

Risk Assessment Models for Breast Cancer

Eun Young Chae

*Department of Radiology, Research Institute of Radiology, Asan Medical Center,
University of Ulsan College of Medicine, Seoul, Korea*

Although mammography is the only proven screening modality that decreases breast cancer mortality, women with an elevated risk for breast cancer may benefit from more intensive surveillance or risk-reduction strategies. For these women, supplemental screening with MRI or ultrasound increases cancer detection. A personalized risk assessment will help provide individual screening protocols according to patient risk. Over the last two decades, various mathematical models have been designed and validated to assess breast cancer risk based on family history and other risk factors. This review examines some of the most commonly used models and discusses their advantages, limitations, and other details.

Index words: Risk assessment model; Breast cancer screening; Risk factors; Breast neoplasms

1. Importance of Risk Assessment

Breast cancer is one of the most common cancers in Korean women. There were 22,395 newly diagnosed breast cancers in Korea in 2017, and, according to the Korea National Cancer Incidence Database, the crude cancer incidence was 86.9 per 100,000 women (1). The breast cancer incidence rate has been continuously increasing since 1999, and research suggests that it will continue to increase significantly due to rapidly changing lifestyles and an increase in the proportion of older individuals

over the last decade (2). This trend emphasizes the importance of efficient breast cancer screening programs in Korea.

According to the recommendations for breast cancer screening from the American Cancer Society (ACS) and the U.S. Preventive Services Task Force (USPSTF) (3, 4), there is a growing consensus that screening should be tailored to patient risk. The paradigm has moved from mammography screening alone to a personalized protocol that considers patient risk factors. Therefore, it is important to identify women that are at increased risk for breast cancer because they need more intensive screening (typically including an MRI) and should start screening at a younger age than the general population. Supplemental screening with MRI or ultrasound has been shown to increase breast cancer detection, particularly in women with an elevated

통신저자: Eun Young Chae, MD
Department of Radiology, Research Institute of Radiology, Asan Medical Center, University of Ulsan College of Medicine, 88, Olympic-ro 43-Gil, Songpa-Gu, Seoul 05505, Korea
Tel. (02) 3010-1709 Fax. (02) 476-0090
E-mail: chaeey@hanmail.net

risk. The ACS recommends an annual screening MRI for women with a greater than 20% lifetime risk of breast cancer (3).

2. Risk Factors

Family History of Breast Cancer

Family history is an important risk factor for breast cancer and should be as detailed as possible. A qualified family history of breast cancer requires the following information: age at the time of breast cancer diagnosis, degree of relationship (first or greater), multiple family members with breast cancer (particularly on one side), bilateral breast cancer, other early-onset tumors (e.g., sarcoma or ovarian cancer), and the number of unaffected family members. For women with a history of breast cancer in a first-degree relative, there is a twofold increase in the risk of developing the disease compared with women without any affected relatives (5). Younger age at the time of breast cancer diagnosis in a family member is also associated with an increased risk of developing breast cancer (6). That is, there is a greater breast cancer risk in women with first-degree relatives diagnosed before age 50 compared to those with relatives diagnosed after age 50.

Reproductive and Hormonal Risk Factors

Reproductive and hormonal factors are well-known to increase breast cancer incidence. Longer exposure to endogenous estrogen, from an early age at menarche (< 12 years) or late age at menopause (> 55 years), increases breast cancer risk (5). The relationship between oral contraceptives and breast cancer has been explored. The use of long-term combined oral contraceptives (i.e., estrogen plus progestin for more than five years) in postmenopausal women is related to an increased risk of breast cancer (7). Different oral contraceptive formulations may affect the risk of breast cancer

differently; that is, the risk associated with an estrogen-only formulation of oral contraceptive is much lower than that associated with combined hormone replacement therapy (8).

Mammographic Density

Breast density on mammography has been recognized as an independent risk factor for breast cancer (9, 10). Research has consistently reported that women with dense breast tissue have a four-to-six-fold increased risk for breast cancer compared to those with fatty breast tissue (9, 11). The relationship between mammographic parenchymal density and the risk of breast cancer is consistent regardless of age at mammography or the ethnic background of the study population (12-14). In Asian countries, dense breast tissue is more common and the proportion of breast cancer in younger individuals is higher than that in the Western population (15-17).

Park et al. (18) investigated the effects of mammographic density on breast cancer risk in a Korean population using nationwide mammographic screening program data. They found that the risk of breast cancer for women with extremely dense breasts was five times higher than that in women with an almost entirely fatty breast, which is consistent with results from research in Western countries. When the effect of breast density was calculated based on menopausal status, the odds ratio of breast cancer for an extremely dense breast was 8.5 (95% confidence interval (CI)=1.9-36.7) for premenopausal women and 3.8 (95% CI=2.8-5.1) for postmenopausal women. The authors reported that the positive relationship between mammographic density and the risk of breast cancer was stronger in premenopausal women.

3. Risk Assessment Models

A risk assessment model is a mathematical tool

to estimate the probability that a currently healthy individual with certain risk factors will develop a specific disease, such as breast cancer, over a certain period (e.g., within five years or over their lifetime) (19). Here we review some of the most commonly used models.

Gail Model

The Gail model is the most commonly used and widely known model for breast cancer risk prediction (20). It was developed in 1989 and modified in 1999 (21). This model was designed using data from the Breast Cancer Detection Demonstration Project (BCDDP), a screening study including almost 300,000 women aged 35 to 74 years between 1973 and 1980. The modified version is called the National Cancer Institute–Gail model or the Breast Cancer Risk Assessment Tool (BCRAT). The modified version differs from the original model in that it contains only invasive breast cancers, whereas the original version contained both DCIS and invasive cancers. The age-specific incidence in the modified version was based on the National Cancer Institute's Surveillance, Epidemiology, and End Results Program rather than the BCDDP, as in the original model.

This model focuses on non-genetic risk factors and contains limited information on family history of breast cancer. Both the original and modified Gail models use six input variables, including age, hormonal or reproductive factors (age at menarche and first live birth), family history (number of first-degree female relatives with breast cancer), and previous history of breast disease (number and results of breast biopsy). It is widely and easily available online at <http://www.cancer.gov/bcrisktool/>, and calculates invasive breast cancer risk over the next five-years and the lifetime breast cancer risk to age 90. The NCCN guidelines support using the Gail model to identify candidates for risk management, such as chemoprevention, if the five-

year risk is 1.67% or greater (22).

The Gail model has some specific limitations. It does not give an accurate risk prediction for women with certain conditions, such as those who already had breast cancer, LCIS, or DCIS. It also cannot be used for women with a previous history of chest wall irradiation treatment or known genetic mutations related to breast cancer. The model cannot be used in women younger than 35 years of age. Another limitation is that the model only asks how many first-degree relatives have had breast cancer and not about the age at breast cancer diagnosis. It also does not account for a family history of breast cancer on the paternal side or a family history of male breast cancer. It therefore may underestimate the risk in women who have 2nd-degree relatives with breast cancer and overestimate the risk for women who have had a benign breast biopsy (23, 24).

Claus Model

The Claus model focuses on a family history of breast cancer. It was designed using data from the Cancer and Steroid Hormone Study (25), a nested population-based case-control study including 4,730 patients with breast cancer aged 20–54 years matched with 4,688 controls between 1980 and 1982. Originally, the model only included data on family history of female breast cancer. A modified version has been validated by adding bilateral disease, three or more affected relatives, and ovarian cancer (26).

The Claus model includes age and a thorough family history, including 1st- and 2nd-degree relatives, and the age at which cancers in those relatives were diagnosed. Unlike the Gail model, the Claus model includes family history on both the maternal and paternal sides of the family. The model results are presented in a series of tables that show the lifetime risk of both invasive cancer and DCIS. Although this table does not provide risk estimates for some relative combinations, an

estimation of breast cancer risk is based on similar combinations, such as paternal aunt and mother.

The major limitation of the Claus model is that it does not include breast cancer risk factors other than family history (e.g., reproductive or hormonal factors) and does not incorporate male relatives with breast cancer. Another potential limitation is that it was developed using data from a North American population between 1980 and 1982. Because the incidence of breast cancer has increased, the model results are lower than the current incidence (27, 28). Therefore, some experts suggest an upward adjustment of 3-4% for the lifetime risk is needed for those lifetime risks below 20%, as calculated by this model. The model may not be applicable for women with known genetic mutations, such as *BRCA*, because it may underestimate their risk.

Tyrer-Cuzick Model

The Tyrer-Cuzick model was developed using data from the United Kingdom's International Breast Intervention Study (IBIS) (29, 30). This model includes a comprehensive family history of breast and ovarian cancer as well as a wide variety of individual risk factors (31). Free software is available at: <http://www.ems-trials.org/riskevaluator/>. The model includes individual risk factors such as age, height, weight, history of breast biopsies and their results, hormone replacement use, and a comprehensive maternal and paternal family history of breast and ovarian cancer, including age at diagnosis. It also includes family size and the number of unaffected relatives, which may affect the risk prediction. It also factors in positive or negative *BRCA* gene mutation test results. The model output is presented as the breast cancer risk over the next 10 years and lifetime risk for both invasive cancer and DCIS. It is based on a dataset from Caucasian high-risk females and does not adjust for a non-Caucasian population. The model may overestimate risk for women with previous LCIS results, atypical

hyperplasia (32), and no family history of breast cancer, and it may underestimate risk for women with a strong family history of breast cancer.

BRCAPRO Model

The BRCAPRO model calculates lifetime breast cancer risk as well as the probability of having a *BRCA* 1 or 2 mutation based on the Bayesian rules (33). One advantage of the BRCAPRO model is that it incorporates information on both affected and unaffected relatives. However, it requires a full pedigree, making it difficult to use in a clinic. Like the Claus model, it does not include information about non-hereditary risk factors and therefore may underestimate risk for some women. Another limitation is that no other genetic mutation that contributes to breast cancer development is included. It may underestimate the risk for women with negative *BRCA* results and a family history of breast cancer only (34).

Model Performance

It is important to validate risk assessment models before using them. Two criteria, calibration and discrimination, are often used to evaluate the performance of prediction models (35). Calibration means how well a model predicts risk overall in a population, and it can be presented as the E/O ratio (expected number (E) from the model compared to the observed number (O) of breast cancer cases.) A well-calibrated model should have an E/O ratio near 1. Discrimination indicates how well a model predicts risk at the individual level. It is the probability that a randomly selected case with breast cancer will have a higher predicted risk than a randomly selected non-case without breast cancer. An AUC equal to 1 indicates perfect discrimination, and an AUC near 0.5 indicates random chance.

A recent study by McCarthy et al. (36) evaluated the performance of the BRCAPRO, Gail, Claus,

Breast Cancer Surveillance Consortium (BCSC), and Tyrer-Cuzick models in predicting risk of breast cancer in 35,921 women aged 40–84 years who underwent mammography screening between 2007 and 2009. In this study, the Gail, BCSC, and BRCAPRO models were well-calibrated (CIs include 1.0), but the Tyrer-Cuzick model slightly overestimated risk, especially when breast density was included (O/E = 0.84, 95% CI = 0.79–0.91). The Claus model substantially underestimated risk (O/E = 1.69, 95% CI = 1.48–1.87). The AUC values showing discrimination accuracy were comparable and moderate (0.59–0.64). Among women with mammographic density information (n = 30,970), the BCSC model had the highest discrimination accuracy (O/E = 0.97, 95% CI = 0.89–1.05, AUC = 0.64, 95% CI = 0.62–0.66), despite its simplicity. The Claus model performed worse than the other models.

The UK Generation Study, a large population-based study including 64,874 women aged 35–74 years (37), provided more data on model performance in a general population. The Gail model slightly underestimated risk (O/E=1.09, 95% CI=1.02–1.16), whereas the Tyrer-Cuzick model slightly overestimated risk (O/E=0.88, 95% CI=0.83–0.94). The respective AUC values were 0.61 and 0.63. The two general-population cohorts (36, 37) indicate that Gail, BCSC, BRCAPRO, and Tyrer-Cuzick are reasonably well-calibrated and have a moderate level of discrimination accuracy in a general population.

4. Breast Cancer Risk Assessment in Korean Women

Many risk-prediction models have been developed, and each has strengths and limitations. Different risk assessment models may provide different outcomes for the same woman (38). Most of the established prediction models estimating breast cancer risk were developed using data from a Western population.

However, there are considerable differences in the epidemiological characteristics of breast cancer cases in Koreans compared to those in a Western population (39). According to the Korean Breast Cancer Society and Korea Central Cancer Registry, the incidence of breast cancer in Korean women is highest in those aged 40–49 years (40), whereas the incidence in the United States is highest in women aged 60–69, followed by those aged 50–59, and the median age at breast cancer diagnosis was 62 (41).

Therefore, a risk assessment model based on the epidemiological characteristics of a Korean population must be developed and validated. A Korean Breast Cancer Risk Assessment Tool (KoBCRAT) was developed in 2013 based on the Seoul Breast Cancer Study (42), a case-control study based on equations developed for the Gail model. This model stratified risk factors by age group (< age 50 and ≥ age 50). The KoBCRAT had a discriminatory accuracy of 0.63 in women aged < 50 years and 0.65 in those aged ≥ 50 years, suggesting modest discrimination power between breast cancer cases and controls. In the validation analyses, the AUC was 0.61 for the Korean Multicenter Cancer cohort and 0.89 for the National Cancer Center cohort. However, this model does not include image-based information, such as breast density, to assess breast cancer risk. To accurately identify Korean women at an elevated breast cancer risk, a breast cancer risk model must be developed that includes breast density.

Summary

Breast cancer risk assessment models estimate risk to develop personalized management options. These models always have limitations and must be improved to individualize breast cancer risk prediction. The risk prediction model that includes breast density will help improve outcomes. In the future, it is necessary to develop a breast cancer risk assessment model targeting Korean females that

includes breast density information.

References

1. Hong S, Won YJ, Park YR, et al. Cancer Statistics in Korea: Incidence, Mortality, Survival, and Prevalence in 2017. *Cancer Res Treat* 2020;52:335-350
2. Lee JE, Lee SA, Kim TH, et al. Projection of Breast Cancer Burden due to Reproductive/Lifestyle Changes in Korean Women (2013-2030) Using an Age-Period-Cohort Model. *Cancer Res Treat* 2018;50:1388-1395
3. Oeffinger KC, Fontham ET, Etzioni R, et al. Breast Cancer Screening for Women at Average Risk: 2015 Guideline Update From the American Cancer Society. *JAMA* 2015;314:1599-1614
4. Siu AL, Force USPST. Screening for Breast Cancer: U.S. Preventive Services Task Force Recommendation Statement. *Ann Intern Med* 2016;164:279-296
5. Collaborative Group on Hormonal Factors in Breast C. Familial breast cancer: collaborative reanalysis of individual data from 52 epidemiological studies including 58,209 women with breast cancer and 101,986 women without the disease. *Lancet* 2001;358:1389-1399
6. Brewer HR, Jones ME, Schoemaker MJ, Ashworth A, Swerdlow AJ. Family history and risk of breast cancer: an analysis accounting for family structure. *Breast Cancer Res Treat* 2017;165:193-200
7. Beral V, Million Women Study C. Breast cancer and hormone-replacement therapy in the Million Women Study. *Lancet* 2003;362:419-427
8. Ross RK, Paganini-Hill A, Wan PC, Pike MC. Effect of hormone replacement therapy on breast cancer risk: estrogen versus estrogen plus progestin. *J Natl Cancer Inst* 2000;92:328-332
9. Boyd NF. Tamoxifen, mammographic density, and breast cancer prevention. *J Natl Cancer Inst* 2011;103:704-705
10. Boyd NF, Martin LJ, Yaffe MJ, Minkin S. Mammographic density and breast cancer risk: current understanding and future prospects. *Breast Cancer Res* 2011;13:223
11. Vachon CM, van Gils CH, Sellers TA, et al. Mammographic density, breast cancer risk and risk prediction. *Breast Cancer Res* 2007;9:217
12. McCormack VA, dos Santos Silva I. Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis. *Cancer Epidemiol Biomarkers Prev* 2006;15:1159-1169
13. Kerlikowske K, Cook AJ, Buist DS, et al. Breast cancer risk by breast density, menopause, and postmenopausal hormone therapy use. *J Clin Oncol* 2010;28:3830-3837
14. Eriksson L, Czene K, Rosenberg L, Humphreys K, Hall P. Possible influence of mammographic density on local and locoregional recurrence of breast cancer. *Breast Cancer Res* 2013;15:R56
15. Youn I, Choi S, Kook SH, Choi YJ. Mammographic Breast Density Evaluation in Korean Women Using Fully Automated Volumetric Assessment. *J Korean Med Sci* 2016;31:457-462
16. Ursin G, Ma H, Wu AH, et al. Mammographic density and breast cancer in three ethnic groups. *Cancer Epidemiol Biomarkers Prev* 2003;12:332-338
17. del Carmen MG, Halpern EF, Kopans DB, et al. Mammographic breast density and race. *AJR Am J Roentgenol* 2007;188:1147-1150
18. Park B, Cho HM, Lee EH, et al. Does breast density measured through population-based screening independently increase breast cancer risk in Asian females? *Clin Epidemiol* 2018;10:61-70
19. Meads C, Ahmed I, Riley RD. A systematic review of breast cancer incidence risk prediction models with meta-analysis of their performance. *Breast Cancer Res Treat* 2012;132:365-377
20. 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;81:1879-1886
21. 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;91:1541-1548
22. National Comprehensive Cancer Network. Genetic/familial high-risk assessment: breast and ovarian (version 1.2020). Available at: https://www.nccn.org/professionals/physician_gls/pdf/genetics_screening.pdf. Accessed November 15, 2020.
23. Euhus DM, Leitch AM, Huth JF, Peters GN. Limitations of the Gail model in the specialized breast cancer risk assessment clinic. *Breast J* 2002;8:23-27
24. Pankratz VS, Hartmann LC, Degnim AC, et al.

- Assessment of the accuracy of the Gail model in women with atypical hyperplasia. *J Clin Oncol* 2008;26:5374-5379.
25. Claus EB, Risch N, Thompson WD. Genetic analysis of breast cancer in the cancer and steroid hormone study. *Am J Hum Genet* 1991;48:232-242
26. Claus EB, Risch N, Thompson WD. The calculation of breast cancer risk for women with a first degree family history of ovarian cancer. *Breast Cancer Res Treat* 1993;28:115-120
27. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. *CA Cancer J Clin* 2020;70:7-30
28. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018;68:394-424
29. Tyrer J, Duffy SW, Cuzick J. A breast cancer prediction model incorporating familial and personal risk factors. *Stat Med* 2004;23:1111-1130
30. Berry DA, Parmigiani G, Sanchez J, Schildkraut J, Winer E. Probability of carrying a mutation of breast-ovarian cancer gene BRCA1 based on family history. *J Natl Cancer Inst* 1997;89:227-238
31. Barke LD, Freivogel ME. Breast Cancer Risk Assessment Models and High-Risk Screening. *Radiol Clin North Am* 2017;55:457-474
32. Boughey JC, Hartmann LC, Anderson SS, et al. Evaluation of the Tyrer-Cuzick (International Breast Cancer Intervention Study) model for breast cancer risk prediction in women with atypical hyperplasia. *J Clin Oncol* 2010;28:3591-3596
33. Parmigiani G, Berry D, Aguilar O. Determining carrier probabilities for breast cancer-susceptibility genes BRCA1 and BRCA2. *Am J Hum Genet* 1998;62:145-158
34. Amir E, Freedman OC, Seruga B, Evans DG. Assessing women at high risk of breast cancer: a review of risk assessment models. *J Natl Cancer Inst* 2010;102:680-691
35. Gail MH. Choosing Breast Cancer Risk Models: Importance of Independent Validation. *J Natl Cancer Inst* 2020;112:433-435
36. 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;112:489-497
37. Pal Choudhury P, 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;112:278-285
38. Ozanne EM, Drohan B, Bosinoff P, et al. Which risk model to use? Clinical implications of the ACS MRI screening guidelines. *Cancer Epidemiol Biomarkers Prev* 2013;22:146-149
39. Kim DY, Park HL. Breast Cancer Risk Prediction in Korean Women: Review and Perspectives on Personalized Breast Cancer Screening. *J Breast Cancer* 2020;23:331-342
40. Kang SY, Kim YS, Kim Z, et al. Breast Cancer Statistics in Korea in 2017: Data from a Breast Cancer Registry. *J Breast Cancer* 2020;23:115-28
41. DeSantis CE, Ma J, Gaudet MM, et al. Breast cancer statistics, 2019. *CA Cancer J Clin* 2019;69:438-451
42. Park B, Ma SH, Shin A, et al. Korean risk assessment model for breast cancer risk prediction. *PLoS One* 2013;8:e76736