Epilepsia

Is the crystal ball broken? Another external validation of the post-withdrawal seizure-relapse prediction model

We read with interest the work of Contento et al.[1] They performed the third external validation of the Lamberink Prediction Model (LPM), which assesses seizure relapse risk following withdrawal of antiseizure medications (ASMs).[2]

In a first external validation,[3] Lin et al. reported an area under the curve (AUC) of 0.71. They showed that the LPM outperformed predictions based upon the single largest randomized-controlled trial (RCT) to date,[4] although somewhat overpredicted observed probabilities. Chu et al.[5] provided a second Chinese cohort (AUC 0.61, again some overprediction). In contrast, Contento et al. concluded that model accuracy was inadequate because no single cutoff point provided high sensitivity and specificity (AUC ~0.5).

We agree that the LPM has limitations. Recruitment occurred mostly pre-2000 thus was lacking newer ASMs, genetics, and magnetic resonance imaging (MRI) studies. In contrast, in Contento et al. all patients underwent MRI, which could influence variables in the model such as electroencephalography (EEG) interpretation or focality explaining some divergence.

However, we have several concerns about their conclusions. First, the LPM was created from a large ($N = 1769$) diverse ($N_{countries} = 7$) data set, using "leave one out" internal-external cross-validation, which essentially performed 10 external validation steps in addition to the two from Lin et al. and Chu et al. We suggest caution before discounting the LPM in light of essentially only 1 of 13 validation steps suggesting poor performance. It would seem very surprising if the 12 variables contained in LPM (epileptiform EEG, number of seizures, duration seizure-free, and so on) predicted relapse no better than chance as the modified receiver-operating characteristic (ROC) curve of Contento et al. suggests. Moreover, validation is only as strong as the external data source. Selection mechanisms going from 4154 patients diagnosed with epilepsy down to just 205 (5%) who discontinued their ASM are not described, and another 36 of 205 were excluded due to missing data or incomplete follow-up. Including only 3% of those diagnosed with epilepsy raises concerns that the strong selection process determining discontinuation could explain divergence from the Chinese results.

Second, sensitivity and specificity are not the only metrics by which to judge a model. Observed vs predicted calibration may be a more intuitive way to assess model fit. In addition, the Discussion in the report of Contento et al. focuses on the inability of LPM to provide a single best cutoff. However, we believe that the predicted probability itself is the quantity of interest, rather than seeking an arbitrary dichotomous prediction. Even a perfectly accurate model could not inform what constitutes "high" vs "low," which varies from patient to patient.

Third, Contento et al. interpreted the decision curve analysis of Lin et al. as showing usefulness only within limited ranges. However, that range (30%–65%) actually contains the majority of patients. It would be interesting if a future study compared the accuracy of clinician predictions vs the LPM, given it is generally the rule rather than the exception that big data-driven individualized prediction models outperform clinician intuition alone (a recent example[6]).

Ultimately, showing predicted probabilities to patients influences decisions,[7] and we acknowledge that the ability to predict outcomes is imperfect,[8,9] encouraging future work. We appreciate the enthusiasm for critically appraising the best available science to move the field forward.

CONFLICT OF INTEREST

AUTHOR CONTRIBUTIONS

All coauthors have been substantially involved in the study and the preparation of the manuscript. There are no undisclosed groups or persons who have had a primary role in the study and/or in manuscript preparation. All co-authors have seen and approved the submitted version of the paper and accept responsibility for its content.

Samuel W. Terman[1,2]
Herm J. Lamberink[3,4]
Geertruida Slinger[4]
Willem M. Otte[4]
James F. Burke[1,2]
Kees P. J. Braun[4]

[1]*Department of Neurology, University of Michigan, Ann Arbor, Michigan, USA*
[2]*University of Michigan Institute for Healthcare Policy and Innovation, Ann Arbor, Michigan, USA*
[3]*Department of Neurology, Haaglanden Medical Center, Den Haag, The Netherlands*
[4]*Department of Child Neurology, University Medical Center, Utrecht University, Utrecht, The Netherlands*

Correspondence
Samuel W. Terman, Department of Neurology, University of Michigan, Taubman 1st Floor, Reception C, 1500 E Medical Center Dr, SPC 5316. Ann Arbor, MI 48109, USA.
Email: sterman@umich.edu

ORCID

Samuel W. Terman https://orcid.org/0000-0001-6179-9467
Willem M. Otte https://orcid.org/0000-0003-1511-6834

REFERENCES

1. Contento M, Bertaccini B, Biggi M, Magliani M, Failli Y, Rosati E, et al. Prediction of seizure recurrence risk following discontinuation of antiepileptic drugs. Epilepsia. 2021;62(9):2159–70.
2. Lamberink HJ, Otte WM, Geerts AT, Pavlovic M, Ramos-lizana J, Marson AG, et al. Individualised prediction model of seizure recurrence and long-term outcomes after withdrawal of antiepileptic drugs in seizure-free patients: a systematic review and individual participant data meta-analysis. Lancet Neurol. 2017;16:523–31.
3. Lin J, Ding S, Li X, Hua Y, Wang X, He R, et al. External validation and comparison of two prediction models for seizure recurrence after the withdrawal of antiepileptic drugs in adult patients. Epilepsia. 2020;61(1):115–24.
4. Medical Research Council. Randomised study of antiepileptic drug withdrawal in patients in remission. Lancet. 1991;337:1175–80.
5. Chu SS, Tan G, Wang XP, Liu L. Validation of the predictive model for seizure recurrence after withdrawal of antiepileptic drugs. Epilepsy Behav. 2021;114:106987.
6. van Doorn WPTM, Stassen PM, Borggreve HF, Schalkwijk MJ, Stoffers J, Bekers O, et al. A comparison of machine learning models versus clinical evaluation for mortality prediction in patients with sepsis. PLoS One. 2021;16(1):1–15.
7. Jacoby A, Baker G, Chadwick D, Johnson A. The impact of counselling with a practical statistical model on patients' decision-making about treatment for epilepsy: findings from a pilot study. Epilepsy Res. 1993;16:207–14.
8. Hoffmann TC, Del Mar C. Clinicians' expectations of the benefits and harms of treatments, screening, and tests: a systematic review. JAMA Int Med. 2020;4229(3):407–19.
9. Gracia CG, Chagin K, Kattan MW, Ji X, Kattan MG, Crotty L, et al. Predicting seizure freedom after epilepsy surgery, a challenge in clinical practice. Epilepsy Behav. 2019;95:124–30.