ORIGINAL ARTICLE

Gail Model Risk Factors: Impact of Adding an Extended Family History for Breast Cancer

Anna Crispo, ScD,* Giuseppe D'Aiuto, MD,† MariaRosaria De Marco, MD,* Massimo Rinaldo, MD,† Maria Grimaldi, ScD,* Immacolata Capasso, MD,† Alfonso Amore, MD,† Cristina Bosetti, ScD,‡ Carlo La Vecchia, MD,‡§ and Maurizio Montella, MD*¶

**Servizio di Epidemiologia, Istituto Tumori "Fondazione Pascale," Naples, Italy; †Breast Unit, Istituto Tumori "Fondazione Pascale," Naples, Italy; ‡Istituto di Ricerche Farmacologiche "Mario Negri," Milan, Italy; §Istituto di Statistica Medica e Biometria, Università degli Studi di Milano, Milan, Italy; and ¶LILT, Comitato Regione Campania, Naples, Italy*

■ **Abstract:** An approach commonly used in estimating breast cancer risk is the Gail model. The objective of this study was to evaluate the feasibility and impact of adding extended family history as a new breast cancer risk factor into the Gail model. The data of the present study include cases with breast cancer and hospitalized controls recruited in the National Cancer Institute of Naples (southern Italy) between 1997 and 2000. We compared the first-degree relative (FDR) risk factor (standard Gail model) with the second-degree relative (SDR) information; and the FDR risk factor (standard Gail model) with the combination of FDR and SDR. We computed the c-statistic by comparing the risks found in our population to those in Gail-US population. The concordance for the model with FDR was 0.55 (95% CI 0.53–0.58), the model with SDR shows a modest but significant discriminatory accuracy (0.56, 95% CI 0.53–0.59), and the combination of FDR+SDR gave the concordance statistic of 0.57 (95% CI 0.54–0.60), indicating a good comparison between the two models. The results of our study show that extended family history information could be useful to improve the discriminatory power of the Gail model risk factors. ■

Key Words: breast cancer, concordance statistic, family history, Gail model

Breast cancer is the most common cancer among women (1,2). The estimated annual incidence of breast cancer worldwide is about one million cases, with ~200,000 cases in the United States (27% of all cancers in women) and ~320,000 cases in Europe (31% of all cancers in women) (1,3,4). Early detection of breast cancer is associated with both increased survival rates and invasive and less physically disfiguring treatments, thus emphasizing the importance of identifying those persons at high risk to screen for the disease appropriately (5).

A family history of breast cancer also increases a woman's risk of developing the disease. Women with one or more affected first-degree relatives (FDR) have an increased breast cancer risk when compared with women who do not have any affected relatives (6,7). In general, a twofold to threefold increase in the risk of breast cancer development has been associated with a family history of breast cancer in the mother or a sister (8–12). Claus et al. indicated that a women's risk of breast cancer development is strongly related not only to the presence of a positive family history of breast cancer but, more specifically, to the number and type of relatives affected with breast cancer (5).

An approach commonly used in estimating breast cancer risk is the Gail model (13). Based on a detailed case–control analysis of the Breast Cancer Detection Demonstration Project data set, Gail et al. determined the variables which identified women at risk for breast cancer—age at menarche; number of affected FDR; age at first live birth; number of biopsies (history of atypical hyperplasia); number of biopsies per age; and number of affected relatives per age at first live birth.

Address correspondence and reprint requests to: Anna Crispo, ScD, Servizio di Epidemiologia, Istituto Nazionale Tumori Fondazione "G. Pascale," Via M. Semola, 80131 Naples, Italy, or e-mail: epidemiologia.int@tin.it.

The Breast Journal, Volume 14 Number 3, 2008 221–227

The objective of this study was to evaluate the feasibility and impact of adding extended family history as a new breast cancer risk factor into the Gail model. We chose to incorporate second-degree relatives (SDR) with breast cancer in the model, and to then calculate relative risk based on the combination of FDR+SDR with breast cancer.

As data from case–control studies cannot be used to determine whether the absolute risk from the Gail model projections are correct, we used the original GM risk factors to determine whether the relative risks, for various risk factors in our data, increased when extended family history information was added to the model.

MATERIALS AND METHODS

Study Population

This case–control study included 1,765 women recruited in Italy between 1997 and 2000. Cases were 558 women, consecutively recruited from the Breast Unit of the National Cancer Institute of Naples, who had invasive breast tumors 3.5 cm or less in their largest diameter, with either negative or positive axillary lymph nodes (stage I or II breast cancer) (14,15). The control group included 1,207 women admitted to hospital in the same period as the first group. The choice of controls was made paying particular attention to exclude the following diagnoses—women admitted for gynecological, hormonal, or neoplastic diseases related to known risk factors for breast cancer. The distribution of cases and controls in terms of age and area of residence were similar, even though the cases and controls were not individually matched.

Definition of Family History

All patients filled out a questionnaire which included information on age, age at menarche, menopausal status, type of menopause, total number of live births, age at first live birth, age at first pregnancy (abortion or birth), height, weight, prior breast biopsies, and a history of atypical hyperplasia. Specific information about family history was also collected, including family history of breast cancer in FDR (mother, sister, daughter), and family history of breast cancer in SDR (aunts, grandmothers). For each relative with a history of cancer, the subjects were asked to specify whether the relative was still alive at the time of the interview, current age or age at death, site of the tumor, and age at diagnosis. We grouped the information on breast cancer among female FDR and SDR into the following categories:

(1) women with no reported first- or SDR with breast cancer (FH-) (reference category);

(2) women with 1 or more FDR with breast cancer (≥1 FDR);

(3) women with 1 or more SDR with breast cancer (≥1 SDR); and

(4) women with 1 or more first-degree and/or 1 or more SDR with breast cancer (≥1 FDR + ≥1 SDR).

Data Analysis: Gail Model

The Gail model risk factors were: age at menarche, age at first live birth, number of biopsies, number of biopsies per age, and number of affected relatives per age at first live birth.

We fitted a logistic model comparing our controls to cases. The model included terms for age, age at menarche, number of biopsies, and the combination of age at first live birth and family history. We used the case–control study for parameter estimation and cross validation.

We compared the coefficients from our logistic model with the original coefficients in Gail et al. (13). We evaluated the discriminatory accuracy of our statistical models through the estimation of the concordance statistic (c-statistic). This is an index of predictive discrimination based on rank correlation between predicted and observed outcomes (16). The accuracy represents the model's ability to separate individuals who will develop the disease from those who will be disease free (17). To do this, receiver operating characteristic (ROC) curves were generated. These are plots of the true positive rate against the false positive rate for different possible cutoff points. For binary outcomes, the c-statistic is indeed equal to the area under the ROC curve (AUC, 16,18). The AUC, and the corresponding 95% confidence intervals (CI), can be estimated nonparametrically from the empirical ROC curve without any distribution assumption of the test results (16,18). Therefore, if the lower bound of the 95% CI of AUC for a test is greater than 0.5, then the test is statistically significantly better (with a significance level of 0.05) than pure chance, which has an AUC of 0.5. All statistical analyses were performed using SAS (version 8.0) (19).

Table 1. Distribution of 558 Women with Breast Cancer and 1,207 Controls According to Selected Variables

	Controls	Breast cancer
Age		
No. reported	1207	558
No. unavailable	0	0
Mean (range) in years	53.1 (19–87)	55 (24–86)
Age at menarche		
No. reported	1207	554
No. unavailable	0	4
Mean (range) in years	12.7 (9–18)	12.5 (9–18)
Age at first live birth		
No. reported	1195	470
No. unavailable	12	88
Mean (range) in years	24.5 (17–41)	25.6 (15–42)
Women with FDR with breast cancer		
No. reported	1207	558
No. unavailable	0	0
1 FDR (%)	8.8	13.8
≥2 FDRs (%)	0.2	0.7
Women with SDR breast cancer		
No. reported	1207	558
No. unavailable	0	0
1 SDR (%)	7.5	13.8
≥2 SDRs (%)	0.8	2.9
Women with FDR and SDR with breast cancer		
No. reported	1207	558
No. unavailable	0	0
1 FDR+SDR (%)	15.7	25.4
≥2 FDRs+SDRs (%)	1.0	3.6
Women with a previous breast biopsy		
No. reported	1207	558
No. unavailable	0	0
Previous biopsy (%)	1.5	8.1
Menopause		
No. reported	1207	558
No. unavailable	0	0
Menopause yes (%)	49.6	41.8
Menopause no (%)	50.4	58.2
Mean (range) in years	49.6 (18–65)	48.8 (28–59)
OC use		
No (%)	66.4	77.8
Yes (%)	33.6	22.2

FH, family history; FDR, first-degree relative; SDR, second-degree relative; OC, oral contraceptive.

RESULTS

Table 1 gives the distribution of breast cancer cases, and the control group by selected variables affecting breast cancer risk. The mean age was 55 years for breast cancer patients (median, 54 years), and 53.1 years (median, 52 years) for the control group. There were no significant differences among groups according to age at menarche (mean = 12.5 and 12.6 years), and age at first live birth (mean = 24.5 and 25.6 years, respectively). Very few women were nulliparous. The proportion of women who reported one FDR with breast cancer was 13.8% in breast cancer cases, and 8.8% in the control group.

Table 2. Odds ratio (OR) and Corresponding 95% Confidence Interval (CI)* of Breast Cancer by Family History (FH)

FH	Breast cancer cases	Controls	OR (95% CI)
FDR with breast cancer			
No FH (FDR)	477	1099	1†
1 FDR	77	106	1.25 (0.83–1.87)
≥2 FDRs	4	2	4.79 (0.77–29.78)
SDR with breast cancer			
No FH (SDR)	465	1106	1†
1 SDR	77	91	2.39 (1.61–3.55)
≥2 SDRs	16	10	3.78 (1.62–8.80)
FDR+SDR with breast cancer			
No FH	396	1005	1†
1 F-S DR	132	184	1.59 (1.15–2.19)
2 F-S DRs	10	6	4.44 (1.49–13.17)
3 F-S DRs	12	8	3.15 (1.20–8.28)
≥4 F-S DRs	8	4	6.14 (1.75–21.56)

FH, family history; FDR, first-degree relative; SDR, second-degree relative; F-S DR, first-second-degree relative.
*Estimated from unconditional logistic regression adjusted for age, age at menarche, age at first live birth, number of births, number of previous biopsy.
†Reference category.

Corresponding values for more than 1 FDR were 0.7% and 0.2%, respectively. The proportion of women who reported one SDR with breast cancer was 13.8% in breast cancer cases, and 7.5% in the control group. Corresponding values for more than 1 SDR were 2.9% and 0.8%, respectively. The proportion of the combinations of FDR+SDR for cases and controls were 25.4% and 15.7% with one affected relative and 3.6% and 1.0% with two or more relatives. A Previous breast biopsy was reported in 8.1% of breast cancer cases and 4.4% in the control group.

Table 2 shows the odds ratio (OR) of breast cancer by family history adjusted for age, age at menarche, age at first live birth, number of births, and previous breast biopsies. The OR was 1.25 for one FDR and 4.41 for two or more FDR. The OR was statistically significant for SDR both for one relative (2.39, 95% CI 1.61–3.55) and two or more relatives (3.78, 95% CI 1.62–8.80). For the combinations of FDR+SDR, the highest risk was found for four or more relatives (OR = 6.14, 95% CI 1.75–21.56).

In Table 3 we present the OR for family history stratifying by menopausal status. The risks were similar for pre- and post-menopausal women. The highest risks were noted for the combination of FDR+SDR of 3.83 (95% CI 1.39–10.6) among premenopausal women and 4.37 (95% CI 1.34–14.27) for postmenopausal women.

Table 3. Odds Ratio (OR) and Corresponding 95% Confidence Interval (CI)* of Breast Cancer by Family History (FH) and Menopausal Status

FH	Breast cancer cases	Controls	OR (95% CI)
Premenopausal women			
No FH (FDR)	204	546	1†
≥1 FDRs	29	53	0.99 (0.52–1.89)
No FH (SDR)	188	536	1†
≥1 SDRs	45	63	2.47 (1.51–4.05)
No FH	166	485	1†
1 F-S DR	56	107	1.37 (0.86–2.17)
≥2 F-S DRs	11	7	3.83 (1.39–10.6)
Postmenopausal women			
No FH (FDR)	273	553	1†
≥1 FDRs	52	55	1.53 (0.92–2.55)
No FH (SDR)	277	570	1†
≥1 SDRs	48	38	2.79 (1.62–4.82)
No FH	230	520	1†
1 F-S DR	86	83	2.05 (1.34–3.14)
≥2 F-S DRs	9	5	4.37 (1.34–14.27)

FH, family history; FDR, first-degree relative; SDR, second-degree relative; F-S DR, first-second-degree relative.
*Estimated from unconditional logistic regression adjusted for age, age at menarche, age at first live birth, number of births, number of previous biopsy.
†Reference category.

In Table 4 we show the OR of the original Gail model and those of the Italian model that were calculated for all women in the case–control study. A significantly increased OR was observed in breast cancer patients for a menarche at age 12–13 years and <12 years when compared with menarche at over 14 years of age. The combination between age and number of previous biopsies shows a high prevalence of previous biopsies for benign breast disease among cancer cases (OR = 5.63 and 6.18 under and over 50 years of age, respectively). Only 8.1% of Italian breast cancer women had a history of previous breast biopsies compared with 23.3% among the cases used to construct the Gail model (13). Among breast cancer cases the only significant risk was observed for women with age at first live birth <20 and women with FDR ≥ 1 (OR = 2.31 and OR = 3.25, respectively).

In Table 5 we present the OR for breast cancer considering the variable age at first live birth with both SDR and the combination of FDR+SDR. Evaluating the SDR model, the OR was significantly increased for breast cancer cases. The combination of FDR+SDR shows that the highest risk for breast

Table 4. Comparison of the Original Gail Model Risk Factor Relative Risks for Breast Cancer with those Calculating Using Mediterranean Women

Risk factors		Gail et al. (13)†	Breast cancer patients*	No. breast cancer cases (n = 558)	No. controls (n = 1,207)
AGEMEN					
≥14		1‡	1‡	135	352
12–13		1.099	1.22 (0.95–1.57)	268	606
<12		1.207	1.68 (1.26–2.24)	151	249
NBIOPS					
Age <50 years					
0		1‡	1‡	176	495
(≥)1		1.698	5.63 (2.48–12.75)	18	9
≥2		2.882			
Age ≥50 years					
0		1.000	1†	337	694
(≥)1		1.273	6.18 (2.87–13.28)	27	9
≥2		1.620			
AGEFLB	FDR				
<20	0	1‡	1‡	37	102
	≥1	2.607	0.96 (0.32–2.87)	5	14
	≥2	6.798			
20–24	0	1.244	0.99 (0.65–1.51)	146	405
	≥1	2.681	1.54 (0.80–2.94)	22	38
	≥2	5.775			
25–29 or null.	0	1.548	1.35 (0.89–2.04)	220	462
	≥1	2.756	2.31 (1.27–4.19)	34	40
	≥2	4.907			
≥30	0	1.927	1.5 (0.92–2.42)	70	128
	≥1	2.834	3.25 (1.52–6.95)	20	16
	≥2	4.169			

FH, family history; FDR, first-degree relative; SDR, second-degree relative; F-S DR, first-second-degree relative.
*Odds ratios and 95% confidence intervals included terms for age, age at menarche, number of biopsies, and the combination of age at first live birth and family history.
†Relative risk compared to an individual of the same age without any risk factors is estimated by locating the person's associated relative risk for AGEMEN, NBIOPS, and combination AGEFLB and NUMREL and multiplying these three numbers together.
‡Reference category.

Table 5. Extended Family History Model and Associated Relative Risks

Risk factors		Breast cancer patients*	No. breast cancer cases (n = 558)	No. controls (n = 1,207)
AGEFLB	SDR			
<20	0	1†	32	107
	≥1	4.44 (1.64–12)	10	9
20–24	0	1.16 (0.75–1.81)	141	413
	≥1	3.25 (1.68–6.29)	27	30
25–29 or null.	0	1.69 (1.09–2.6)	217	453
	≥1	2.8 (1.56–5.04)	37	49
≥30	0	1.83 (1.12–3.0)	71	131
	≥1	5.01 (2.21–11.35)	19	13
AGEFLB	FDR+SDR			
<20	0	1†	30	93
	1	1.73 (0.77–3.91)	9	23
	≥2		3	0
20–24	0	1.0 (0.63–1.59)	120	377
	1	1.91 (1.07–3.41)	39	64
	≥2	14.3 (2.91–70.3)	9	2
25–29 or null.	0	1.46 (0.93–2.29)	187	415
	1	2.5 (1.47–4.26)	61	80
	≥2	2.57 (0.79–8.34)	6	7
≥30	0	1.46 (0.86–2.47)	55	118
	1	4.42 (2.24–8.71)	33	23
	≥2	1.88 (0.3–11.9)	2	3

FH, family history; FDR, first-degree relative; SDR, second-degree relative; F-S DR, first-second-degree relative.
*Odds ratios and 95% confidence intervals included terms for age, age at menarche, number of biopsies, and the combination of age at first live birth and family history;
†Reference category.

cancer patients was found when age at first live birth was >30 years with one FDR and/or SDR (OR = 4.42).

The c-statistic was calculated for the three different models. The c-statistic for the first model with FDR was 0.55, (95% CI 0.53–0.58); for SDR the c-statistic indicates a modest but significant discriminatory accuracy (0.56, 95% CI 0.53–0.59); and the concordance statistic for the combination of FDR+SDR was 0.57 (95% CI 0.54–0.60). The area under the curve (equivalent to the c-statistic) for the combined model is modestly greater than for the Gail model. A formal significance-based test (Contrast Test Result) comparing the concordance 0.57 against the concordance 0.56 was statistically significant ($\chi^2 = 6.564$, $p = 0.03$). In order to make a comparison between two or more models derived from the same subjects, the implicit correlation between the curves should be taken into account (18).

DISCUSSION

The identification of preventive interventions for women at high risk for breast cancer is more likely to lead to early diagnosis and treatment (20). To the extent that this model will facilitate research into breast cancer prevention, the Gail model represents the first step toward achieving breast cancer control (21).

The Gail model was initially proposed to assist in counseling women, and there has been a demand for informed counseling of patients at elevated risk (13). The Gail model calculation classifies a woman at the lowest risk level for all factors as RR = 1.0, but relative risks can vary up to 20.0 for those with more than two relatives with breast cancer and with a positive breast biopsy (11). In our study, the result for comparing women who had already developed breast cancer with control group (representing the general population) was that family history appears to affect the estimation of breast cancer risk greatly.

The frequency of previous breast biopsies among controls is lower for Italian women than for US white women. Moreover, the relative difference of previous benign breast biopsies between controls and women with breast cancer in Italy is fourfold (1.5% versus 8.1%) compared with controls and breast cancer cases in the United States (17% versus 23.3%) (13).

According to our findings, the Gail model underestimates risk in women with an extended family history. While the use of the variable number of relative is not consistent with our present understanding of hereditary breast cancer, other models have attempted to calculate genetic risk better. The model developed by Claus et al. (8) takes the number of SDR into consideration. This is consistent with the assumption that affected SDR almost certainly carry a mutation in an incompletely penetrant dominant gene, the parents have 50% chance of carrying the mutation, and the child has a 25% chance; signifying the chance of paternal transmission of an autosomal dominant gene (5,22,23).

The present study has evaluated the concordance of three models to determine whether a more elaborate family history improves discrimination. The addition of SDR and the combination of FDR+SDR reinforced the discriminatory power of the Gail model risk factors. The determination of SDR information is easy to obtain and has no extra cost.

Our results regarding discriminatory accuracy show that the concordance statistic for the model was better than a pure-chance 0.5 but was still relatively low. The fact that the c-statistic for the three models in this data set was similar to that calculated for the Nurses

Health Study (24) suggests that there was no significant bias against the Gail model in our analysis. Because our model was developed using the same data set, our estimates for the c-statistic are likely to be overly optimistic. Of course it is possible that models that incorporate additional predictive variables, such as plasma estrogen levels (25–27), mammographic density (28,29), or more complex information on family history of breast and ovarian cancer (30), may perform somewhat better at individual discrimination.

Recall and selection biases are possible, as in most case–control studies. The questionnaire was administered to cases and controls by the same interviewers under similar conditions in a hospital setting, thus minimizing information bias; in particular the interviewers paid particular attention to obtain information on family history by controls to limit the use of hospital controls as a potential limitation.

Therefore, it appears useful to provide additional information that would allow an approximate statistical comparison of the GM risk factors; but only in addition to, not in place of, the standard model. In fact, the Gail model can be improved with additional information on the women being evaluated. Recent studies report that the inclusion of breast density, and perhaps other modifiable risk factors are favorable in the ongoing evolution of breast cancer prediction (31,32).

The present study contains information that could be of great use to the Italian Health Authorities. In fact we suggest adding extended family history information as an indicator of breast cancer risk for women who participate in screening campaigns. One should take into account not only the number of episodes of breast cancer that occur in the family, but also the number of women in the family who are at risk.

Acknowledgment

The authors gratefully acknowledge the Italian League Against Cancer. Thanks to Mitchell H. Gail, M.D., Ph.D. Chief of Biostatistics Branch Division of Cancer Epidemiology and Genetics National Cancer Institute, for the suggestions and advices.

REFERENCES

1. Parkin DM. International variation. *Oncogene* 2004;23:6329–40.

2. Stewart SL, King JB, Thompson TD, Friedman C, Wingo PA. Cancer mortality surveillance – United States, 1990–2000. *MMWR Surveill Summ* 2004;53:1–108.

3. Jemal A, Thomas A, Murray T, Thun M. Cancer statistics, 2002. *CA Cancer J Clin* 2002;52(1):23–47.

4. Bray F, Sankila R, Ferlay J, Parkin DM. Estimates of cancer incidence and mortality in Europe in 1995. *Eur J Cancer* 2002;38:99–166.

5. Claus EB, Risch N, Thompson WD. Autosomal dominant inheritance of early-onset breast cancer. Implications for risk prediction. *Cancer* 1994;73:643–51.

6. Collaborative Group on Hormonal Factors in Breast Cancer. 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;358:1389–99.

7. McPherson K, Steel CM, Dixon JM. ABC of breast diseases. Breast cancer – epidemiology, risk factors, and genetics. *BMJ* 2000;321:624–8.

8. Claus EB, Risch N, Thompson WD. Age at onset as an indicator of familial risk of breast cancer. *Am J Epidemiol* 1990;131:961–72.

9. Ottman R, Pike MC, King MC, Casagrande JT, Henderson BE. Familial breast cancer in a population based series. *Am J Epidemiol* 1986;123:15–21.

10. Sattin RW, Rubin GL, Webster LA, *et al.* Family history and the risk of breast cancer. *JAMA* 1985;253:1908–13.

11. Lynch HT, Watson P, Conway T, Fitzsimmons SL, Lynch J. Breast cancer family history as a risk factor for early onset breast cancer. *Br Cancer Res Treat* 1988;11:263–7.

12. Colditz GA, Willett WC, Hunter DJ, *et al.* Family history, age, and risk of breast cancer: prospective data from the Nurse's Health Study. *JAMA* 1993;270:338–43.

13. 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;81:1879–89.

14. Montella M, Crispo A, Botti G, *et al.* An assessment of delays in obtaining definitive breast cancer treatment in Southern Italy. *Breast Cancer Res Treat* 2001;66:209–125.

15. Montella M, Schittulli F. Reduction in the number of women with advanced breast cancer stage at diagnosis in Italy. *Eur J Cancer Prev* 2005;14:79–80.

16. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988;44:837–45.

17. Harrell FE Jr., Lee KL, Pollock BG. Regression models in clinical studies: determining relationships between predictors and response. *J Natl Cancer Inst* 1988;80:1198–202.

18. Hanley JA, McNeil BJ. The meaning and use of the area under the receiver operating characteristic (ROC) curve. *Radiology* 1982;143:29–36.

19. SAS Institute, Inc. *SAS/STAT/IML Software. Changes and Enhancements Through Release 8.2*. Cary, NC: SAS Institute, Inc., 2001.

20. Pickle LM, Johnson KA. Estimating the long-term probability of developing breast cancer. *J Natl Cancer Inst* 1989;81:1854–5.

21. Amir E, Evans DG, Shenton A, *et al.* Evaluation of breast cancer risk assessment packages in the family history evaluation and screening programme. *J Med Genet* 2003;40:807–14.

22. McTiernan A, Kuniyuki A, Yasui Y, *et al.* Comparisons of two breast cancer risk estimates in women with a family history of breast cancer. *Cancer Epidemiol Biomarkers Prev* 2001;10:333–833.

23. Bondy ML, Lustbader ED, Halabi S, Ross E, Vogel VG. Validation of a breast cancer risk assessment model in women with a positive family history. *J Natl Cancer Inst* 1994;86:620–5.

24. Rockhill B, Spiegelman D, Byrne C, Hunter DJ, Colditz GA. Validation of the Gail *et al.* model of breast cancer risk prediction and implications for chemoprevention. *J Natl Cancer Inst* 2001;93:358–66.

25. Pike MC, Spicer DV, Dahmoush L, Press MF. Estrogens, progestogens, normal breast cell proliferation, and breast cancer risk. *Epidemiol Rev* 1993;15:17–35.

26. Hankinson SE, Willett WC, Manson JE, *et al.* Plasma sex steroid hormone levels and risk of breast cancer in postmenopausal women. *J Natl Cancer Inst* 1998;90:1292–9.

27. Toniolo PG, Levitz M, Zeleniuch-Jacquotte A, *et al.* A prospective study of endogenous estrogens and breast cancer in postmenopausal women. *J Natl Cancer Inst* 1995;87:190–7.

28. Byrne C. Studying mammographic density: implications for understanding breast cancer. *J Natl Cancer Inst* 1997;89:531–3.

29. Kerlikowske K, Carney PA, Geller B, *et al.* Performance of screening mammography among women with and without a first-degree relative with breast cancer. *Ann Intern Med* 2000;133:855–63.

30. 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;28:115–20.

31. Barlow WE, White E, Ballard-Barbash R, *et al.* Prospective breast cancer risk prediction model for women undergoing screening mammography. *J Natl Cancer Inst* 2006;98:1204–14.

32. Chen J, Pee D, Ayyagari R, *et al.* Projecting absolute invasive breast cancer risk in white women with a model that includes mammographic density. *J Natl Cancer Inst* 2006;98:1215–26.