**Breast Cancer Research**

Open Access

# Letter to the editor: Response to Giardiello D, Antoniou AC, Mariani L, Easton DF, Steyerberg EW

Chang Ming[1*], Valeria Viassolo[2], Nicole Probst-Hensch[3], Pierre O. Chappuis[2,4], Ivo D. Dinov[5,6,7,8] and Maria C. Katapodi[1,8]

We appreciate the opportunity to submit a response to the Letter to the Editor by Giardiello and colleagues [1] addressing our publication in Breast Cancer Research [2].

Giardiello and colleagues mentioned that our machine learning (ML) models were not specific for survival data. BCRAT/BOADICEA were developed and validated using survival data with binary outcomes and retrospective case control/cross-sectional data, respectively [3]. Their clinical application requires only cross-sectional data. Our ML models included the same risk factors and data structure in each comparison as BCRAT/BOADICEA. To avoid exaggerating the function of ML models, we generated the probability of whether a woman at a given age would develop breast cancer in her life, and not specific time frame risks (5-year or 10-year risk).

Giardiello and colleagues mentioned that our validation was unfair because we applied only internal validation processes. Cross-validation is not equivalent to internal validation; it is a statistical out-of-sample testing technique, which pools the results across many iterations, while each fold and each iteration do not blend training and testing data. A slight bias (aka surrogate problem) occurs because the cross-validation training sets are smaller than the original dataset. A 10-fold cross-validation relies on training sets that include 90% of the original dataset. In our study, this translated into two considerable sample sizes, $n_1 = 1029$ from the US population-based data and $n_2 = 2233$ from the Swiss clinic-based data. This lower-sample-size bias often translates into more conservative fit/prediction estimates [4].

Giardiello and colleagues mentioned that a fair comparison of the final models requires reporting parameter estimates and calibration. Reporting parameter estimates and their confidence intervals in the final model is not always possible [5]. We generated 80 parameter estimates for each risk factor based on different ML algorithms and different cross-validation summary approaches. The interpretation and usefulness of these estimates for each risk factor is limited without reference values from BCRAT/BOADICEA. Moreover, better/worse calibration does not lead to better/worse class-based or probability-based predictions [6]. Calibration comparisons was not our aim. ML may generate "aggressive" prediction calibration for minor classes due to "increased" sample size through rebalancing processes. Several recalibration methods can be applied and significantly improve some of the ML calibrations and predicted probabilities [6], making calibration comparisons of ML to BRCAT/BOADICEA unfair. Calibrated predicted probabilities should also fit clinically meaningful sensitivity and specificity for patient stratification, instead of one cutoff (cancer/no cancer) [7].

A prediction model cannot be developed, validated, and tested for utility at once. However, the development and validation of our ML models improved model predictive accuracy efficiently, i.e., using less time and fewer resources. Investing into promising new analytic approaches would improve research in the field of disease prediction and significantly further our knowledge about the potential application of ML in personalized medicine.

* Correspondence: chang.ming@unibas.ch
[1]Department of Clinical Research, Faculty of Medicine, University of Basel, Missionsstrasse 64, 2 OG – Room 007, 4055 Basel, Switzerland
Full list of author information is available at the end of the article

Ming *et al. Breast Cancer Research*     (2020) 22:35

Page 2 of 2

## Author details
[1]Department of Clinical Research, Faculty of Medicine, University of Basel, Missionstrasse 64, 2 OG – Room 007, 4055 Basel, Switzerland. [2]Oncogenetics and Cancer Prevention, Geneva University Hospitals, Geneva, Switzerland. [3]Swiss Tropical and Public Health Institute, University of Basel, Basel, Switzerland. [4]Genetic Medicine, Geneva University Hospitals, Geneva, Switzerland. [5]Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI, USA. [6]Michigan Institute for Data Science, University of Michigan, Ann Arbor, MI, USA. [7]Statistics Online Computational resource, University of Michigan, Ann Arbor, MI, USA. [8]University of Michigan School of Nursing, Ann Arbor, MI, USA.

## References
1. Giardiello D, Antoniou AC, Mariani L, Easton DF, Steyerberg EW. Letter to the editor: a response to Ming's study on machine learning techniques for personalized breast cancer risk prediction. Breast Cancer Res. 2020;22(1):17.
2. Ming C, Viassolo V, Probst-Hensch N, Chappuis PO, Dinov ID, Katapodi MC. Machine learning techniques for personalized breast cancer risk prediction: comparison with the BCRAT and BOADICEA models. Breast Cancer Res. 2019;21(1):75.
3. Wang X, Huang Y, Li L, Dai H, Song F, Chen K. Assessment of performance of the Gail model for predicting breast cancer risk: a systematic review and meta-analysis with trial sequential analysis. Breast Cancer Res. 2018;20(1):18.
4. Steyerberg EW, Harrell FE, Borsboom GJJM, Eijkemans MJC, Vergouwe Y, Habbema JDF. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. J Clin Epidemiol. 2001;54(8):774–81.
5. Luo W, Phung D, Tran T, Gupta S, Rana S, Karmakar C, et al. Guidelines for developing and reporting machine learning predictive models in biomedical research: a multidisciplinary view. J Med Internet Res. 2016; 18(12):e323.
6. Niculescu-Mizil A, Caruana R. Predicting good probabilities with supervised learning. Proceedings of the 22nd international conference on Machine learning. Bonn: Association for Computing Machinery; 2005. p. 625–32.
7. Brinton JT, Hendrick RE, Ringham BM, Kriege M, Glueck DH. Improving the diagnostic accuracy of a stratified screening strategy by identifying the optimal risk cutoff. Cancer Causes Control. 2019;30(10):1145–55.

## Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.