# Breast Cancer Risk Prediction Model for Health Risk Assessment Module
## Executive Summary

[Adapted from the report by DR ROZA SARIMIN]

**Authors:**
Dr Roza Sarimin
Dr Izzuna Mudla Mohamed Ghazali
Dr Khadijah Abd Rahim

**Expert Committee:**
Datin Dr Rugayah Bakri
Dr Nor Aina Emran
Dr Anita Baghawi
Dr Normayah Kitan
Dr Keng Wee Teik
Dr Ch'ng Gaik Siew
D Suzaini Mat Daud
Dr Nor Saleha Ibrahim Tamin
Dr Zaleha A Hamid
En Azman Mohammed
Pn Norhaslida Mohd Noor
Dr Junainah Sabirin

**External Reviewers:**
Dato' Dr Cheng-Har Yip
Professor Dr Nur Aishah Taib
Professor Dr Teo Soo Hwang

**Information Specialist:**
Dr Nur Farhana Mohamad



For further information please contact:
Health Technology Assessment Section (MaHTAS)
Medical Development Division
Ministry of Health Malaysia
Level 4, Block E1, Precinct 1
Government Office Complex
62590 Putrajaya.
Tel: 603 8883 1246
Fax: 603 8883 1230

Available at the following website:
http://www.moh.gov.my

**2015**

### Background

Breast cancer is the commonest cancer among women and the commonest cause of cancer death worldwide. Similarly in Malaysia, it was reported to be the commonest cancer in women with overall age standardized incidence rate (ASR) of 29.1 per 100,000 population (National Cancer Registry 2007), higher than incidence in other developing countries (ASR of 20 per 100,000 population). About 30 to 40% of Malaysian women presented at a later stage of breast cancer (stage III and IV). Hence, affordable and effective approaches in cancer control are needed for early detection, diagnosis and treatment of breast cancer particularly in the less developed countries. The growing public awareness of breast cancer and its risk factors, with availability of medical and surgical risk reduction options has led consultation of many women on their breast cancer risk. Considerable effort has been directed at identifying risk factors and developing risk prediction models for breast cancer. Prediction model is a mathematical equation designed to quantify the risk an individual woman would develop a particular cancer in a defined period. It provides an estimation of disease risk that can be used to guide management for women at all level of risks. Multiple prediction models have been developed to assist with breast cancer risk prediction efforts. Reliable accurate prediction models can inform future disease burdens, health policies, individual decisions on future screening behaviour and adoption of risk reduction strategies, counsel those at risk, design prevention strategies for at risk populations and plan intervention trials. Subsequently, the adoption of such models should guide decision-making, improve patient outcomes and the cost-effectiveness of care. Given the variance in breast cancer risk, surveillance and primary prevention adapted to individual risk level may be the most effective use of resources for preventing, detecting and improving breast cancer survival. In Malaysia, currently opportunistic screening policy is being practiced for breast cancer by means of both primary and secondary prevention approach. There is no risk prediction model/health risk assessment (HRA) in predicting individual risks of developing breast cancer in existing cancer control approaches. Therefore, introduction of a risk assessment/risk prediction model for breast cancer is timely in addressing unmet needs of identifying high risk individuals in the Malaysian community, towards enhancing early detection of breast cancer in facilitating more effective cancer control strategies. This assessment was requested by a Senior Principal Assistant Director, Health Education Division, Ministry of Health.

### Technical features

A risk prediction model is a statistical tool for estimating the probability that a currently healthy individual with specific risk factors will develop a future condition, such as breast cancer, within a certain time. It uses multiple predictors (covariates) to estimate the probability or risk that a certain outcome is present. The goal of risk prediction is to provide individualised risk (absolute risk) with associated measures of uncertainty thus stratifying individuals by these risks. It is accomplished by combining the baseline risk of developing the condition with an individual's risk score. The baseline risk of the condition represents the underlying population risk for patients whose risk factor values are not present, which is usually estimated from a prospective population-based cohort study. The risk-score component shows how much the baseline risk is multiplied for increasing values of the risk factors, which may also be estimated using a cohort study or for rare conditions, a case-control study. The variables in the model can be any combination of environmental, behavioural, genetic or psychological attributes of the person. Developed models need to provide accurate and validated (internally and externally; temporal, geographical and domain/setting) estimates of disease probabilities. Performance of predictive tests is commonly measured by means of calibration (the ability to predict the number of events in subgroups of the

population) and discrimination (the ability to distinguish at the individual level between those who will develop the disease and those who will not). Calibration performance is commonly reported by E/O statistics comparing expected (E) and observed (O) number of events, and discrimination performance by concordance (c)-statistics. Performance of the model may also vary according to the population they are applied to. Cancer risk prediction model is classified into absolute risk prediction model aimed at assessing probability that an individual with given risk factors and a given age will develop cancer over a defined period of time (such as Gail model); and gene carrier status risk prediction models aimed at assessing mutation probability of an individual or carrying a gene mutation that predisposes to a particular cancer (such as BOADICEA, BRCAPRO, Cuzick-Tyrer (IBIS models) and Manchaster Scoring System). Each model has unique attributes stemming from the methodology, sample size and population characteristics used to create the model.

**Policy question**

- In Ministry of Health, should a breast cancer risk prediction model for health risk assessment (HRA) module be introduced as one of the strategies in the prevention of breast cancer under the Malaysia National Cancer Control Programme?
- If breast cancer risk prediction model for HRA module is to be introduced, which risk prediction model should be adopted in Malaysia?

**Objectives**

- To assess the effectiveness in term of predictive accuracy of breast cancer risk assessment/prediction models among women
- To assess the safety, organizational, ethical issues and economic implications related to risk assessment/prediction models for breast cancer among women

**Methods**

Studies were identified by searching electronic databases. The following databases were searched through the Ovid interface: MEDLINE(R) In-process and other Non-Indexed Citations and Ovid MEDLINE(R) 1946 to present. EBM Reviews-Cochrane Database of Systematic Reviews (2005 to May 2015), EBM Reviews-Cochrane Central Register of Controlled Trials (May 2015), EBM Reviews – Database of Abstracts of Review of Effects (2nd Quarter 2015), EBM Reviews-Health Technology Assessment (2nd Quarter 2015), EBM Reviews-NHS Economic Evaluation Database (2nd Quarter 2015). Parallel searches were run in PubMed. Appendix 3 showed the detailed search strategies. No limits were applied to the search. The last search was run on 20 March 2015. From cross referencing of retrieved articles, additional articles were identified. All relevant literature was appraised using the Critical Appraisal Skills Programme (CASP) tool. All full text articles were graded based on guidelines from the U.S./Canadian Preventive Services Task Force.

**Results and conclusions**

A total of 830 titles were identified through Ovid interface, Pubmed and references of retrieved articles. 123 abstracts were screened using the inclusion and exclusion criteria resulting in 58 potentially relevant articles. After the critical appraisal, only 14 full texts out of the above were finally included in this review comprising of two systematic reviews, one randomised controlled trial and eleven observational studies (eight cohorts and three cross sectional). No evidence on cost-effectiveness or cost-utility analysis was retrieved. Of these, eleven articles were related to the effectiveness (predictive accuracy) of risk assessment/ risk prediction/ health risk assessment models for breast cancer, which were two systematic reviews and nine observational studies (eight cohort and two cross sectional). The other two articles were related to safety of risk prediction models

for breast cancer. The articles were published between 2009 and 2015. The included systematic reviews were both published in 2012. Most of the observational studies were conducted in the U.S.A., with one each from Italy, Singapore, Thailand and Korea. The total pooled sample size of included studies was 437,049 subjects. Sample sizes for each of the observational studies ranged from 690 to 135,329 subjects. The length of follow-up ranged from five years to ten years.

**Effectiveness (predictive accuracy)**
Models being assessed were Gail model (also known as Breast Cancer Risk Assessment Tool) including Gail 1 and Gail 2 (updated) model, Contraceptive and Reproductive Experience (CARE) model, model by Petracci, model by Pfeiffer, Vermont model, model by Anothaisintawee and BWHS model.

Performance of prediction model is commonly measured by means of calibration and discrimination. A well-fitted model has Expected/Observed (E/O) ratio close to 1, a number lower underestimates the condition's incidence and a number higher overestimates the incidence. The concordance (c)-statistics measure model discrimination performance which is similar to area under the receiver operating characteristic curve (ROC). A c-statistics of 1.0 indicates perfect discrimination and 0.5 equivalents to no discrimination between people who develop the condition and those who do not.

**Performance of Gail model**
(also known as Breast Cancer Risk Assessment Tool, BCRAT)
Gail 1 model estimates the absolute risk for both invasive and insitu breast cancer (ductal and lobular), derived from data of the Breast Cancer Detection Demonstration project. Gail 2 model is a modification of Gail 1, used to estimate the risk of invasive breast cancer using data from Surveillance, Epidemiology and End Result (SEER) Programme of the US National Cancer Institute.

**i.Calibration performance**

- Validation cross-population showed underestimation of breast cancer cases in two studies, overestimation of breast cancer cases was observed in one study and pool result showed it has good calibration
- Pool E/O ratio was 1.13(95% CI 0.80 to 1.60) for Gail 1 model (I-squared: 95.0%) and 0.95 (95% CI 0.92 to 1.02) for Gail 2 model (I-squared: 92.5%)
- E/O ratio ranges from 0.86(95% CI 0.82 to 0.90), 0.87(95% CI 0.85 to 0.89)33 in white post-menopausal women aged 50 to 74 years, to E/O ratio of 0.85(95% CI 0.83 to 0.88) in Hispanic and non-Hispanic post-menopausal women aged 50 to 79 years; both in US for Gail 2 model
- E/O ratio was 1.85(95% CI 1.68 to 2.04) among women aged 50 to 64 in Singapore 37 for Gail 2 model

    - Updating the Gail 2 model improved models calibration
- E/O ratio of 1.03 (95% CI 1.00 to 1.05)
  [using Surveillance Epidemiology End Result (SEER) breast cancer incidence and mortality 1995 to 2003)]
- E/O ratio of 1.00 (95% CI 0.94 to 1.08) (using SEER 2003 to 2006)

**ii.Discriminative ability**

- Pool concordance (c)-statistics was 0.63 for Gail 2 model (I-squared: 94.5%)
- Overall area under curve (AUC) and c-statistics ranges from 0.54 to 0.63 for Gail 2 model in the three validation studies in the US
- [0.58 in the US white post-menopausal women aged 50 to 74 years, between 0.58 (95% CI 0.56 to 0.58) for non-Hispanic white and 0.63 (95% CI 0.58 to 0.67) for Hispanic US postmenopausal women aged 50 to 7936

and 0.54 (95% CI 0.52 to 0.56) among Vermont women aged 70 years and older]

**Performance of Contraceptive and Reproductive Experience (CARE) model**

**i. Calibration performance**

- Validation in a single study showed underestimation of breast cancer cases with E/O ratio of 0.88 (95% CI 0.82 to 0.94) among black women aged 30 to 69 years 38

**ii. Discriminative ability**

- Age-specific c-statistics for total invasive breast cancer was 0.57 (95% CI 0.55 to 0.59)
- Age-adjusted c-statistics for specific breast cancer subtypes:
  0.59(95% CI 0.57 to 0.61)(Estrogen receptor positive (ER+) breast cancer)
  0.54(95% CI 0.50 to 0.57)(Estrogen receptor negative (ER-) breast cancer)

**Performance of model by Petracci**

**i. Calibration performance**

- Validation in a single study showed the model was well calibrated with E/O ratio of 1.10 (95% CI 0.96 to 1.26) among Italian women aged 35 to 64 years

**ii. Discriminative ability**

- c-statistics was 0.62 (95% CI 0.55 to 0.69)(women younger than 50 years)
- c-statistics was 0.57(95% CI 0.52 to 0.61)(women more than 50 years)

**Performance of model by Pfeiffer**

**i. Calibration performance**

- Validation in a single study showed the model was well calibrated with E/O ratio of 1.00 (95% CI: 0. 96 to1.04) among white women aged 30 to 55 years

**ii. Discriminative ability**

- Overall AUC was 0.58 (95% CI 0.57 to 0.59) in the above validated population

**Performance of Vermont model**

**i. Discriminative ability**

Validation in a single study showed the model discriminative power was modest with c-statistics of 0.55 (95% CI 0.53 to 0.58) among Vermont women aged 70 years and older

**Performance of model by Anothaisintawee**

**i. Calibration performance**

- Validation in a single study showed the derived model was well calibrated with O/E ratio of 1.00 (95% CI 0.82 to 1.21)

**ii. Discriminative ability**

- C-statistics (overall) was 0.65 (95% CI 0.59 to 0.70) for prediction of breast cancer cases

**Performance of BWHS model**

**i. Calibration performance**

- Validation in a single study showed the BWHS model was well calibrated with E/O ratio of 0.96 (95% CI 0.88 to 1.05)

**ii. Discriminative ability**

- AUC (overall) was 0.59 (95% CI 0.56 to 0.61)

**Safety**
The only reported adverse event was anxiety. Women who had higher risk status had an odd of having increased anxiety about five times greater than women who had lower risk status (OR 5.03; 95% CI 1.54 to 16.43)

**Cost implication**
No evidence retrieved on breast cancer risk assessment/prediction model. Potential direct cost implicated on the designing, developing, testing and commissioning of available one breast cancer prediction model into a health risk assessment module was given at approximately RM75,000

**Organizational**
Computerized risk estimate using any model requires computer literate user/patient and internet access.

Cancer risk prediction model needs to be continually calibrated and revalidated. The complexity of prediction modelling research from developing and internally validating a prediction model, testing, adjusting or updating the model for other individuals (external validation); and assessing its impact on therapeutic management and patient outcomes need a dedicated research expertise. Uncertainties associated with risk estimates should be addressed and informed particularly when clinical decision has serious consequences.

Among ethical issues that arose following cancer risk assessment was psychological harms resulting from being labelled 'at risk', and additional diagnostic procedures that can artificially increase sense of risk. The 'at risk' label also has implication for future health care cost. Theoretically increased psychological distress from risk labelling may contribute to other healthcare demands and raising the health care cost.

**Conclusion**
There was sufficient good level of retrievable evidence for breast cancer risk prediction model. There were six models identified for predicting breast cancer risk with Gail model is the most widely studied and validated model in various population. The Gail model appeared to have good calibration in validation studies done cross-population; however there is considerable heterogeneity across studies. This model showed moderate performance in terms of discriminatory ability.

For other risk prediction models, there was insufficient good level of retrievable evidence with only one study each of those other models (CARE model, model by Petracci, model by Pfeiffer, Vermont model, model by Anothaisintawee and BWHS model). The models were well calibrated in the validated population however appeared modest in discriminating woman who will be having breast cancer, than for those who will not in the study population.

There was insufficient evidence on the safety aspect of cancer risk prediction models for the detection of women who will develop breast cancer. Although there was minor adverse psychological sequale reported among high risk women who demonstrated to be five times more likely to have increased anxiety, it may be considered relatively safe.

There was no retrievable evidence on economic evaluation of health risk assessment or risk prediction model for breast cancer, or cost implication involved in developing a new risk prediction model without genetic component for breast cancer retrieved. The cost involved in validating a model by a prospective cohort validation study could be very costly depending on the number of study participants and years of follow up. However potential direct cost implicated to the

designing, developing, testing and commissioning of available one breast cancer prediction model was given approximately at RM75,000.

Cancer risk prediction models need continual validation to give meaningful risk estimate and to ensure its applicability in the setting it will be used. The complexity to develop and subsequently validate any breast cancer risk prediction model specifically models without genetic component is reflected in the necessary local data required and availability of dedicated research expertise to create a robust model with consistent performance.

**Recommendation**

Although the above review showed that the Gail model had good calibration and moderate discriminative ability, it is not suitable to be introduced as one of the strategy in the prevention of breast cancer under the Malaysian National Cancer Control Programme yet as it needs further validation until a well-fitted model that would have better predictive ability tailored to Malaysian population established. In addition, this model needs continual validation to determine the consistency of its performance.