JNCI J Natl Cancer Inst (2020) 112(5): djz180

doi: 10.1093/jnci/djz180
First published online September 26, 2019
Editorial

EDITORIAL

Choosing Breast Cancer Risk Models: Importance of Independent Validation

Mitchel H. Gail

See the Notes section for the full list of author's affiliations.
Correspondence to: Mitchell H. Gail, MD, PhD, National Cancer Institute, 9609 Medical Center Dr, Rockville, MD 20850-9780 (e-mail: gailm@mail.nih.gov).

Several widely used risk models estimate the chance that a woman will develop breast cancer over a defined time interval such as 5 years. Before using such a model, one should consider the nature of the target population and whether the model has been validated for that population. In this issue of the Journal, McCarthy and colleagues (1) present valuable data comparing the performance of five risk models in 35 921 predominantly white women who came to the Newton-Wellesley Hospital in Massachusetts for mammographic screens between 2007 and 2009. This study adds importantly to two other recent large validation studies. The UK Generation Study cohort (2) of 64 874 women was recruited between 2003 and 2012 from the general UK population. The Breast Cancer Prospective Family Study Cohort (3) (ProF-SC) included 15 732 women recruited from Australia, Canada, and the United States between 1992 and 2011 from high-risk clinics or as relatives of women in breast cancer registries. Of the women in ProF-SC, 82% had at least one affected first-degree relative, and 6.83% carried a mutation in *BRCA1* or *BRCA2* (4). These two general-population cohorts and one high-risk cohort enable assessment of risk model performance.

Four widely used models incorporate a rare autosomal dominant genetic component: Claus (5), BRCAPRO (6), International Breast Cancer Intervention Study (IBIS) (or Tyrer-Cuzick) (7), and Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm (BOADICEA) (8). These models require detailed family history. IBIS and BOADICEA allow for residual familial correlation not explained by a rare autosomal dominant gene. IBIS also includes many standard breast cancer risk factors, and a recent version added mammographic density. Two other models assessed by McCarthy and colleagues do not assume an autosomal dominant component. The National Cancer Institute's Breast Cancer Risk Assessment Tool (BCRAT) (9,10), sometimes called the Gail model, requires only age, age at menarche, age at first live birth, number of previous benign breast biopsies, presence of atypical hyperplasia on biopsy, number of affected mother or sisters, and race or ethnicity. The Breast Cancer Surveillance Consortium (BCSC) model, sometimes called the Tice model, uses age, race or ethnicity, mammographic density (BIRADS), history of breast cancer in a first-degree female relative, and biopsy history.

Two criteria are often used to assess model performance: calibration and discriminatory accuracy (11). In a cohort study, one can compare the expected number of breast cancers based on the model (E) to the observed number of breast cancers (O). A model is well calibrated if the ratio O/E is near 1. Discriminatory accuracy, namely the area under the receiver operating characteristic curve (AUC), is the probability that a randomly selected case will have a larger projected risk than a randomly selected noncase. An AUC near 0.5 indicates low discriminatory accuracy, and an AUC equal to 1 indicates perfect discriminatory accuracy.

Table 1 summarizes results on O/E and AUC from the three validation studies. In the Newton-Wellesley Mammography Cohort, BCRAT, BCSC, and BRCAPRO were well calibrated (confidence intervals include 1.0), but IBIS overestimated risk slightly, especially when mammographic density was used (O/E = 0.84, 95% confidence interval [CI] = 0.79 to 0.91). The Claus model underestimated risk substantially (O/E = 1.69, 95% CI = 1.48 to 1.87). The AUC values were modest (between 0.59 and 0.64). Despite its simplicity, BCRAT had an AUC as high BCSC's, which used BIRADS, and higher than other models. The UK Generation Study provides additional data on model performance in a general population. BCRAT underestimated risk slightly (O/E = 1.09, 95% CI = 1.02 to 1.16), whereas IBIS again overestimated risk slightly (O/E = 0.88, 95% CI = 0.83 to 0.94). Respective AUC values were 0.61 and 0.63. These data support the conclusions of McCarthy and colleagues that BCRAT, BCSC, BRCAPRO, and IBIS are reasonably well calibrated in general populations of largely white women, and have comparable moderate discriminatory accuracy. The Claus model underestimated risk by a factor of 1/1.69 = 0.59 in the Newton-Wellesley Mammography Cohort.

Received: August 28, 2019; **Accepted:** September 4, 2019

Table 1. Calibration and discriminatory accuracy of seven risk models in three independent validation cohorts

	Newton-Wellesley Mammography Cohort (1)		UK Generation Study (2)		ProF-SC (3) (all)		ProF-SC (3) (*BRCA* negative)	
Risk model*	O/E (95% CI)	AUC	O/E (95% CI)	AUC	O/E (95% CI)	AUC	O/E (95% CI)	AUC
BCRAT (Gail)	0.98 (0.91 to 1.06)	0.64	1.09 (1.02 to 1.16)‡	0.61‡	1.27 (1.18 to 1.37)	0.60	1.03 (0.94 to 1.12)	0.64
BCSC (Tice)†	0.97 (0.89 to 1.05)	0.64	–	–	–	–	–	–
BRCAPRO	0.94 (0.91 to 1.02)	0.61	–	–	1.69 (1.56 to 1.82)	0.68	1.89 (1.72 to 2.04)	0.62
IBIS (no MD)	0.90 (0.84 to 0.96)	0.61	–	–	0.97 (0.89 to 1.04)	0.71	1.00 (0.91 to 1.09)	0.66
IBIS (with MD)	0.84 (0.79 to 0.91)	0.62	0.88 (0.83 to 0.94) ‡	0.63‡	–	–	–	–
Claus§	1.69 (1.48 to 1.87)	0.59	–	–	–	–	–	–
BOADICEA	–	–	–	–	0.95 (0.88 to 1.03)	0.70	0.98 (0.89 to 1.08)	0.65

*Risk models are defined in the text with references. AUC = area under the receiver operating characteristic curve; BCRAT = Breast Cancer Risk Assessment Tool; BCSC = Breast Cancer Surveillance Consortium; BOADICEA = Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm; *BRCA* = mutation in *BRCA1* or *BRCA2* gene; CI = confidence interval; E = expected breast cancers from model; IBIS = International Breast Cancer Intervention Study; MD = mammographic density; O = observed breast cancers; ProF-SC = Breast Cancer Prospective Family Study Cohort.

†In 30 970 women with BIRADS, and other data needed for BCSC.

‡Data from Figures 1 and 2 in (2) are used to compute O/E and 95% CI. AUC is computed from the age-weighted average of AUC values in Table 1 of (2) for women younger than 50 years and women 50 years and older.

§In 11 873 women with affected family members. This includes invasive cancer and ductal carcinoma in situ.

ProF-SC permits evaluation in a high-risk population. IBIS (without mammographic density) and BOADICEA were well calibrated in the entire ProF-SC, but BRCAPRO substantially underestimated risk (O/E = 1.69, 95% CI = 1.56 to 1.82) as did BCRAT (O/E = 1.27, 95% CI = 1.18 to 1.37) (Table 1). However, BCRAT does not calculate risk for known carriers of *BRCA1* or *BRCA2* mutations (hereafter called *BRCA* mutations), and it recommends BOADICEA for *BRCA* carriers. BCRAT was well calibrated (O/E = 1.03) in the *BRCA*-negative subset of ProF-SC overall, but it underestimated risk in *BRCA*-negative women aged younger than 50 years (O/E = 1.28, 95% CI = 1.12 to 1.47 from Table 3 in Terry et al. [3]), whereas BOADICEA and IBIS were well calibrated in *BRCA*-negative women in both age cohorts (those younger than 50 years and those 50 years or older) (12). BRCAPRO underestimated risk in *BRCA*-negative women (O/E = 1.89, 95% CI = 1.72 to 2.04). Thus, BOADICEA and IBIS are preferred in this setting.

AUC values for BRCAPRO, BOADICEA, and IBIS ranged from 0.68 to 0.71 in ProF-SC, which is higher than in the general populations. However, in the *BRCA*-negative subset of ProF-SC, the AUC values ranged from 0.62 to 0.66 and differed little from the value 0.64 for BCRAT. This suggests that much of the discriminatory ability of BRCAPRO, BOADICEA, and IBIS in ProF-SC derives from their incorporation of data on *BRCA* mutations, which are extraordinarily prevalent (6.83%). Indeed, a model that included only *BRCA* status would have an AUC above 0.7 in ProF-SC, but only 0.53 in the general population (4,12), for which the prevalence is only 0.32% (8).

Breast cancer risk models with modest discriminatory accuracy are useful in counseling to provide general perspective on risk and to help weigh the risks and benefits of preventive interventions. For example, a 40-year-old woman may decide to begin mammographic screening because her risk is greater than that of a 50-year-old woman, for whom screening is recommended (13), or a 60-year-old woman may decide not to take tamoxifen because risks outweigh potential benefits (14). Risk models are also useful for designing prevention trials (15), and models with modifiable risk factors such as alcohol consumption (16–18) can be used for estimating the reductions in population absolute risk from reducing modifiable exposures. Other applications in public health require higher discriminatory accuracy, such as allocating preventive resources under cost constraints or identifying subsets of women who need not be screened for breast cancer (19). Recent versions of BOADICEA (20) and IBIS (21) incorporate mammographic density and polygenic risk scores to improve discriminatory accuracy. I estimated that the latest BOADICEA (20) had an AUC near 0.7, even in a population with few mutations (12), in line with previous predictions (22). But even higher discriminatory accuracy and safer preventive interventions will be needed to have a major impact on population absolute risk (19).

McCarthy and colleagues provide much-needed data on model performance in a population consisting largely of white women. More such data are needed for ethnic and racial subgroups and for women from various parts of the world, where breast cancer rates can be much lower. Some risk models take race and ethnicity into account, but care is needed to adapt models for use in other countries. Users of risk models also need to recognize special situations in which the models should not be used. For example, women who received radiation to the chest for childhood Hodgkin lymphoma have breast cancer risks comparable to *BRCA* carriers (23); the risk models evaluated by McCarthy and colleagues do not account for such women. As models that incorporate additional risk factors are developed, including iCARE (2) and the latest versions of BOADICEA and IBIS, independent validation studies such as that by McCarthy and colleagues will be needed to assess their performance.

Funding

This work was supported by the Intramural Research Program, National Institutes of Health, National Cancer Institute, Division of Cancer Epidemiology and Genetics.

Notes

Affiliation of author: Division of Cancer Epidemiology and Genetics, National Cancer Institute, Rockville, MD.

The funders had no role in the writing of this editorial or the decision to submit it for publication. The author has no disclosures.

References

1. McCarthy AM, Guan Z, Welch M, et al. Performance of breast cancer risk assessment models in a large mammography cohort. *J Natl Cancer Inst*. 2020;112(5):djz177.
2. Choudhury PP, Wilcox AN, Brook MN, et al. Comparative validation of breast cancer risk prediction models and projections for future risk stratification. *J Natl Cancer Inst*. 2020;112(3):djz113.
3. Terry MB, Liao Y, Whittemore AS, et al. 10-year performance of four models of breast cancer risk: a validation study. *Lancet Oncol*. 2019;20(4):504–517.
4. Gail MH. Performance of BCRAT in high-risk patients with breast cancer. *Lancet Oncol*. 2019;20(6):e285.
5. Claus EB, Risch N, Thompson WD. Autosomal-dominant inheritance of early-onset breast-cancer-implications for risk prediction. *Cancer*. 1994;73(3):643–651.
6. Parmigiani G, Berry DA, Aguilar O. Determining carrier probabilities for breast cancer-susceptibility genes BRCA1 and BRCA2. *Am J Hum Genet*. 1998;62(1):145–158.
7. Tyrer J, Duffy SW, Cuzick J. A breast cancer prediction model incorporating familial and personal risk factors. *Stat Med*. 2005;24(1):156.
8. Antoniou AC, Cunningham AP, Peto J, et al. The BOADICEA model of genetic susceptibility to breast and ovarian cancers: updates and extensions. *Br J Cancer*. 2008;98(8):1457–1466.
9. Costantino JP, Gail MH, Pee D, et al. Validation studies for models projecting the risk of invasive and total breast cancer incidence. *J Natl Cancer Inst*. 1999;91(18):1541–1548.
10. 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.
11. Pfeiffer RM, Gail MH. *Absolute Risk: Methods and Applications in Clinical Management and Public Health*. Baton Rouge, LA: Chapman and Hall/CRC Taylor and Francis Group; 2017.
12. Gail MH. Commentary on Terry et al., 10-year performance of four models of breast cancer risk: a validation study. *Lancet Oncol*. 2019;4(4):e190017.
13. Gail MH, Schairer C. Comments and response on the USPSTF recommendation on screening for breast cancer. *Ann Intern Med*. 2010;152(8):540.
14. Freedman AN, Yu B, Gail MH, et al. Benefit/risk assessment for breast cancer chemoprevention with raloxifene or tamoxifen for women age 50 years or older. *J Clin Oncol*. 2011;29(17):2327–2333.
15. Fisher B, Costantino JP, Wickerham DL, et al. Tamoxifen for prevention of breast cancer: report of the National Surgical Adjuvant Breast and Bowel Project P-1 study. *J Natl Cancer Inst*. 1998;90(18):1371–1388.
16. Maas P, Barrdahl M, Joshi AD, et al. Breast cancer risk from modifiable and nonmodifiable risk factors among white women in the United States. *JAMA Oncol*. 2016;2(10):1295–1302.
17. Petracci E, Decarli A, Schairer C, et al. Risk factor modification and projections of absolute breast cancer risk. *J Natl Cancer Inst*. 2011;103(13):1037–1048.
18. Pfeiffer RM, Park Y, Kreimer AR, et al. Risk prediction for breast, endometrial, and ovarian cancer in white women aged 50 y or older: derivation and validation from population-based cohort studies. *PLoS Med*. 2013;10(7):e1001492.
19. Gail MH, Pfeiffer RM. Breast cancer risk model requirements for counseling, prevention, and screening. *J Natl Cancer Inst*. 2018;110(9):994–1002.
20. Lee A, Mavaddat N, Wilcox AN, et al. BOADICEA: a comprehensive breast cancer risk prediction model incorporating genetic and nongenetic risk factors. *Genet Med*. 2019;21(6):1462.
21. Cuzick J, Brentnall A. Commentary on Terry et al., 10-year performance of four models of breast cancer risk: a validation study. *Lancet Oncol*. 2019;20(4):504–17. *Med One*. 2019;4:e190006.
22. Garcia-Closas M, Gunsoy NB, Chatterjee N. Combined associations of genetic and environmental risk factors: implications for prevention of breast cancer. *J Natl Cancer Inst*. 2014;106(11):dju305.
23. Travis LB, Hill D, Dores GM, et al. Cumulative absolute breast cancer risk for young women treated for Hodgkin lymphoma. *J Natl Cancer Inst*. 2005;97(19):1428–1437.