OXFORD

JNCI J Natl Cancer Inst (2021) 113(6): djaa181

doi: 10.1093/jnci/djaa181
First published online December 10, 2020
Editorial

EDITORIAL

Identifying Preferred Breast Cancer Risk Predictors: A Holistic Perspective

Ruth Etzioni, PhD,[1,*] Yu Shen, PhD[2] Ya-Chen Tina Shih, PhD[3]

[1]Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA, USA; [2]Department of Biostatistics, University of Texas MD Anderson Cancer Center, Houston, TX, USA and [3]Section of Cancer Economics and Policy, Department of Health Services Research, University of Texas MD Anderson Cancer Center, TX, USA

*Correspondence to: Ruth Etzioni PhD, Fred Hutchinson Cancer Research Center, Seattle, WA, USA (e-mail: retzioni@fredhutch.org).

Cancer risk prediction is the cornerstone of precision cancer control, and breast cancer risk prediction stands out as a prototype. Risk calculators based on clinical, biological, behavioral, and epidemiologic factors are available for many cancers, but breast cancer has the most established history and arguably the largest number of available tools for risk prediction.

Although there are various metrics for quantifying risk, since the 1989 publication by Gail et al. (1) that became the National Cancer Institute's Breast Cancer Risk Assessment Tool (2), the field has been largely focused on absolute breast cancer risk over an interval as a measure for targeting prevention and screening interventions.

In this issue of *JNCI*, MacInnis and colleagues (3) address a contentious issue, namely, whether short-term (5-year) risk should be preferred as the driver of intervention decisions over the traditionally used long-term (lifetime) risk. The investigators zero in on a definition of "preferred" that speaks to the accuracy of the predicted risk, comparing the diagnostic performance of 5-year and lifetime risk predictions from the IBIS (v8b) (4) and BOADICEA (v3) (5) tools in data from the Breast Cancer Prospective Family Study Cohort (PFSC) (6).

In this editorial, we discuss the process of evaluating risk prediction tools as exemplified by the study of MacInnis et al. (3) while raising a broader question of what constitutes a preferred tool in the context of risk prediction and communication. Beyond being accurate, we propose 2 additional qualities for tools to be preferred—they should be meaningful and actionable. We define each and argue that all 3 properties should be considered when identifying preferred prediction tools for informing targeted cancer control strategies.

The type of validation study represented by that of MacInnis et al. (3) is both common and necessary. Every risk prediction tool is developed and calibrated within a specific cohort before being offered to the field for potentially much broader use. Given inevitable differences across cohorts in population composition, screening practices, and length of follow-up, there is no guarantee that absolute risks from a model calibrated to 1 cohort will match a different cohort. In general, predicted long-term risks will exceed those observed in a cohort with short-term follow-up and vice versa. This may explain why MacInnis et al. (3) found that 5-year risk predictions seemed to perform better than lifetime risk predictions (at least for women younger than 40 years) in the PFSC, which had a follow-up interval close to 10 years. Still, their suggestion that 5-year risk is to be preferred in a clinical setting based on this finding bears further scrutiny.

First, once a risk prediction tool has been developed, whether it is for predicting 5-year, 10-year, or lifetime risk, it is essentially an algorithm calibrated to a specific training dataset. A new validation dataset is agnostic to the time horizon of the prediction tool; we might as well call our 5-year prediction algorithm A and our lifetime prediction algorithm B. Further, the standard validation metric, the concordance index or AUC, is a check on the ordering of the predicted risks rather than their magnitude. In principle, therefore, a lifetime risk prediction could perform better or worse in terms of Area Under the Curve (AUC) than a 5-year risk prediction on a validation set with relatively short-term follow-up. In the study of MacInnis et al. (3), the 5-year risk prediction happened to perform better, at least for women younger than 40 years. The point is that this is not necessarily a consequence of the time horizon of the prediction tool; rather, it is a reflection that algorithm A happens to align better with the validation data than algorithm B. Indeed, the fact that algorithm A is preferred over B when validating performance against a certain cohort could be a feature of that cohort and not a generalizable property of the algorithms themselves.

Beyond accuracy, however, there are clear reasons to prefer a tool with a specific time horizon. From a clinical perspective, there are 2 key questions: Why is the predicted risk needed, and what decision rests on its result? The answers could land the user squarely on the side of preferring a short-term over a long-term prediction. If a 45-year-old woman is deliberating about whether to start breast cancer screening or wait until age 50 years, then a 5-year risk is more meaningful than a lifetime

Received: October 28, 2020; **Accepted:** November 2, 2020

risk. This is similarly the case for a 65-year-old woman deliberating about whether to stop or continue screening until age 70 years. Naturally, even for an average-risk woman, these 2 risks will differ considerably. Comparison of the predicted risk against the risk in the general population at a similar age will be helpful in contextualizing the predicted risk and determining whether altered management is called for. An absolute risk that can seem relatively tiny could in fact reflect a dramatically elevated relative risk in the context of the age-matched population, and conversely. Ultimately, clinical meaningfulness may be as important as predictive accuracy in determining the preferred risk prediction time horizon.

Beyond accuracy and meaningfulness, there is the matter of how the risk prediction tool translates into a targeted cancer control strategy. Targeted strategies may focus on cancer prevention or early detection. In the case of early detection, a targeted strategy may screen only cases with predicted risk above a prespecified threshold, or it may intensify screening among such high-risk cases. The appropriate way of intensifying screening will depend on the mechanism by which risk differs across population strata. A higher risk of disease onset may be addressed by lowering the age to start screening to provide an equal opportunity for benefit; higher risk of disease progression may require more frequent screening to maintain benefit. And differential detectability may call for changing the screening modality. For example, if a 45-year-old woman has an elevated risk because of high breast density, then the appropriate intensification of screening may be different than if she has the same risk because of family history. A focus on predictive accuracy glosses over the fact that when it comes to intervening, being able to predict risk is not enough—we need to be able to explain it. An actionable risk prediction tool is one that is explainable in a manner that facilitates the identification of appropriately targeted interventions.

In conclusion, validation studies of predictive accuracy such as the one conducted by MacInnis et al. (3) occupy an established niche in precision prevention research. But predictive accuracy is only 1 desirable feature of risk prediction algorithms. These algorithms must also be meaningful and actionable if targeted strategies are to improve over one-size-fits-all approaches and do more good than harm.

Funding

This work was supported by the National Institutes of Health grant numbers R01CA242735, U54CA132381 (RE), CA016672 (YS), and CA207216 (YCTS).

Notes

Role of the funder: The funder had no role in the writing of this editorial or the decision to submit it for publication.

Disclosures: The authors have no conflicts of interest to disclose.

Author contributions: Ruth Etzioni, writing—original draft; Ya-Chen Tina Shih, writing—review & editing; Yu Shen, writing—review & editing.

Acknowledgments: We are grateful to Noel S Weiss, Mia Gaudet, and Marc Ryser for comments on an earlier draft of this article. RE's work is partially supported by the Rosalie and Harold Rea Brown Endowment.

Data Availability

Not applicable.

References

1. Gail MH, Brinton LA, Byar DP, et al. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. *J Natl Cancer Inst*. 1989;81(24):1879–1886.
2. National Cancer Institute. Breast Cancer Risk Assessment Tool. https://bcrisktool.cancer.gov/. Accessed July 1, 2020.
3. MacInnis RJ, Knight JA, Chung WK, et al. Comparing five-year and lifetime risks of breast cancer in the prospective family study cohort. *J Natl Cancer Inst*. 2020.
4. Wolfson Institute of Preventive Medicine - IBIS Breast Cancer Risk Assessment Tool v8. http://www.ems-trials.org/riskevaluator/. Accessed October 28, 2020.
5. University of Cambridge - Centre for Cancer Genetic Epidemiology. BOADICEA https://ccge.medschl.cam.ac.uk/boadicea/. Accessed October 28, 2020.
6. Terry MB, Phillips K-A, Daly MB, et al. Cohort profile: the breast cancer Prospective Family Study Cohort (ProF-SC). *Int J Epidemiol*. 2016;45(3):683–692.