Prognostic models in first-episode psychosis - Authors' reply

Leighton, Samuel P.; Birchwood, Max; Mallikarjun, Pavan K.

DOI:
10.1016/S2589-7500(19)30238-9

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (Harvard):
Leighton, SP, Birchwood, M & Mallikarjun, PK 2020, 'Prognostic models in first-episode psychosis - Authors' reply', *The Lancet Digital Health*, vol. 2, no. 2, pp. e61-e61. https://doi.org/10.1016/S2589-7500(19)30238-9

Prognostic models in first-episode psychosis

Authors' reply

We thank Daniel Whiting and Seena Fazel for their insightful comments about our work and important suggestions for improvement in our analyses.

We completely agree that machine learning might not be appropriate for all prediction problems, though the distinction between what constitutes machine learning and what constitutes traditional statistical approaches is less clearcut, and some have suggested to consider them as being on a continuum.[1] We agree with Whiting and Fazel about models being clinical acceptable if prognostic factors are included as a starting point. However, given that a 2017 meta-regression did not find any significant predictors of remission in first-episode psychosis,[2] one could argue that data-driven approaches, including machine learning, might offer a much better opportunity than traditional statistical approaches.

Whiting and Fazel rightly point out the requirement of ten events per predictor parameter (EPP) for machine learning. We had 163 predictor parameters available for model development, which included most of the predictors that have been mentioned in the literature. We did not feel able to further reduce these variables on the basis of previous expert knowledge, because no consensus on reliable predictors of poor outcome in psychosis exists. We felt that it would be clinically less useful to combine individual items from rating scales into a single total score because this would necessitate carrying out the whole rating scale, causing the models to be non-parsimonious. We took a non-hypothesis driven approach to feature selection by using leave-one-site-out-cross-validation (LOSOCV) with elastic net for feature selection on our development data.[3] For example, for the remission outcome for external validation in the OPUS study, our final logistic regression model had 22 predictors and was built on 673 participants with 320 events, giving a EPP of 14·5. We apologise if this method was not clear in our manuscript.

Regarding the second point of Whiting and Fazel, we agree that measures of calibration and assessment of clinical usefulness are essential to evaluate a model's performance, along with measures of discrimination, before considering the application of any prediction model to clinical practice.[4] This is missing from our study and we thank Whiting and Fazel for highlighting this important limitation. It is one we will address in our future work.

We declare no competing interests.

*Samuel P Leighton, Max Birchwood, *Pavan K Mallikarjun*
p.mallikarjun@bham.ac.uk

Institute of Health and Wellbeing, University of Glasgow, Glasgow, UK (SPL); Mental Health and Wellbeing, Warwick Medical School, University of Warwick, Coventry, UK (MB); and Institute for Mental Health, University of Birmingham, Birmingham B15 2SA, UK (PKM)

1 Beam AL, Kohane IS. Big data and machine learning in health care. *JAMA* 2018; **319:** 1317–18.
2 Lally J, Ajnakina O, Stubbs B, et al. Remission and recovery from first-episode psychosis in adults: systematic review and meta-analysis of long-term outcome studies. *Br J Psychiatry* 2017; **211:** 350–58.
3 Zou H, Hastie T. Regularization and variable selection via the elastic net. *J R Stat Soc Series B Stat Methodol* 2005; **67:** 301–20.
4 Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. *Eur Heart J* 2014; **35:** 1925–31.