

LETTER

Open Access



# Trustworthy or flawed clinical prediction rule?

Anders Granholm<sup>1</sup>, Anders Perner<sup>1,2</sup>, Aksel Karl Georg Jensen<sup>2,3</sup> and Morten Hylander Møller<sup>1,2\*</sup>

See related research by Hilder et al., <https://ccforum.biomedcentral.com/articles/10.1186/s13054-017-1888-6>.

We read with interest the recently published paper by Hilder et al. [1], where the authors present the PRESET-Score, a new clinical prediction rule for patients with acute respiratory distress syndrome treated with extracorporeal membrane oxygenation (ECMO). While the topic is clinically relevant and interesting, we are worried that spurious findings, biased results, and overstated findings are presented.

First, both the new and the four existing scores assessed are at high risk of being underpowered. Multivariable risk prediction models should be based on an effective sample size (lowest number of events/non-events) of at least 10, often more, per predictor variable assessed [2, 3]. Using 11 variables and 41 non-events (3.7 per predictor) results in overfitting of the development sample and inflated performance estimates [2]. This will be evident upon use of the score in other populations.

Second, comparing the performance of the new score with four existing scores using the development dataset is against recommendations [2], as this is biased to favor the new score due to overfitting. For comparison with other scores, an independent cohort not used to develop any of the scores must be used [2].

Third, internal validation is performed to quantify overfitting, and should be done by bootstrap resampling of the development dataset [2]. The authors state that they used logistic regression analysis to “reassess” the score, which essentially is a *recalibration* resulting in a new model generating new predictions. This is neither internal nor external validation, which requires assessment of predictions made by the score *without* modifications in a new sample [2].

Fourth, it is recommended to assess calibration by graphical methods or regressions of the predicted versus observed outcomes [2, 4], not by the Hosmer-Lemeshow

$\hat{C}$ -test, as  $P > 0.05$  is more likely to indicate lack of power than proper model fit when used on small samples.

While we agree that clinical prediction rules may be valuable for clinicians considering ECMO, it is a prerequisite that such scores are developed and validated using appropriate methodology [2] and sufficient sample sizes, and that all relevant features are transparently reported with adequate discussion of the limitations [5]. Developing and sufficiently validating a clinical prediction rule for this highly selected patient group likely requires a large, multi-centre collaboration to ensure trustworthy predictions that will benefit patients and relatives, the healthcare system, researchers, and society.

#### Abbreviations

ECMO: Extracorporeal membrane oxygenation

#### Acknowledgements

Not applicable.

#### Funding

None.

#### Availability of data and materials

Not applicable.

#### Authors' contributions

AG and MHM were 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

<sup>1</sup>Department of Intensive Care 4131, Copenhagen University Hospital—Rigshospitalet, Blegdamsvej 9, 2100 Copenhagen, Denmark

\* Correspondence: [mortenhylander@gmail.com](mailto:mortenhylander@gmail.com)

<sup>1</sup>Department of Intensive Care 4131, Copenhagen University Hospital—Rigshospitalet, Blegdamsvej 9, 2100 Copenhagen, Denmark

<sup>2</sup>Centre for Research in Intensive Care, Blegdamsvej 9, 2100 Copenhagen, Denmark

Full list of author information is available at the end of the article

<sup>2</sup>Centre for Research in Intensive Care, Blegdamsvej 9, 2100 Copenhagen, Denmark. <sup>3</sup>Section of Biostatistics, University of Copenhagen, Øster Farimagsgade 5, 1014 Copenhagen, Denmark.

Received: 23 December 2017 Accepted: 18 January 2018

Published online: 08 February 2018

#### References

1. Hilder M, Herbstreit F, Adamzik M, Beiderlinden M, Bürschen M, Peters J, et al. Comparison of mortality prediction models in acute respiratory distress syndrome undergoing extracorporeal membrane oxygenation and development of a novel prediction score: the PREdiction of Survival on ECMO Therapy-Score (PRESET-Score). *Crit Care*. 2017. <https://doi.org/10.1186/s13054-017-1888-6>.
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, Combescure 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. Van Calster B, Nieboer D, Vergouwe Y, De Cock B, Pencina MJ, Steyerberg EW. A calibration hierarchy for risk models was defined: From utopia to empirical data. *J Clin Epidemiol*. 2016. <https://doi.org/10.1016/j.jclinepi.2015.12.005>.
5. 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*. 2014. <https://doi.org/10.1136/bmj.g7594>.