Letter to the Editor

Stroke *welcomes Letters to the Editor and will publish them, if suitable, as space permits. Letters must reference a* Stroke *published-ahead-of-print article or an article printed within the past 4 weeks. The maximum length is 750 words including no more than 5 references and 3 authors. Please submit letters typed double-spaced. Letters may be shortened or edited.*

Response by Guida and Scrutinio to Letter Regarding Article, "Development and Validation of a Predictive Model for Functional Outcome After Stroke Rehabilitation: The Maugeri Model"

We thank Dijkland et al for their interest in our article[1] on prediction of functional outcome after stroke rehabilitation. Dr Dijkland raises some issues about the outcome measures and the predictive model for the secondary outcome.

In general, models predictive of dichotomous clinical outcomes are most relevant in medical applications. Consistently, dichotomized outcome measures, including the Barthel index and the modified Rankin Scale, have been used to assess the effects of stroke treatment in clinical trials. For patients undergoing rehabilitation after a moderate-to-severe stroke, regaining minimal levels of disability or physical independence are clinically highly relevant outcomes. According to the article of Altman cited by Dijkland, we did not convert continuous explanatory variables to 2 (dichotomization) or multiple (categorization) groups, but kept them as continuous in the analysis.

The number of events relative to the number of regression coefficients estimated (excluding intercept) is known as events per variables (EPV).[2] The rule of thumb of 10 EPV for logistic regression models is based on two influential articles by Harrell et al[3] and Peduzzi et al[4] published in 1996. The simulation study of Peduzzi et al[4] was based on data from a trial of 673 patients, in which 252 deaths occurred during 10 years of follow-up. Seven binary variables were cogent predictors of mortality.[4] In that study, the accuracy and precision of regression coefficients declined as the number of EPV fell <10.[3] Vittinghoff and McCulloch,[5] however, showed that the rule of 10 EPV can be relaxed, as problems are uncommon with 5 to 9 EPV. The authors concluded, "discounting of results, in particular statistically significant associations, from any model with 5 to 9 EPV does not appear to be justified."[5] Moreover, in a recent article, van Smeden et al[2] concluded that "the evidence underlying the EPV = 10 rule as a minimal sample size criterion for binary logistic regression analysis is weak." Miscalibration of predictions in external data sets is common in prognostic research. According to Janssen et al,[6] in general, when the discrimination of a model is sufficient, simple recalibration methods can improve the calibration. Although overfitted in the derivation cohort, the model for the secondary outcome retained an excellent discriminative ability in the validation cohort (area under the curve = 0.85). Thus, we have recalibrated the model by updating both the intercept and the overall calibration slope. This new analysis yielded a χ^2 of 6.33 (*P* value for lack of fit = 0.097). It is relevant to note that recalibration does not affect the model's discrimination.

Miscalibration of the prediction model for the secondary outcome should not divert the attention from the key finding of our study that the model developed to predict the probability of a patient having mild disability at discharge from inpatient rehabilitation (primary outcome) exhibited an excellent discrimination and was well calibrated in both derivation and validation cohorts. Furthermore, we provided a supplementary model to predict Minimal Clinically Important Difference for Functional Independence Measure motor score. Even this model exhibited good discriminative ability and was well calibrated in both cohorts. Finally, further validation in other data sets was recommended in our article.[1]

Disclosures

None.

Pietro Guida, PhD
Domenico Scrutinio, MD
Department of Cardiology
Istituti Clinici Scientifici Maugeri-SPA SB. I.R.C.C.S.
Institute of Cassano Murge (Bari) Italy

References

1. Scrutinio D, Lanzillo B, Guida P, Mastropasqua F, Monitillo V, Pusineri M, et al. Development and validation of a predictive model for functional outcome after stroke rehabilitation: the Maugeri model. *Stroke*. 2017;48:3308–3315. doi: 10.1161/STROKEAHA.117.018058.
2. van Smeden M, de Groot JA, Moons KG, Collins GS, Altman DG, Eijkemans MJ, et al. No rationale for 1 variable per 10 events criterion for binary logistic regression analysis. *BMC Med Res Methodol*. 2016;16:163. doi: 10.1186/s12874-016-0267-3.
3. Harrell FE Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med*. 1996;15:361–387.
4. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. *J Clin Epidemiol*. 1996;49:1373–1379.
5. Vittinghoff E, McCulloch CE. Relaxing the rule of ten events per variable in logistic and Cox regression. *Am J Epidemiol*. 2007;165:710–718. doi: 10.1093/aje/kwk052.
6. Janssen KJ, Moons KG, Kalkman CJ, Grobbee DE, Vergouwe Y. Updating methods improved the performance of a clinical prediction model in new patients. *J Clin Epidemiol*. 2008;61:76–86. doi: 10.1016/j.jclinepi.2007.04.018.

(*Stroke*. 2018;49:00-00. DOI: 10.1161/STROKEAHA.117.020318.)

Stroke
Journal of the American Heart Association

Response by Guida and Scrutinio to Letter Regarding Article, ''Development and Validation of a Predictive Model for Functional Outcome After Stroke Rehabilitation: The Maugeri Model''

Pietro Guida and Domenico Scrutinio

Stroke. published online February 12, 2018;
Stroke is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231

Print ISSN: 0039-2499. Online ISSN: 1524-4628