# Letters

COMMENT & RESPONSE

## Prediction of Platelet Transfusion Outcomes in Preterm Newborns With Severe Thrombocytopenia

**To the Editor** A recent study[1] developed dynamic models that predicted individualized bleeding or mortality risk under 2 different transfusion strategies, aiming to individualize transfusion decisions, adjusting for both static and time-varying confounding to circumvent causal blind spots.[2] The authors note that the differences in estimated risks between the 2 strategies indicated estimates of individualized treatment effects.

Their models showed reasonably good discrimination and calibration performance in external validation for predicting outcome risks under both transfusion strategies, which may be interpreted as evidence that the individualized treatment effect estimates were accurate as well. However, reasonably well-calibrated estimates of outcome risks do not always correspond to reasonably well-calibrated estimates of individualized treatment effects (both with observational data or in a perfect randomized clinical trial). For example, a 2020 study[3] on predicting individualized treatment effects of percutaneous coronary intervention vs coronary artery bypass grafting on survival showcased a model (SYNTAX-II 2013) with little to no alignment between observed and predicted treatment effects, despite reasonable performance in predicting outcomes in both treatment groups (C indexes of 0.64-0.69 and calibration curves closely following the diagonal; eFigure 3 in the supplementary appendix).

This discrepancy has several possible explanations. For example, even small prediction errors in each of the treatment groups can compound in the estimation of absolute treatment effects and these errors are magnified because the scale of absolute treatment effects is much smaller than the scale of outcome risks.[4] Moreover, building separate models for each treatment group—hence, modeling all possible treatment-covariate interactions, as presented by van der Staaij et al[1]—risks overestimating the heterogeneity of treatment effects,[5] which would remain obscured unless treatment effect predictions are directly evaluated.

For these reasons, in line with the PATH (Predictive Approaches to Treatment Effect Heterogeneity) Statement,[4] we advocate for directly evaluating individualized treatment effect predictions beyond conventional outcome prediction metrics. Pending further research on the optimal method, one practical option is to examine the calibration of predicted treatment effects by stratifying the validation cohort based on predicted treatment effect quartiles and comparing the observed treatment effects with the predicted effects in each of these strata, as shown in the aforementioned example.[3]

For any model that aims to personalize treatment decisions, we advocate to consider its individualized treatment effects estimates and evaluate these directly, ideally using randomized clinical trial data.

**Jim M. Smit, MS**
**David M. Kent, MD, MS**
**Jesse H. Krijthe, PhD**

**Author Affiliations:** Data Science Group, Institute for Computing and Information Sciences, Radboud University, Nijmegen, the Netherlands (Smit); Predictive Analytics and Comparative Effectiveness (PACE) Center, Institute for Clinical Research and Health Policy Studies, Tufts Medical Center, Boston, Massachusetts (Kent); Pattern Recognition & Bioinformatics Group, Delft University of Technology, Delft, the Netherlands (Krijthe).

**Corresponding Author:** Jim Smit, MS, Data Science Group, Institute for Computing and Information Sciences, Radboud University, Mercator 1 Building, Toernooiveld 212, 6525 EC Nijmegen, the Netherlands (Jim.m.smit@outlook.com).

**Published Online:** January 21, 2026. doi:10.1001/jama.2025.22903

**Conflict of Interest Disclosures:** Dr Krijthe reported receiving funding for the project Safe Causal Inference with file No. OCENW.XL.23.123 of the research programme Open ENW-XL, which is financed by the Dutch Research Council. No other disclosures were reported.

**1**. van der Staaij H, Prosepe I, Caram-Deelder C, et al. Individualized prediction of platelet transfusion outcomes in preterm infants with severe thrombocytopenia. *JAMA*. 2025;334(14):1267-1277. doi:10.1001/jama.2025.14194

**2**. van Geloven N, Keogh RH, van Amsterdam W, et al. The risks of risk assessment: causal blind spots when using prediction models for treatment decisions. *Ann Intern Med*. 2025;178(9):1326-1333. doi:10.7326/ANNALS-24-00279

**3**. Takahashi K, Serruys PW, Fuster V, et al; SYNTAXES, FREEDOM, BEST, and PRECOMBAT trial investigators. Redevelopment and validation of the SYNTAX score II to individualise decision making between percutaneous and surgical revascularisation in patients with complex coronary artery disease: secondary analysis of the multicentre randomised controlled SYNTAXES trial with external cohort validation. *Lancet*. 2020;396(10260):1399-1412. doi:10.1016/S0140-6736(20)32114-0

**4**. Kent DM, van Klaveren D, Paulus JK, et al. The Predictive Approaches to Treatment Effect Heterogeneity (PATH) statement: explanation and elaboration. *Ann Intern Med*. 2020;172(1):W1-W25. doi:10.7326/M18-3668

**5**. van Klaveren D, Balan TA, Steyerberg EW, Kent DM. Models with interactions overestimated heterogeneity of treatment effects and were prone to treatment mistargeting. *J Clin Epidemiol*. 2019;114:72-83. doi:10.1016/j.jclinepi.2019.05.029