



Important methodological flaws in the recently published clinical prediction model the REMEMBER score

Granholt, Anders; Perner, Anders; Jensen, Aksel Karl Georg; Moller, Morten Hylander

Published in:
Critical Care

DOI:
[10.1186/s13054-019-2363-3](https://doi.org/10.1186/s13054-019-2363-3)

Publication date:
2019

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

Document license:
[CC BY](#)

Citation for published version (APA):
Granholt, A., Perner, A., Jensen, A. K. G., & Moller, M. H. (2019). Important methodological flaws in the recently published clinical prediction model the REMEMBER score. *Critical Care*, 23, [71].
<https://doi.org/10.1186/s13054-019-2363-3>

LETTER

Open Access



Important methodological flaws in the recently published clinical prediction model the REMEMBER score

Anders Granholm^{1*} , Anders Perner^{1,2}, Aksel Karl Georg Jensen^{2,3} and Morten Hylander Møller^{1,2}

See related research by Wang et al., <https://ccforum.biomedcentral.com/articles/10.1186/s13054-019-2307-y>

We have with interest read the recently published paper by Wang et al. in *Critical Care*, which proposes a new clinical prediction model—the REMEMBER score—to predict in-hospital mortality in patients undergoing venoarterial extracorporeal membrane oxygenation (VA-ECMO) after coronary artery bypass grafting [1]. The topic is clinically relevant; however, the study suffers from important methodological shortcomings, which hamper the validity of the findings and conclusions presented.

First, the study is critically underpowered. An effective sample size (minimum number of events or non-events) of ≥ 10 per candidate variable is recommended [2], and often more are required [3]. This number is 4–5 for the REMEMBER score (74 non-events/17 candidate variables). Consequently, the risk of random errors, overfitting, and inflated performance estimates is high. The performance will almost certainly deteriorate when used in other populations, and the single-centre design increases this risk.

Second, calibration was only assessed using the Hosmer-Lemeshow \hat{C} test, which is highly sensitive to sample size and unable to indicate lack of fit when power is low [2]. Lack of significance for this test does not equal adequate fit, and calibration plots or regressions of the predicted versus observed outcomes are recommended [2].

Third, comparing a newly developed prediction model with existing models using the development dataset is not recommended [2], as this approach is biased towards favouring the new model, especially when the risk of overfitting is high. A different cohort independent of model development must be used [2].

Additional important limitations include the long recruitment period where standards of care may have changed, the use of a non-fixed-time mortality outcome affected by discharge practices, the lack of external validation, and the reporting, which lacks information on sample size considerations and missing data handling, and inadequately acknowledges important limitations [4].

We are worried about the consequences if the REMEMBER score is used to select patients for VA-ECMO as suggested [1]. VA-ECMO is a costly and highly invasive treatment with severe potential adverse effects, and we agree that research within this area is highly needed [5]. Prediction models are relevant in this context; however, it is paramount that such models are developed, validated, and reported appropriately [2, 4]. If not, their use may put patients at risk and lead to inappropriate use of resources. Developing and validating a trustworthy clinical prediction model in this very selected patient population likely requires international, multicentre collaboration [5].

Authors' response

Liangshan Wang, Hong Wang and Xiaotong Hou

We are grateful to Granholm et al. for their interesting and valuable comments on our paper.

The REMEMBER score was developed according to the SAVE score [6], ENCOURAGE score [7], and two

published recommendations [8, 9]. Actually, we have mentioned that the single-centre design and absence of external validation may limit the generalizability of the REMEMBER score. Theoretically, the number of patients included in this study might be a little limited regarding the number of variables included in the model, which is similar in the ENCOURAGE score. Thus, we performed a bootstrap analysis, and it showed similar

* Correspondence: andersgran@gmail.com

¹Department of Intensive Care 4131, Copenhagen University Hospital – Rigshospitalet, Inge Lehmanns Vej 5, 2100 Copenhagen, Denmark
Full list of author information is available at the end of the article



results, confirming the stability of the original model. In addition, a post hoc random forest analysis [10] was performed, in which the six pre-ECMO parameters of the REMEMBER score were the top 10 risk factors. As for calibration, we also investigated the relationship between the predicted and observed outcomes grouped by REMEMBER score classes using a method that was similar to the calibration plots and found that there was a very good overlap between observed and expected mortality in all four groups. Calibration regression of the predicted versus observed outcomes is always recommended for external validation. Over the last 14 years, our surgical approach to CABG and standards of care have not changed much over the last 14 years, and year of ECMO was not associated with mortality by univariable analysis ($p = 0.698$). Given the absence of external validation, we have moderated the conclusion and did not use confirmatory terms in the paper. Prospective studies are needed to externally validate the scoring system before it can be widely applied.

Abbreviations

VA-ECMO: Venoarterial extracorporeal membrane oxygenation

Acknowledgements

Not applicable.

Funding

None.

Availability of data and materials

Not applicable.

Authors' contributions

AG was responsible for the conception of the letter and drafted the manuscript. All authors revised, read, and approved submission of the final manuscript.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Department of Intensive Care 4131, Copenhagen University Hospital – Rigshospitalet, Inge Lehmanns Vej 5, 2100 Copenhagen, Denmark. ²Centre for Research in Intensive Care, Blegdamsvej 9, 2100 Copenhagen, Denmark. ³Section of Biostatistics, University of Copenhagen, Øster Farimagsgade 5, 1014 Copenhagen, Denmark.

Received: 18 January 2019 Accepted: 17 February 2019

Published online: 07 March 2019

References

1. Wang L, Yang F, Wang X, Xie H, Fan E, Ogino M, et al. Predicting mortality in patients undergoing VA-ECMO after coronary artery bypass grafting: the REMEMBER score. *Crit Care*. 2019. <https://doi.org/10.1186/s13054-019-2307-y>.
2. Labarère J, Bertrand R, Fine MJ. How to derive and validate clinical prediction models for use in intensive care medicine. *Intensive Care Med*. 2014. <https://doi.org/10.1007/s00134-014-3227-6>.
3. Courvoisier DS, Combes C, Agoritsas T, Gayet-Ageron A, Perneger TV. Performance of logistic regression modeling: beyond the number of events per variable, the role of data structure. *J Clin Epidemiol*. 2011. <https://doi.org/10.1016/j.jclinepi.2010.11.012>.
4. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMJ*. 2015. <https://doi.org/10.1136/bmj.g7594>.
5. Schellongowski P, Combes A, Møller MH. Focus on extracorporeal life support. *Intensive Care Med*. 2018. <https://doi.org/10.1007/s00134-018-5465-5>.
6. Schmidt M, Burrell A, Roberts L, Bailey M, Sheldrake J, Rycus PT, et al. Predicting survival after ECMO for refractory cardiogenic shock: the survival after veno-arterial-ECMO (SAVE)-score. *Eur Heart J*. 2015;36:2246–56.
7. Muller G, Flecher E, Lebreton G, Luyt CE, Trouillet JL, Bréchet N, et al. The ENCOURAGE mortality risk score and analysis of long-term outcomes after VA-ECMO for acute myocardial infarction with cardiogenic shock. *Intensive Care Med*. 2016;42:370–8.
8. Collins GS, Reitsma JB, Altman DG, Moons KG, TRIPOD group. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *Circulation*. 2015;131:211–9.
9. Labarere J, Renaud B, Fine MJ. How to derive and validate clinical prediction models for use in intensive care medicine. *Intensive Care Med*. 2014;40:513–27.
10. Raja S, Rice TW, Ehrlinger J, Goldblum JR, Rybicki LA, Murthy SC, et al. Importance of residual primary cancer after induction therapy for esophageal adenocarcinoma. *J Thorac Cardiovasc Surg*. 2016;152(3):756–61 e5.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

