Prediction is very difficult, especially about the future

Anton, T.; Klautz, R.J.M.

more appropriate end-point to assess the peri-operative mortality compared to the in-hospital mortality.

REFERENCES

[1] Nezic D. Is it appropriate to use EuroSCORE II for prediction of 30-day mortality in cardiac surgery? Eur J Cardiothorac Surg 2021;60:434.

[2] Carino D, Denti P, Ascione G, Del Forno B, Lapenna E, Ruggeri S, *et al.* Is the EuroSCORE II reliable in surgical mitral valve repair? A single-centre validation study. Eur J Cardiothorac Surg 2020; doi:10.1093/ejcts/ezaa403.

[3] Siregar S, Groenwold R, de Mol B, Speekenbrink R, Versteegh M, Brandon Bravo Bruinsma G, *et al.* Evaluation of cardiac surgery mortality rates: 30-day mortality or longer follow-up? Eur J Cardiothorac Surg 2013;44:875–883.

[4] Mori M, Yousef S, Pichert M, Vinholo TF, Geirsson A. Elevated risk of death persists beyond 30 days after mitral valve surgery. J Thorac Cardiovasc Surg Elsevier Inc. 2020;159:e171–3.

*Corresponding author. Department of Cardiac Surgery, San Raffaele Scientific Institute, Via Olgettina 60, Milan 20132, Italy. Tel: +39-02-26437108; e-mail: carino.davide@hsr.it (D. Carino).

doi:10.1093/ejcts/ezab052
Advance Access publication 9 February 2021

The first external validation of GERAADA score for the prediction of 30-day mortality in patients with acute type A aortic dissection

Dusko Nezic *

Department of Cardiac Surgery I, "Dedinje" Cardiovascular Institute, Belgrade, Serbia

Received 8 January 2021; accepted 24 January 2021

Keywords: Aortic dissection surgery • Outcomes • Risk analysis

I read with great interest the article by Luehr *et al.* [1] who are the first who performed the external validation of recently launched GERAADA (the German Registry for Acute Type A Aortic Dissection) score [2] for the prediction of the 30-day mortality rate for patients undergoing surgery for acute aortic dissection type A (AADA).

The main finding of their study (Take-home message) is: The new GERAADA score is an easy, useful and accurate tool for predicting 30-day mortality in patients with AADA [1]. Their conclusion was based on retrospective analysis of 371 consecutive AADA patients operated in their institution between 2010 and 2020.

During the external validation of a risk prediction model, we have to check 2 key aspects, which characterize the performance of a prediction model: discrimination and calibration.

Discrimination differentiates low-risk from high-risk patients. It can be assessed by the area under the receiver operating characteristic curve. The area under the receiver operating characteristic curve (AUC) represents the percentage of randomly drawn pairs (a patient with an event paired with one without) in which the patient who had an event (in this case 30-day mortality following AADA surgery) had a higher risk score than a patient without. The discriminative power is thought to be excellent if the AUC is >0.80, very good if >0.75 and good (acceptable) if >0.70 [3]. Although calibration refers to the agreement between the observed number of perioperative (up to 30 days) deaths and predicted probability of occurrence of these events, the authors [1] confirmed good calibration ability of GERAADA score (actual mortality of 15.1% versus predicted probability by GERAADA score of 15.7%). However, discriminative power was poor, with an AUC of 0.67 (95% confidence interval 0.60–0.75). Therefore, when discrimination is not good, as it is true for validation of reported cohort [1], tested risk stratification model (GERAADA score) cannot be proclaimed as accurate tool for predicting 30-day mortality in patients with AADA. Furthermore, when discrimination of the model is good, but the calibration is not, the model can be made more accurate by recalibration. However, the opposite is not possible [4].

The authors [1] reported that preoperative resuscitation ($P < 0.001$), advanced age ($P = 0.042$) and other/unknown malperfusion ($P = 0.032$) were identified as independent risk factors for 30-day mortality. They also underlined that the third independent risk factor (other/unknown malperfusion) remained controversial. The authors have offered some logical explanations for such an outcome (i.e. well known negative impact of renal malperfusion on postoperative outcome in AADA surgery [5]). However, when you access on-line GERAADA Score calculator (https:/www.dgthg.de/de/GERAADA_Score), it is completely irrelevant whether you mark or not the field entitled—other/unknown maplerfusion (section Preoperative organ malperfusion), because the GERAADA score retains the same value. In other words, GERAADA score shows the same value no matter whether field—no malperfusion or other/unknown malperfusion is marked. Thus, is it logical that variable that does not have any influence on final score value can be an independent risk factor for prediction of 30-day mortality?

REFERENCES

[1] Luehr M, Merkle-Storms J, Gerfer S, Li Y, Krasivsky I, Vehrenberg J *et al.* Evaluation of the GERAADA score for prediction of 30-day mortality in patients with acute type A aortic dissection. Eur J Cardiothorac Surg 2020; doi:10.1093/ejcts/ezaa455.

[2] Czerny M, Siepe M, Beyersdorf F, Feisst M, Gabel M, Pilz M *et al.* Prediction of mortality rate in acute type A dissection: the German Registry for Acute Type A Aortic Dissection score. Eur J Cardiothorac Surg 2020;58:700–6.

[3] Moons K, Altman D, Reitsma J, Ioannidis J, Macaskill P, Steyerberg E *et al.* Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): explanation and elaboration. Ann Intern Med 2015;162:W1–73.

[4] Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N *et al.* Assessing the performance of prediction models: a framework for some traditional and novel measures. Epidemiology 2010;21:128–38.

[5] Nishigawa K, Fukui T, Uemura K, Takanashi S, Shimokawa T. Preoperative renal malperfusion is an independent predictor for acute kidney injury and operative death but not associated with late mortality after surgery for acute type A aortic dissection. Eur J Cardiothorac Surg 2020;58:302–8.

*Corresponding author. "Dedinje" Cardiovascular Institute, Milana Tepića 1, Belgrade 11000, Serbia. Tel: +381-11-3601631/3601724; fax: +381-11-266 6392; e-mail: nezic@eunet.rs, nezic@ikvbd.com (D. Nezic).

doi:10.1093/ejcts/ezab056
Advance Access publication 10 February 2021

Prediction is very difficult, especially about the future

Anton Tomšič* **and Robert J.M. Klautz**

Department of Cardiothoracic Surgery, Leiden University Medical Center, Leiden, Netherlands

Received 10 January 2021; accepted 26 January 2021

Prognostic risk models are an important clinical tool that can be used for research purposes, monitoring performance and to evaluate the risk of event

occurrence on individual patient level. Validation is needed to determine the generalizability of risk predictions tools in various populations and real-life scenarios. Luehr *et al.* [1, 2] performed an external validation of the German Registry of Acute Aortic Dissection Type A (GERAADA) score that was designed to predict 30-day mortality in patients suffering from acute type A aortic dissection.

The two primary measures used to assess the performance of a risk prediction tool are calibration and discrimination. Calibration (the ability of the prognostic risk model to predict the absolute risk level) was excellent as the actual 30-day mortality observed in the study (15.1%) was nearly identical to the mortality predicted by the GERAADA score (15.7%). Based on this result, the model could be used for monitoring centre performance. However, the observed model discrimination power (the ability of a prognostic risk model to accurately identify patients at high risk of experiencing the event of interest) was rather low with receiver operating characteristics analysis demonstrating an area under the curve value of 0.673 (95% confidence interval 0.595–0.751). A higher area under the curve value indicates increasing model discrimination power and a value above 0.7 is generally considered acceptable. As the authors propose the model to be used as a bed side tool to explain the situation to the patient and the patient's relatives, this poor discrimination power should raise concerns. Moreover, additional analyses in various subgroups from the GERAADA score demonstrated poorest model calibration for 'Previous cardiac surgery', 'Preoperative ventilation' and 'Inotropes at referral' subgroups. As the latter two variables are not completely objective but are likely influenced by clinical decision-making, this might not come as a surprise.

Luehr *et al.* are to be congratulated for their efforts to validate the recently designed GERAADA score risk model. We believe that their results have revealed some important limitations of the model, providing important data to further improve its performance in the future.

Conflict of interest: none declared.

REFERENCES

[1] Luehr M, Merkle-Storms J, Gerfer S, Li Y, Krasivskyi I, Vehrenberg J *et al.* Evaluation of the GERAADA score for prediction of 30-day mortality in patients with acute type A aortic dissection. Eur J Cardiothorac Surg 2020; doi:10.1093/ejcts/ezaa455.

[2] Czerny M, Siepe M, Beyersdorf F, Feisst M, Gabel M, Pilz M *et al.* Prediction of mortality rate in acute type A dissection: the German Registry for Acute Type A Aortic Dissection score. Eur J Cardiothorac Surg 2020;58:700–6.

*Corresponding author. Department of Cardiothoracic Surgery, Leiden University Medical Center, K6-S, PO Box 9600, 2300 RC Leiden, Netherlands. Tel: +31-71-5263445; fax: +31-71-5266899; e-mail: a.tomsic@lumc.nl (A. Tomšič).

doi:10.1093/ejcts/ezab066
Advance Access publication 10 February 2021

Reply to Nezic and Tomsic and Klautz

Maximilian Luehr [a,*], **Yupeng Li** [b] **and Thorsten Wahlers** [a]

[a]Department of Cardiothoracic Surgery, Heart Center, University of Cologne, Cologne, Germany
[b]Department of Political Science & Economics, Rowan University, Glassboro, NJ, USA

Received 21 January 2021; accepted 24 January 2021

Keywords: German Registry of Acute Aortic Dissection Type A • Acute type A aortic dissection • Risk score • Aortic surgery

We thank Drs Nezic, Tomsic and Klautz [1, 2] for their interest in our work and appreciate their comments from the two letters. To recapitulate, the aim of our retrospective single-centre study was to compare our institutional results to the new German Registry of Acute Aortic Dissection Type A (GERAADA) score to evaluate its accuracy in prediction of 30-day mortality in a consecutive series of patients suffering from acute aortic dissection type A (AADA) [3].

Both letters pointed out the result summary from the calibration and discrimination tools and demonstrated concerns mainly on the area under the curve value and the interpretations of the new GERAADA score in the clinical setting. Even though calibration was excellent, as acknowledged by Nezic, Tomsic and Klautz [1, 2], the area under the curve value result did not surpass the 0.7 threshold for the overall patient sample (0.673; 95% confidence interval 0.595–0.751) [3]. However, a considerable number of subgroups demonstrated acceptable discrimination power with the area under the curve values above 0.7 or even higher. Indeed, our results may have revealed some limitations of the risk model; however, this was only the first attempt to evaluate the new GERAADA score in a relatively small retrospective clinical cohort ($n = 314$).

In addition, Tomsic and Klautz [2] mistakenly criticize that the authors proposed to use the GERAADA score, despite the poor discrimination power, to explain the hazardous situation to AADA patients and the patients' relatives preoperatively. We believe that the GERAADA score may be used as a bedside tool for improved decision-making by the operating surgeon with regard to the expected operation time and the predicted mortality (e.g. hemiarch vs total arch replacement), but clearly it needs further evaluation before being actually used to explain the operative risk to affected patients and their relatives. These limitations, such as the need for further analyses in larger patient cohorts or subgroups in a prospective clinical trial to rule out potential over- and underestimation by the risk score model, were already addressed in the manuscript [3].

Finally, Dr Nezic questions the logic in reporting the variable 'other/unknown malperfusion' as an independent risk factor for 30-day mortality of our clinical study cohort. His argument that 'it is completely irrelevant whether you mark or not mark the field' for the GERAADA score result when using the score calculator is true—and the reason for it has already been explained by Czerny and Feisst previously [4]—but the identification of independent risk factors for 30-day mortality in the present study cohort was performed additionally and independent from the actual GERAADA score evaluation. However, this specific and small subgroup of AADA patients—without the clinical information on preoperative malperfusion—often present emergently from an external site without an electrocardiogram-guided computed tomography angiography or the time for a thorough physical examination preoperatively. As clinicians and surgeons, we know that AADA patients with unrecognized or belatedly diagnosed malperfusion are unquestionably at a higher risk for 30-day mortality [5]. So yes, we strongly believe in the logic of reporting all identified risk factors even if they may not directly influence the calculated result of a newly introduced scoring model.

The GERAADA score has been introduced only recently, and yet, it is still the only score specifically designed to predict 30-day mortality for AADA patients in a preoperative setting. We strongly believe that—despite the potential limitations rightly addressed by Nezic, Tomsic and Klautz—it will prove sufficient after further prospective evaluation and may serve as a useful tool in the future.

REFERENCES

[1] Nezic D. The first external validation of GERAADA score for the prediction of 30-day mortality in patients with acute type A aortic dissection. Eur J Cardiothorac Surg 2021;60:435.

[2] Tomsic A, Klautz RJM. Prediction is very difficult, especially about the future. Eur J Cardiothorac Surg 2021;60:435–6.

[3] Luehr M, Merkle-Storms J, Gerfer S, Li Y, Krasivskyi I, Vehrenberg J *et al.* Evaluation of the GERAADA score for prediction of 30-day mortality in patients with acute type A aortic dissection. Eur J Cardiothorac Surg 2020; doi:10.1093/ejcts/ezaa455.

[4] Czerny M, Feisst M. Reply to Sartipy. Eur J Cardiothorac Surg 2021;59:282–3.

[5] Bachet J. Acute type A aortic dissection: can we dramatically reduce the surgical mortality? Ann Thorac Surg 2002;73:701–3.

*Corresponding author. Department of Cardiothoracic Surgery, Heart Center of the University of Cologne, Kerpener Strasse 62, 50937 Cologne, Germany. Tel: +49-221-47832487; fax: +49-221-47832648; e-mail: maximilian.luehr@uk-koeln.de (M. Luehr).

doi:10.1093/ejcts/ezab076
Advance Access publication 15 February 2021