeutics*, 106(1), 72–74. <https://doi.org/10.1002/cpt.1481>
- [4] Hauben, M., Aronson, J. K., & Ferner, R. E. (2016). Evidence of Misclassification of Drug-Event Associations Classified as Gold Standard 'Negative Controls' by the Observational Medical Outcomes Partnership (OMOP). *Drug safety*, 39(5), 421–432. <https://doi.org/10.1007/s40264-016-0392-2>
- [5] Knowledge Base workgroup of the Observational Health Data Sciences and Informatics (OHDSI) collaborative (2017). Large-scale adverse effects related to treatment evidence standardization (LAERTES): an open scalable system for linking pharmacovigilance evidence sources with clinical data. *Journal of biomedical semantics*, 8(1), 11. <https://doi.org/10.1186/s13326-017-0115-3>
- [6] Voss, E. A., Boyce, R. D., Ryan, P. B., van der Lei, J., Rijnbeek, P. R., & Schuemie, M. J. (2017). Accuracy of an automated knowledge base for identifying drug adverse reactions. *Journal of biomedical informatics*, 66, 72–81. <https://doi.org/10.1016/j.jbi.2016.12.005>