

LETTER

Open Access



# Letter to the editor: a response to Ming's study on machine learning techniques for personalized breast cancer risk prediction

Daniele Giardiello<sup>1,2\*</sup> , Antonis C. Antoniou<sup>3</sup>, Luigi Mariani<sup>4</sup>, Douglas F. Easton<sup>3,5</sup> and Ewout W. Steyerberg<sup>2,6</sup>

A recent paper [1] compared two well-known breast cancer risk prediction models (BCRAT and BOADICEA) with eight different machine learning (ML) methods. The authors found a striking improvement in cancer prediction with ML. While their comparative assessment against more classical approaches is timely, we are skeptical about the results presented.

A recent review on ML methods in a clinical epidemiological context shows that benefits of ML tend to arise in biased comparisons [2]. In the analyses of Ming et al., the ML methods were not specific for survival data and the validation process was unfair. While the ML used fits to binary outcomes prediction (having the disease or not), BOADICEA/BCRAT computes the probability of developing breast cancer over time. Regarding the second aspect, a fair comparison of the validity of the models would require data on unaffected women with prospective follow-up, with like for like risk predictions (over the same time period) for all methods. The comparisons in [1] were based on retrospective data of families of unaffected/affected individuals, and in the context of the BCRAT/BOADICEA, it is unclear what the observed and predicted events are. Furthermore, for the existing models, the study assessed external validity, while the ML methods were fitted on the same samples incorporating a tenfold cross-validation procedure, which is only equivalent to internal validation [3]. Internal validation is often overoptimistic in comparison to external validation studies [4]. Although the authors indicate the most important risk predictors in the ML approaches, the final models are not provided. A fair comparison would require the comparison of the final models from the ML with the existing models in external, prospective datasets. Moreover, discrimination is only one

measure of model performance: good calibration and clinical utility assessment are also crucial [5]. Last but not least, Ming et al. did not mention which version of BOADICEA was used for the comparison with ML methods. In conclusion, the practical relevance of ML methods needs to be further investigated in this specific context, based on more rigorous methodology.

#### Acknowledgements

Not applicable

#### Authors' contributions

DG, AAC, LM, DFE, and EWS conceived and drafted the letter to the editor. All authors read and approved the final manuscript.

#### Funding

Not applicable

#### Availability of data and materials

Not applicable

#### Ethics approval and consent to participate

Not applicable

#### Consent for publication

Not applicable

#### Competing interests

The authors declare that they have no competing interests.

#### Author details

<sup>1</sup>Division of Molecular Pathology, The Netherlands Cancer Institute - Antoni van Leeuwenhoek Hospital, Amsterdam, The Netherlands. <sup>2</sup>Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, The Netherlands. <sup>3</sup>Department of Public Health and Primary Care, Centre for Cancer Genetic Epidemiology, University of Cambridge, Cambridge, UK. <sup>4</sup>Unit of Clinical Epidemiology and Trial Organization, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy. <sup>5</sup>Department of Oncology, Centre for Cancer Genetic Epidemiology, University of Cambridge, Cambridge, UK. <sup>6</sup>Department of Public Health, Erasmus MC Cancer Institute, Rotterdam, The Netherlands.

Received: 28 December 2019 Accepted: 26 January 2020

Published online: 10 February 2020

#### References

1. Ming C, Viassolo V, Probst-Hensch N, Chappuis PO, Dinov ID, Katapodi MC. Machine learning techniques for personalized breast cancer risk prediction:

\* Correspondence: [d.giardiello@nki.nl](mailto:d.giardiello@nki.nl)

<sup>1</sup>Division of Molecular Pathology, The Netherlands Cancer Institute - Antoni van Leeuwenhoek Hospital, Amsterdam, The Netherlands

<sup>2</sup>Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, The Netherlands

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



comparison with the BCRA and BOADICEA models. *Breast Cancer Res.* 2019;21(1):75.

2. Christodoulou E, Ma J, Collins GS, Steyerberg EW, Verbakel JY, Van Calster B. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *J Clin Epidemiol.* 2019; 110:12–22.
3. Steyerberg EW, Harrell FE Jr. Prediction models need appropriate internal, internal-external, and external validation. *J Clin Epidemiol.* 2016;69:245–7.
4. Siontis GC, Tzoulaki I, Castaldi PJ, Ioannidis JP. External validation of new risk prediction models is infrequent and reveals worse prognostic discrimination. *J Clin Epidemiol.* 2015;68(1):25–34.
5. Vickers AJ, van Calster B, Steyerberg EW. A simple, step-by-step guide to interpreting decision curve analysis. *Diagn Progn Res.* 2019;3:18.

### **Publisher's Note**

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