



Letter to the Editor

Reply to commentary on “Predicting seizure freedom after epilepsy surgery, a challenge in clinical practice”


Keywords:

Epilepsy surgery
Outcome prediction
p-Value

Reply:

We thank Drs. Goldenholz, Sun, and Westover for raising a controversial and counter intuitive issue that we elaborate on here, the issue of statistical significance in prediction model evaluation. While much has been written lately about the overreliance on p-values in the inference setting [1], less attention has been paid to the prediction setting. In our opinion, the situation is more straightforward with respect to prediction: p-values are generally not helpful [2]. Suppose one would like to obtain predicted outcomes, and two methods are available for obtaining them, Methods A and B. The Methods are equally costly, invasive, etc., and all that matters for evaluation is predictive accuracy. In this case, one should generally prefer the Method with the higher predictive accuracy, regardless of the p-value. If the difference between Methods A and B is not statistically significant, one should not then be indifferent with respect to the choice (you should bet on the one that appears better), nor is it sensible to argue in favor of the one with the lower predictive accuracy (even though not statistically significant). This is our logic for preferring one Method over another. Although we remain in an era where this is difficult to achieve, one could easily argue here that p-values for these differences should not even be reported.

In the context of epilepsy surgery outcome prediction, a lot remains to be accomplished. Our unease with a status quo of blind reliance on clinician judgment and “experience” is only dwarfed by our inability to accept that improvement is neither necessary nor possible. One may flip the argument presented by Dr. Goldenholz and colleagues backwards: if a basic, simple, statistical model, published as a proof of concept [3] can perform as well as experienced epilepsy specialists,

then: 1)- further gains from clinical experience are doubtful, while 2)- there is a lot to be hoped for by collaborations and efforts to develop the model further by incorporating and testing the predictive value of additional promising outcome determinants (imaging, electrophysiology, genetics, etc.). Our published work is a small step – with “appropriately muted conclusions” – as stated by Dr. Goldenholz and colleagues. It is, however, a step in the right direction. We invite others to take similar steps, on whichever path they feel is more promising: our goal is to improve our ability to select and counsel our patients for surgery, rather than profess that we know the answer.

We apologize for a typo in Fig. 6b. Although the legend at the bottom has the correct AUC for the doctors' curve (0.466), the legend within the figure misstates this as 0.539. A corrected figure (legend) is attached.

Declaration of Competing Interest

No competing interests.

References

- [1] Wasserstein Ronald L, Lazar Nicole A. The ASA statement on p-values: context, process, and purpose. *Am Stat* 2016;70(2):129–33. <https://doi.org/10.1080/00031305.2016.1154108>.
- [2] Kattan M, Gonen M. The prediction philosophy in statistics. *Urol Oncol* 2008;26(3):316–9. <https://doi.org/10.1016/j.urolonc.2006.12.002>.
- [3] Jehi L, Yardi R, Chagin K, Tassi L, Russo GL, Worrell G, et al. Development and validation of nomograms to provide individualized predictions of seizure outcomes after epilepsy surgery: a retrospective analysis. *Lancet Neurol* 2015 Mar;14(3):283–90. [https://doi.org/10.1016/S1474-4422\(14\)70325-4](https://doi.org/10.1016/S1474-4422(14)70325-4) [Epub 2015 Jan 29. PubMed PMID: 25638640].

Lara Jehi

Epilepsy Center, Cleveland Clinic, United States of America
Corresponding author at: Cleveland Clinic, 9500 Euclid Ave, Cleveland, OH 44195, United States of America.
E-mail address: jehil@ccf.org

Michael Kattan

Quantitative Health Sciences department, Cleveland Clinic, United States of America

13 July 2019