ORIGINAL ARTICLE

Evaluation of breast cancer risk assessment packages in the family history evaluation and screening programme

E Amir,* D G Evans,* A Shenton,* F Lalloo, A Moran, C Boggis, M Wilson, A Howell

J Med Genet 2003;**40**:807–814

*The first three authors contributed equally to the paper

See end of article for authors' affiliations

Correspondence to: Professor D G Evans, Academic Unit of Medical Genetics and Regional Genetics Service, St. Mary's Hospital, Hathersage Road, Manchester M13 0JH, UK; gareth.evans@cmmc.nhs.uk

Submitted for publication 17 April 2003
Accepted for publication 26 June 2003

Introduction: Accurate individualised breast cancer risk assessment is essential to provide risk–benefit analysis prior to initiating interventions designed to lower breast cancer risk. Several mathematical models for the estimation of individual breast cancer risk have been proposed. However, no single model integrates family history, hormonal factors, and benign breast disease in a comprehensive fashion. A new model by Tyrer and Cuzick has addressed these deficiencies. Therefore, this study has assessed the goodness of fit and discriminatory value of the Tyrer–Cuzick model against established models namely Gail, Claus, and Ford.

Methods: The goodness of fit and discriminatory accuracy of the models was assessed using data from 1933 women attending the Family History Evaluation and Screening Programme, of whom 52 developed cancer. All models were applied to these women over a mean follow up of 5.27 years to estimate risk of breast cancer.

Results: The ratios (95% confidence intervals) of expected to observed numbers of breast cancers were 0.48 (0.37 to 0.64) for Gail, 0.56 (0.43 to 0.75) for Claus, 0.49 (0.37 to 0.65) for Ford, and 0.81 (0.62 to 1.08) for Tyrer–Cuzick. The accuracy of the models for individual cases was evaluated using ROC curves. These showed that the area under the curve was 0.735 for Gail, 0.716 for Claus, 0.737 for Ford, and 0.762 for Tyrer–Cuzick.

Conclusion: The Tyrer–Cuzick model is the most consistently accurate model for prediction of breast cancer. The Gail, Claus, and Ford models all significantly underestimate risk, although the accuracy of the Claus model may be improved by adjustments for other risk factors.

Women who are at increased risk for breast cancer can be identified on the basis of their individual risk factors. However, such an approach does not permit combination of multiple risk factors or calculation of a woman's lifetime probability of breast cancer. Therefore, multivariate risk models have been introduced. These models allow determination of a woman's composite relative risk for breast cancer as well as her cumulative lifetime risk adjusted for all risk factors. Such models therefore, provide an individualised breast cancer risk assessment, which is an essential component of the risk–benefit analysis from which decisions regarding the implementation of frequent surveillance, chemoprevention or prophylactic surgery can be made.

Two frequently used models are the Gail and Claus models. Gail *et al*,[1 2] described a risk assessment model that focuses on non-genetic risk factors with limited information on family history, while Claus *et al*[3] assessed the risk level based on family history of breast cancer. A third model, Ford[4] is based on personal and family history characteristics to identify the presence of any germline mutation of the *BRCA* genes. Both the Ford and Claus models are outputs of the BRCAPro software package.

There is evidence to support the use of these mathematical models.[5] However, until recently, no single model integrated family history, surrogate measures of endogenous oestrogen exposure and benign breast disease in a comprehensive fashion. A new model, Tyrer–Cuzick[6] is based partly on a dataset acquired from the International Breast Intervention Study (IBIS).[7] As can be seen in table 1, the Tyrer–Cuzick model addresses many of the pitfalls of the above models; significantly, the combination of extensive family history, endogenous oestrogen exposure, and benign breast disease (atypical hyperplasia).

The absence of a comprehensive risk assessment model and the choice of other models has led to much debate as to which package is best in the family history setting. Therefore, this study was designed to compare the predictive value of the Gail, Claus, Ford, and Tyrer–Cuzick risk assessment models using a cohort of women attending the Family History Evaluation and Screening Programme. These computerised models were also compared with a risk assessment undertaken by clinicians based on Claus tables with adjustment for other risk factors, particularly hormonal and reproductive factors (the Manual model).

METHODS

Study population

Since 1987, 4536 women have been assessed in the Family History Clinic at the University Hospital of South Manchester for their family history of breast and other cancers, and their hormonal and reproductive factors. These women completed a comprehensive breast cancer risk assessment during which risk factor information was collected, analysed, and archived to a database. Breast examination and mammography were also carried out.

Archived information includes: demographic details, family pedigree including any history of cancer (current age or age of death of any relative, type of cancer, and age at diagnosis), reproductive history (age at menarche, age at first pregnancy, duration of episodes of lactation, and age at menopause if applicable), history of benign breast disease (including number of benign biopsies), artificial oestrogen

Abbreviations: ROC, receiver operating characteristic

Table 1 Variables used in the Gail, Claus, Ford, Tyrer–Cuzick, and Manual models

Variable	Gail	Claus	Ford	Tyrer–Cuzick	Manual
Personal information					
Age	Yes	Yes	Yes	Yes	Yes
Body mass index	No	No	No	Yes	No
Hormonal factors					
Menarche	Yes	No	No	Yes	Yes
First live birth	Yes	No	No	Yes	Yes
Menopause	No	No	No	Yes	Yes
Personal breast disease					
Breast biopsies	Yes	No	No	Yes	Yes
Atypical hyperplasia	Yes	No	No	Yes	Yes
LCIS	No	No	No	Yes	Yes
Family history					
First degree relatives	Yes	Yes	Yes	Yes	Yes
Second degree relatives	No	Yes	Yes	Yes	Yes
Age of onset of cancer	No	Yes	Yes	Yes	Yes
Bilateral breast cancer	No	No	Yes	Yes	Yes
Ovarian cancer	No	No	Yes	Yes	Yes
Male breast cancer	No	No	Yes	No	Yes

exposure (duration of oral contraceptive pill usage, hormone replacement, or fertility drugs), and morphometric information (height and weight). Ethnic origin has been collected since 1995. In addition, the database stores an absolute lifetime risk calculated using the Manual model.

The database also contains information regarding breast cancer incidence until at least January 2001. All women previously assessed for whom vital and tumour status was not available as of January 2001 and whose address showed residency within Greater Manchester or Lancashire were traced for such information on the North West Cancer Registry. The cancer registry data in combination with tumour and vital status from the database were used as the observed numbers of breast cancers for the purpose of comparison.

For analysis, 3170 women had all elements of hormonal, reproductive, and computerised pedigree available, and 20 of these had cancer at initial assessment. Of these, 1933 women were followed up in a regular 12–18 monthly mammography screening programme and 1217 women were discharged to routine follow up by their general practitioner and the National Breast Cancer Screening programme if aged over 50 years (the NHS programme offers 3 yearly screening from 50 years of age). There was a missing element to the dataset for 1366 women, which did not allow analysis in every risk package. The most frequent reason was that women were sent from the Regional Genetics Department with a photocopied pedigree.

Study tools

The computerised risk assessment packages Gail, BRCAPro (Claus and Ford) and Tyrer–Cuzick were tested on this population. Lifetime and 10 year risk for developing breast cancer were estimated using each of the models. Methods for risk estimation differed between the models (table 1). The Gail risk was established by importing the relevant data (current age, age at menarche, number of previous breast biopsies, age at first live birth, and the number of first degree relatives affected with breast cancer) into an external application. Claus and Ford risks were calculated using a supplementary plug-in for the Cyrillic 3 package, a software package designed to display family pedigrees for use in clinical genetics and genetic counselling. The required data was imported into the Cyrillic package by using an intermediate GEDCOM file (GEnealogical Data COMmunication). Example patients were assessed from the direct BRCAPro model and near identical risks to those obtained from the Cyrillic package were obtained. The Tyrer–Cuzick model was also tested by importing the relevant details into an external application. These computerised models were then compared with a manual risk calculation (the Manual model), which was calculated in clinic. The Manual model uses the Claus tables[3] and curves[8 9] to prospectively calculate a heterozygote and lifetime risk (which includes a population risk element). The latter of these can then used to compute 10 year risks. Families fulfilling Breast Cancer Linkage Consortium (BCLC) criteria were given risks based on the penetrance and 10 year risks from the BCLC.[4 10] All other women were given 10 year risks calculated from the equivalent figure given in the Claus tables after a clinician's modification for hormonal and reproductive factors. This adjustment allowed for a maximum 50% upwards or downwards change in risk based on a combination of hormonal and reproductive factors. In practice, this meant a greater reduction than increase in the higher risk categories. As it is not possible to increase lifetime risk of 80–85% above 100%, only a maximum 25% upwards variation of the heterozygote risk is possible.[8] For example, if a woman has a 1 in 5 heterozygote risk, her hereditary element is 1/5 of 80%, that is, 16%. This can only rise to a maximum of 20%. However her population element of 6–8% (4/5ths of 8–10% population risk) can rise to 9–12%. Therefore her risk calculated from family history can vary upwards from 22–24% to 29–32%, but may drop by 50% to 11–12%, that is, a population risk of 3–4% and a hereditary risk of 8%. A more detailed description of the Manual model is described elsewhere.[8 9]

The outputs of all these models were used as the expected numbers of breast cancers.

Inclusion criteria

The sample was limited to the women for whom breast cancer risk estimation could be derived by all models. Only breast cancer incidence from time of initial assessment to January 2001 was assessed for the entire cohort. Data analysis was carried out on both the full population that had ever visited the Family History Clinic and again among those women still enrolled in the 12–18 monthly mammographic screening programme.

Statistical analysis

Predicted risk during the follow up period was compared against the observed numbers of breast cancers (database and cancer registry data). In order to express the predicted risk in terms of the follow up period, projections of the absolute 10 year risk were obtained from the models. These were then converted first into an annual risk (division by 10), and then into a follow up risk by multiplication by the number of years of follow up. Follow up was taken as the period of time from initial assessment with clinical breast examination and mammogram to the most recent examination or 31 January 2001, whichever was the later. All prevalence cancers were excluded. The expected number (E) of breast cancers within the cohort was calculated as the sum of these predicted risks. E was then compared with the observed number (O) of women with breast cancer, and 95% confidence intervals (CI) of the ratio of expected to observed numbers (E/O) were obtained with the use of the Poisson variance. This was carried out by first solving the 95% CI for O, specifically the values of O_L for the lower limit and O_U for the upper limit, and then dividing E by the values of O_L and O_U to obtain the upper and lower CI for this ratio.

These analyses were performed in two parts: (*a*) part 1, comparison based on the whole study population; and (*b*)

part 2, comparison based on subgroups of the study population, using the following risk factor categories.

Family history:

- One first degree relative (FDR) only;
- Two FDRs only (no other breast cancer family history);
- One FDR plus two or more breast cancers in family (can include another FDR);
- Ovarian cancer in any relative (any family containing an ovarian cancer excluded from other categories);
- Any other family history combination.

Menarche:

- At or before 12 years of age;
- After 12 years of age.

First live birth:

- At or before 30 years of age;
- After 30 years of age and nulliparity.

In addition to the overall accuracy in terms of the numbers of cancers predicted, the accuracy for individual cases was also evaluated, that is, the ability of the models to separate individuals who will go on to develop disease from those who will be disease free.[11] To do this, the probability of developing cancer for the period at risk was calculated for each woman using the different models, and receiver operating characteristic (ROC) curves were generated. These are plots of the true positive rate against the false positive rate for different possible cut off points. They show the trade off between sensitivity and specificity of the different models (any increase in sensitivity will be accompanied by a decrease in specificity). In order to distinguish formally high risk from low risk subjects, the C statistic was calculated. For binary outcomes, Hanley and McNeil[12] quantified the concordance statistic (C statistic) as equal to the area under the ROC curve. Therefore, the Wilcoxon estimate of the area under the ROC curve was calculated for all models. The larger the area under the curve, the more accurate the test, with an area under the curve of 1 representing a perfect test, whereas an area of 0.5 or lower represents a poor test.

RESULTS

Analysis of the whole study population was carried out on data from 3170 women (range 21–73 years; median 44 years) (tables 2–5). Twenty prevalent cancers (table 6) were excluded (rate 20/3170 = 6.31 per thousand), leaving 3150 women for follow up. The mean time of follow up was 5.27 years (range 0.10–15.00 years). Fifty five per cent of the population had a follow up of more than 5 years, while only 9.7% had less than 1 year of follow up. During the course of follow up, 64 cancers were diagnosed, giving an incidence of 3.86 cancers per 1000 women per year. Ethnic origin was available on 2398/3150 (76%) women, of whom 95.7% were white Northern European, 2.6% Jewish, and 1.7% other (including Afro-Caribbean and Asian).

Analysis of the screening programme study population was carried out on data from 1933 women (range 25–73 years; median 46 years). The mean time of follow up was 6.39 years (range 0.28–15.00 years). Seventy per cent of the population had a follow up of more than 5 years while only 1.3% had less than 1 year of follow up. During the course of follow up, 52 cancers were diagnosed (including interval cancers), giving an incidence of 4.21 cancers per 1000 women per year.

Table 2 Breakdown of ages of women in the total study population and the screening programme by age at first assessment and breast cancer diagnosis

	Total study population			Screening population		
Age groups	Disease free	Breast cancer	Total	Disease free	Breast cancer	Total
10–19	29	0	29	1	0	1
20–29	484	1	485	162	1	163
30–39	1294	22	1316	866	20	886
40–49	989	30	1019	685	24	709
50–59	275	10	285	162	7	169
60–69	15	1	16	5	0	5
Total	3086	64	3150	1881	52	1933

Table 3 Breakdown of family history of women in the total study population and the screening programme

Family history	Total number	Number in screening group	Mean age at assessment	Breast cancers
a. 1 FDR BC<40 only	510	302	35.49	9
b. 1 FDR BC 40+ only	728	401	38.58	16
c. 2 FDR BC<50 only	82	61	43.34	1
d. 2 FDR BC only not fulfilling c	127	81	45.72	3
e. 2 BC affected relatives not in c or d	755	485	37.74	7
f. 3 BC affected relatives	451	285	39.23	9
g. 4+ BC affected relatives	272	181	39.52	9
h. 1 OC+BC	154	98	39.67	4
i. 2 OC+BC	42	30	39.61	5
j. Any other history	29	9	39.07	1
Total	3150	1933		64

BC, breast cancer; OC, ovarian cancer; FDR, first degree relative. Categories are more extensive in this table than for the full analysis and we have excluded breast cancers occurring in distant relatives in categories a–d.

Table 4 Comparison of expected cases of breast cancer predicted from Gail, Claus, Ford (BRCAPRO), Tyrer–Cuzick, and the Manual models against the observed cases for the total study population (n = 3150)

	Observed (O)	Expected (E)	E/O	95% CI
Gail	64	44.3037	0.69	0.54 to 0.90
Claus	64	48.5565	0.76	0.59 to 0.99
Ford	64	42.2790	0.66	0.52 to 0.86
Tyrer–Cuzick	64	69.5653	1.09	0.85 to 1.41
Manual	64	77.9232	1.22	0.95 to 1.58

CI, confidence interval.

Table 5 Comparison of expected cases of breast cancer predicted from Gail, Claus, Ford (BRCAPro), Tyrer–Cuzick, and the Manual models against the observed cases for women still enrolled in the screening programme (n = 1933)

	Observed (O)	Expected (E)	E/O	95% confidence intervals
Gail	52	25.0312	0.48	0.37 to 0.64
Claus	52	29.1489	0.56	0.43 to 0.75
Ford	52	25.4029	0.49	0.37 to 0.65
Tyrer–Cuzick	52	42.0492	0.81	0.62 to 1.08
Manual	52	46.4261	0.89	0.68 to 1.20

Table 6 Comparison of expected and observed cancers for categories defined by breast cancer risk factors for the full study population

Variable	N	O	Prevalent cancers	Gail			Claus			Ford			Tyrer–Cuzick			Manual		
				E	E/O	95% CI	E	E/O	95% CI	E	E/O	95% CI	E	E/O	95% CI	E	E/O	95% CI
Family history																		
1 FDR	1224	25	4	13.69	0.55	0.37 to 0.85	10.93	0.44	0.30 to 0.68	10.29	0.41	0.28 to 0.64	18.78	0.75	0.51 to 1.16	22.56	0.90	0.61 to 1.39
2 FDRs	204	4	2	6.56	1.64	0.64 to 6.02	4.08	1.02	0.40 to 3.74	3.16	0.79	0.31 to 2.90	6.28	1.57	0.61 to 5.76	5.41	1.35	0.53 to 4.96
1 FDR+2 other relatives	555	14	6	11.02	0.79	0.47 to 1.44	16.73	1.20	0.71 to 2.19	12.36	0.88	0.53 to 1.61	17.90	1.28	0.76 to 2.34	19.94	1.42	0.85 to 2.61
OC	196	9	2	2.75	0.31	0.16 to 0.67	3.40	0.38	0.20 to 0.83	5.53	0.61	0.32 to 1.34	6.83	0.76	0.40 to 1.66	6.07	0.67	0.36 to 1.47
Other history	971	12	6	10.28	0.86	0.49 to 1.66	13.41	0.76	0.64 to 2.16	10.95	0.91	0.52 to 1.77	19.76	1.65	0.94 to 3.19	23.95	2.00	1.14 to 3.86
Menarche																		
⩽12 years	1391	26	5	19.49	0.75	0.51 to 1.15	19.79	0.76	0.52 to 1.17	17.30	0.67	0.45 to 1.02	29.43	1.13	0.77 to 1.73	33.90	1.30	0.89 to 2.00
>12 years	1759	38	15	24.81	0.65	0.48 to 0.92	28.76	0.76	0.55 to 1.07	24.98	0.66	0.48 to 0.93	40.13	1.06	0.77 to 1.49	44.02	1.16	0.84 to 1.64
First live birth																		
⩽30 years	2026	38	13	32.31	0.85	0.62 to 1.20	33.81	0.89	0.65 to 1.26	29.34	0.77	0.56 to 1.09	43.92	1.16	0.84 to 1.63	48.35	1.27	0.93 to 1.80
>30 years or nulliparous	1124	26	7	12.00	0.46	0.31 to 0.71	14.75	0.57	0.39 to 0.87	12.94	0.50	0.34 to 0.76	25.65	0.99	0.67 to 1.61	29.58	1.14	0.78 to 1.74

OC, ovarian cancer; FDR, first degree relative.

Table 2 shows the breakdown of ages of women in the total study population and the screening programme by age at initial assessment and breast cancer diagnosis. The age range of diagnosed cancers was 32–64 years in the overall population and 32–61 years in the screening group, although one cancer was detected in a 32 year old woman who commenced screening at 29 years. The degree of family history is presented in table 3.

All 12 cancers in the 1217 women not actively followed up were identified from the Cancer Registry more than 2 years after initial evaluation, and only four were detected by mammographic screening in the NHS Breast Screening programme.

Risk assessment using all five models was not possible in 1366 women on the database because of missing values for one or more of the criteria used by the models. Analysis of this group produced an identical expected versus observed ratio using the manual model (1.22 for both the non-included (29 cancers observed) and the included population).

The expected against observed counts for all breast cancers are shown in tables 4 and 5.

With the Gail model, 44.30 breast cancers were predicted from the total study population compared with 64 observed. This is an expected to observed ratio (E/O) of 0.69 (95% CI 0.54 to 0.90). Among the screening programme population, 25.03 cancers were predicted compared with 52 observed (E/O = 0.48, 95% CI 0.37 to 0.64). Therefore, the Gail model significantly underestimated the absolute risk of women.

With the Claus computer model, 48.56 breast cancers were expected from the total study population compared with the 64 observed cancers. This corresponds to an E/O ratio of 0.76 (95% CI 0.59 to 0.99). Among the screening programme population 29.15 cancers were expected compared with 52 observed (E/O = 0.56, 95% CI 0.43 to 0.75). Therefore, the Claus model also significantly underestimated the absolute risk of women.

Further analysis was carried out to calculate expected and observed rates for categories defined by breast cancer risk factors (tables 6 and 7). The results showed that the Gail and Claus models significantly underestimate risk across the entire risk factor categories. Importantly, they markedly underestimated risk among women in the screening programme with a single first degree relative with breast cancer (Gail E/O = 0.39, 95% CI 0.25 to 0.65; Claus E/O = 0.32, 95% CI 0.20 to 0.54). Furthermore, the Gail model appeared to show an increase in risk of breast cancer with early age at first pregnancy in the familial setting. Indeed, example cases showed that the model gives a substantially increased risk in the category of "patients with two affected first degree relatives" with decreasing age at first pregnancy and shows very little protective effect in the "patients with one affected first degree relative" category.

The Ford model performed worst, with only 42.28 breast cancers expected compared with 64 observed. This corresponds to an E/O ratio of 0.66 (95% CI 0.52 to 0.86). Among the screening programme population, 25.40 cancers were expected compared with 52 observed (E/O = 0.49, 95% CI 0.37 to 0.65). Therefore, the Ford model also significantly underestimated the absolute risk of women.

The Tyrer–Cuzick model performed best, with 69.57 breast cancers expected compared with 64 observed. This corresponds to an E/O ratio of 1.09 (95% CI 0.85 to 1.41). Among the screening programme population, 42.05 cancers were expected compared with 52 observed (E/O 0.81, 95% CI 0.62 to 1.08). Therefore, the Tyrer–Cuzick model did not show any significant deviation from unity for either of the study populations.

On further analysis (table 6 and 7), this pattern of accurate prediction was continued with most risk factor categories

having E/O ratios, which did not statistically differ from 1.0. The exception to this was once again the women in the screening programme with a single first degree relative with breast cancer. In this sub-group, the Tyrer–Cuzick model also underestimated risk, however, its E/O ratio of 0.56 (95% CI 0.35 to 0.94) was still best of the computerised models.

Finally, the Manual model also performed well, with 77.92 breast cancers expected compared with 64 observed. This corresponds to an E/O ratio of 1.22 (95% CI 0.95 to 1.58). Among the screening programme population, 46.43 cancers were expected compared with 52 observed (E/O = 0.83, 95% CI 0.68 to 1.20). Therefore, the Manual model also did not show significant statistical deviation from 1.0.

Further analysis of the Manual model (table 6 and 7) showed results similar to those of the Tyrer–Cuzick model, with most risk factor categories having E/O ratios, which did not exhibit a statistically significant deviation from unity. Significantly, the Manual model was the only model to predict risk accurately in women in the screening programme with a single first degree relative with breast cancer, achieving an E/O ratio of 0.67 (95% CI 0.43 to 1.14).

As there did not appear to be a significant effect of menarche on breast cancer risk using the age of 12 years to define the two subgroups, analysis was repeated with more extreme ages for menarche. Comparison of risk for women with menarche before 12 years with after 13 years was therefore undertaken. The E/O ratio in the Claus and Ford models (those not incorporating hormonal factors) was 0.64:0.72 and 0.56:0.64 respectively, implying a 12–14% increase in incidence in the earlier onset cohort.

Table 8 shows the median and range of estimated 10 year risks for women who remained disease free and for women who developed breast cancer using each of the models. With all models, the median of the estimated 10 year risks was significantly greater among women who developed breast cancer than in women who remained disease free.

Analysis of the accuracy of the models for individual cases (fig 1, table 9) showed that of the models tested, the Tyrer–Cuzick model performed best, with an area under the curve of 0.762 (95% CI 0.700 to 0.824). All the other models also showed an area under the curve significantly greater than 0.5.

DISCUSSION

Accurate individualised breast cancer risk assessment is essential for the provision of risk-benefit analysis prior to the initiation of any preventative interventions. Until recently, there have been no stand alone comprehensive breast cancer risk assessment models able to account for all significant permutations of family history, reproductive history, and personal history of benign breast disease. The Tyrer–Cuzick model is a new risk assessment model that appears to address these deficiencies. However, it has not been independently validated and its value in the family history setting is still unclear. Furthermore, there is no clear evidence that even the established risk assessment models (Gail, Claus, and Ford) have ever been fully tested among women attending family history clinics. Indeed the only such attempts we could identify[13–15] simply compared the risks obtained from using Gail and Claus,[14] Gail, Claus, and BRCAPro,[13] and Claus and a UK model, Houlston–Murday,[15] in women from a breast cancer risk assessment clinic. As the risk was only increased above the Gail estimate in 13% of 213 women, the authors of one report[13] concluded that the Gail model was accurate in this setting. The second report[14] also reported a higher risk in most of 491 women using the Gail model rather than Claus and concluded that both models should be used in the family history setting. The present study would suggest that none of these three established

Table 7 Comparison of expected and observed cancers for categories defined by breast cancer risk factors for the women enrolled in the screening programme

			Gail			Claus			Ford			Tyrer–Cuzick			Manual		
Variable	N	O	E	E/O	95% CI	E	E/O	95% CI	E	E/O	95% CI	E	E/O	95% CI	E	E/O	95% CI
Family history																	
1 FDR	686	18	6.98	0.39	0.25 to 0.65	5.79	0.32	0.20 to 0.54	5.36	0.30	0.19 to 0.50	10.05	0.56	0.35 to 0.94	12.14	0.67	0.43 to 1.14
2 FDRs	137	3	3.56	1.19	0.41 to 5.75	2.45	0.82	0.28 to 3.96	1.88	0.63	0.21 to 3.04	3.51	1.17	0.40 to 5.67	3.45	1.15	0.39 to 5.58
1 FDR+2 other relatives	405	13	6.87	0.53	0.31 to 0.99	10.51	0.81	0.47 to 1.52	7.82	0.60	0.35 to 1.13	11.41	0.88	0.51 to 1.65	12.74	0.98	0.57 to 1.84
OC	128	8	1.64	0.21	0.10 to 0.47	2.27	0.28	0.14 to 0.66	3.74	0.47	0.24 to 1.08	4.83	0.60	0.31 to 1.40	3.88	0.49	0.25 to 1.12
Other history	577	10	5.98	0.60	0.33 to 1.25	8.13	0.81	0.44 to 1.70	6.60	0.66	0.36 to 1.38	12.25	1.23	0.67 to 2.55	14.20	1.42	0.77 to 2.96
Menarche																	
⩽12 years	840	21	11.12	0.53	0.35 to 0.86	11.82	0.56	0.37 to 0.91	10.34	0.49	0.32 to 0.80	17.49	0.83	0.54 to 1.35	20.12	0.96	0.63 to 1.55
>12 years	1093	31	13.91	0.45	0.32 to 0.66	17.33	0.56	0.39 to 0.82	15.06	0.49	0.34 to 0.71	24.56	0.79	0.56 to 1.17	26.31	0.85	0.60 to 1.25
First live birth																	
⩽30 years	1292	28	18.00	0.64	0.44 to 0.97	19.95	0.71	0.49 to 1.07	17.16	0.61	0.42 to 0.92	25.87	0.92	0.64 to 1.39	29.28	1.05	0.72 to 1.57
>30 years or nulliparous	641	24	7.03	0.29	0.20 to 0.46	9.20	0.38	0.26 to 0.60	8.24	0.34	0.23 to 0.54	16.17	0.67	0.45 to 1.05	17.14	0.71	0.48 to 1.11

FDR, first degree relative; OC, ovarian cancer.

models provides an accurate risk estimate for most women at high and moderate risk of breast cancer. The third study[15] compared the Claus with the Houlston–Murday model in 200 women from breast cancer genetic risk assessment clinics. Only 27% were ascribed ''high risk'' using the Claus model compared with 53% using Houlston–Murday. While the Houlston–Murday model could be accurate in our dataset this was not available to us as a computer downloaded model. It also does not have an adjustment for other risk factors.

The present study has evaluated both the goodness of fit (as reflected in E/O ratios) and the discriminatory value of five methods for projecting invasive breast cancer (Gail, Claus, Ford, Tyrer–Cuzick, and the Manual model) using a cohort of women from the Family History Evaluation and Screening Programme.

It was found that only the Tyrer–Cuzick and Manual models predicted risk accurately. Of these, the Tyrer–Cuzick model showed the better overall agreement between expected and observed counts of breast cancer among the total study population (69.57 versus 64), while the Manual model was stronger among the screening population (46.43 versus 52). Considering all of the comparisons by categories of risk factors (tables 6 and 7), few E/O ratios deviated significantly from unity with the Tyrer–Cuzick model. The Gail, Claus, and Ford models all significantly underestimated risk, although some comparisons by risk factor categories did not reach a statistically significant deviation from unity. Therefore, the previous work comparing the latter three models favourably[13 14 15] is likely to be misleading, because while they give similar risk values in a risk assessment setting, all three underestimate risk.

In this study, the Gail, Claus, and Ford models all particularly underestimated risk in women with a single first degree relative affected with breast cancer. Tyrer–Cuzick and the Manual model were both accurate in this sub-group. Conversely, all the models accurately predicted risk in women with multiple relatives affected with breast cancer (two first degree relatives, or one first degree plus two other relatives). This implies that the effect of a single affected first degree relative is higher than may have been previously thought. The Gail model is likely to have underestimated in this group, as it does not take into account age of breast cancer and nearly all women in our single FDR category had a relative diagnosed at less than 50 years of age.

The Ford, Tyrer–Cuzick, and Manual models were the only models to predict risk accurately in women with a family history of ovarian cancer. As these are the only models to take account of ovarian cancer in their risk assessment algorithm, this confirms that ovarian cancer has a significant effect on breast cancer risk.

Although the Claus computer model has performed poorly, we do not believe that the model's algorithm contributed to its deficiency, as it is the same model we used in the manual calculation. It appears that that the BRCAPro models do not add in a non-*BRCA1/2* element of risk. This accords with our own direct comparison with the Claus model (data not shown) and would explain why it does not perform well in the lower risk categories such as Group 1. Another possibility is that downward adjustments to heterozygote risk made due to the presence of unaffected female relatives are not counterbalanced by increases in risk in those without such unaffected relatives.

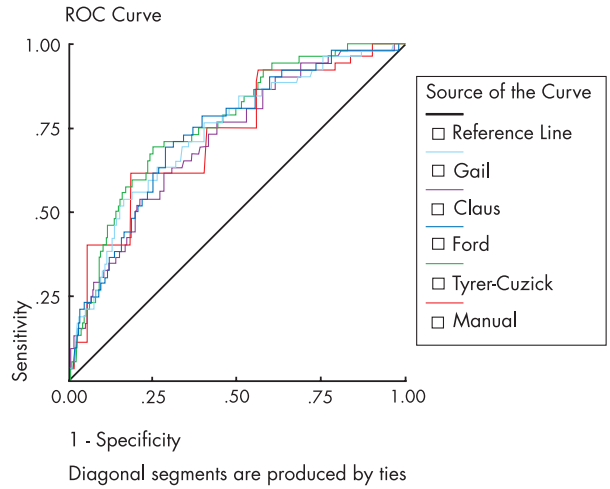

Figure 1 ROC curves for Gail, Claus, Ford (BRCAPro), Tyrer-Cuzick and the Manual models.

The Gail, Claus, and Ford models all significantly underestimated risk in women who were nulliparous or whose first live birth occurred after the age of 30 years. Moreover, the Gail model appeared to show increased risk with pregnancy <30 years in the familial setting. It is not clear why such a modification to the effects of age at first birth should be made unless it is as a result of modifications made to the model after early results suggested an increase with *BRCA1/2* mutation carriers.[16] However, the Gail model has determined an apparent increase in risk with early first pregnancy and it would appear to be misplaced from our results. Furthermore, Gail, Claus, and Ford models also underestimated risk in women whose menarche occurred after the age of 12 years. The Tyrer–Cuzick and Manual models accurately predicted risk in these sub-groups. These results suggest that age at first live birth also has an important effect on breast cancer risk, while age at menarche perhaps has a lesser effect. The effect of pregnancy <30 years appeared to reduce risk by 40–50% compared with the older and nulliparous groups, whereas at the extremes of menarche there was only a 12–14% effect.

The accuracy of these models for individual cases has significant implications in their use in clinical counselling.

Table 8 Median and range of estimated 10 year risks for women who remained disease free and for women who developed breast cancer

	Breast cancer (%)		Disease free (%)		
Model	Range	Median	Range	Median	p Value
Gail	0.2–12.0	3.40	0.0–26.0	1.50	<0.0001
Claus	0.3–15.4	3.25	0.2–35.2	1.60	<0.0001
Ford	0.2–14.0	2.65	0.2–19.2	1.40	<0.0001
Tyrer–Cuzick	1.1–19.2	5.60	0.2–52.3	2.54	<0.0001
Manual	1.2–15.0	6.00	1.2–20.0	3.40	<0.0001

Table 9 Area under the ROC curve values and confidence intervals for the Gail, Claus, Ford (BRCAPro), Tyrer–Cuzick, and the Manual models

Risk assessment model	Area	Asymptotic 95% confidence interval	
		Lower bound	Upper bound
Gail	0.735	0.666	0.803
Claus	0.716	0.648	0.784
Ford	0.737	0.671	0.803
Tyrer–Cusick	0.762	0.700	0.824
Manual	0.727	0.656	0.798

The analysis of the value of the models in predicting risk for individual cases showed that all models were compromised by the trade off between sensitivity and specificity, that is, any increase in sensitivity was accompanied by a decrease in specificity and vice versa. However, of the models, the Tyrer–Cuzick model was again the most accurate, with an area under the ROC curve of 0.762 (95% CI 0.700 to 0.824). All of the other models also showed a significant deviation from 0.5.

The use of the Tyrer–Cuzick model in clinical counselling would lead to a large majority of women who are going to develop breast cancer being offered the option of chemoprevention or prophylactic surgery. However, a large number of women who will remain disease free may also be advised to take such action. The effects of this on disease burden appear to be beneficial. However, as was discussed by Gail *et al*,[17] the potential public health consequences of widespread chemoprevention and prophylactic surgery may be detrimental.

Testing the various models in the family history setting provided very good access to a well maintained cohort sample. However, the use of this sample did have some methodological weaknesses that should be addressed. First, the sample could only be tested over a maximum follow up period of 15.00 years (mean 5.27 years). A longer follow up period might yield better and/or more significant results. Secondly, there was a lack of data on atypical hyperplasia status for women who reported a history of breast biopsy. In this small subset of the cohort (140 women), risk estimation would have been artificially low in the Gail, Tyrer–Cuzick, and Manual models. However, it is likely that women identified on biopsy in other units would have been followed up if they had a biopsy showing proliferative disease and such information was sought in clinic. Indeed, only one cancer occurred in this group (0.7%), which is less than the 2% (63/3010) for the remainder of the full study population. The cohort used in this study comprised a moderate and high risk population with cancer incidence rates of 3–10 times higher than the general population. Hence, our findings may not be representative for women of lower risk. There is unlikely to be any bias in the study relating to the women excluded from the main study, as these women had the same E/O ratio as the study group.

Only nine cancers occurred in the 1216 women in the whole study group for whom we do not have information on screening. This may be an underestimate as patients may have moved area and cancers may not have been registered. This group was predominantly in the lower risk categories and many were not recommended for ongoing screening. The effect of the prevalence scan in taking out the lead time of 1.5 to 3 years (this could potentially have halved the number compared with what may have been found if the 1216 had all been screened for the 5 year mean follow up) may also make the whole study group a less definitive population than the screened one. Although screening is likely to increase the incidence of breast cancer through early detection, most of this effect will be offset by excluding the prevalence cases. Moreover, the Gail and Tyrer–Cuzick models have been derived from a screened population. Given that screening is likely to be recommended in women at moderate and high risk we believe that our Manual model and the Tyrer–Cuzick models are the most appropriate in the family history setting. As the Manual model contains the human approximation, which is difficult to fully elaborate in this paper (it is taught over 5 days on our risk assessment course), the Tyrer–Cuzick model provides the most useful and accurate tool for clinicians involved in breast cancer risk assessment. Nonetheless, geneticists should be able to calculate a manual risk even if they use a computer model, as a slip in data input can produce very inaccurate predictions, and comparison with a manual assessment would improve confidence in the risk figure. Although previous studies have validated the Gail model in a population unselected for family history,[18] the results are not really relevant to a family history/cancer genetics clinic. The most recent study validating in the context of the Nurses Health study[18] produced an E/O ratio of 0.94 (0.89–0.99) compared with our screening population results, which were much worse at 0.48 (0.37–0.64). The concordance statistic (C statistic) was 0.56 (95% CI 0.56 to 0.60), whereas in our study it was better at 0.74 (95% CI 0.67 to 0.80). The study concluded that the Gail model fits well in this sample in terms of predicting numbers of breast cancer cases in specific risk factor strata but had modest discriminatory accuracy at the individual level. Our findings in the familial setting show that Gail is better at discriminatory value than at predicting risk. The difference is most probably explained by the fact that our population was high risk and not in a general population setting (Nurses' Health Study).

In conclusion, the Tyrer–Cuzick model appears to provide the most consistently accurate risk estimation for women at high risk based on family history and hormonal factors. Gail, Claus, and Ford all appear to have limited predictive validity in this cohort.

Authors' affiliations

Eitan Amir, University of Manchester, UK
D G Evans, A Shenton, F Lalloo, Academic Unit of Medical Genetics and Regional Genetics Service, St. Mary's Hospital, Manchester, UK
A Moran, Centre for Cancer Epidemiology, University of Manchester, UK
C Boggis, M Wilson, Department of Radiology, South Manchester University Hospital Trust, Manchester, UK
A Howell, Department of Medical Oncology, Christie Hospital, Manchester, UK

REFERENCES

1 **Gail MH**, Brinton LA, Byar DP, Corle DK, Green SB, Schairer C, Milvihill JJ. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. *J Natl Cancer Inst* 1989;**81**:1879–86.
2 **Costantino JP**, Gail MH, Pee D, Anderson S, Redmond CK, Benichou J, Wieand HS. Validation studies for models projecting the risk of invasive and total breast cancer incidence. *J Natl Cancer Inst* 1999;**91**:1541–48.
3 **Claus EB**, Risch N, Thompson WD. Autosomal dominant inheritance of early onset breast cancer: implications for risk prediction. *Cancer* 1994;**73**:643–51.
4 **Ford D**, Easton DF, Bishop DT, Narod SA, Goldgar DE, the Breast Cancer Linkage Consortium. Risk of cancer in BRCA-1 mutation carriers. *Lancet* 1994;**343**:692–5.
5 **Euhus DM**. Understanding mathematical models for breast cancer risk assessment and counselling. *Breast J* 2001;**7**:224–32.
6 **Tyrer JP**, Duffy SW, Cuzick J. A breast cancer prediction model incorporating familial and personal risk factors. *Stat Med* (in press).
7 **IBIS investigators**. First results from the International Breast Cancer Intervention Study (IBIS-I): a randomised prevention trial. *Lancet* 2002;**360**:817–24.
8 **Evans DGR**, Fentiman IS, McPherson K, Asbury D, Ponder BAJ, Howell A. Familial breast cancer. *BMJ* 1994;**308**:183–7.
9 **Evans DGR**, Lalloo F. Risk assessment and management of hereditary breast cancer. *J Med Genet* 2002;**39**:865–71.
10 **Ford D**, Easton DF, Stratton M, Narod S, Goldgar D, Devilee P, Bishop DT, Weber B, Lenoir G, Chang-Claude J, Sobol H, Teare MD, Struewing J, Arason A, Scherneck S, Peto J, Rebbeck TR, Tonin P, Neuhausen S, Barkardottir R, Eyfjord J, Lynch H, Ponder BA, Gayther SA, Zelada-Hedman M, the Breast Cancer Linkage Consortium. Genetic heterogeneity and penetrance analysis of the BRCA1 and BRCA2 genes in breast cancer families. *Am J Hum Genet* 1998;**62**:676–89.
11 **Harrell FE**, Lee KL, Pollock BG. Regression models in clinical studies; determining relationships between predictors and response. *J Natl Cancer Inst* 1988;**80**:1198–202.
12 **Hanley JA**, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982;**143**:29–36.
13 **Euhus DM**, Leitch AM, Huth JF, Peters GN. Limitations of the Gail model in the specialized breast cancer risk assessment clinic. *Breast* 2002;**8**:23–7.
14 **McTiernan A**, Kuniyuki A, Yasui Y, Bowen D, Burke W, Culver JB, Anderson R, Durfy S. Comparisons of two breast cancer risk estimates in women with a family history of breast cancer. *Cancer Epidemiol Biomarkers Prev* 2001;**10**:333–8.
15 **Tischkowitz M**, Wheeler D, France E, Chapman C, Lucassen A, Sampson J, Harper P, Krawczak M, Gray J. A comparison of methods currently used in

clinical practice to estimate familial breast cancer risks. *Ann Oncol* 2000;**11**:451–4.
16 **Jernstrom H**, Lerman C, Ghadirian P, Lynch HT, Weber B, Garber J, Daly M, Olopade OI, Foulkes WD, Warner E, Brunet JS, Narod SA. Pregnancy and risk of early breast cancer in carriers of BRCA1 and BRCA2. *Lancet* 1999;**354**:1846–50.
17 **Gail MH**, Costantino JP, Bryant J, Croyle R, Freedman L, Helzlsouer K, Vogel V. Weighing the risks and benefits of tamoxifen treatment for preventing breast cancer. *J Natl Cancer Inst* 1999;**91**:1829–46.
18 **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.

ECHO

Prospective Belgian study of neurodegenerative and vascular dementia: APOE genotype effects

S Engelborghs, B Dermaut, J Goeman, J Saerens, P Mariën, BA Pickut, M Van den Broeck, S Serneels, M Cruts, C Van Broeckhoven, P P De Deyn

Please visit the *Journal of Medical Genetics* website [www.jmedgenet.com] for a link to the full text of this article.

Objective: The authors conducted a prospective study of neurodegenerative and vascular dementia in Belgium. Strict diagnostic inclusion criteria were used to include well defined patients and controls. The results of apolipoprotein E (APOE) genotype effect on risk and clinical characteristics are presented.
Methods: APOE genotyping was performed in patients with probable Alzheimer's disease (AD) (n = 504), frontotemporal dementia (FTD) (n = 47), vascular dementia (VaD) (n = 152), mixed dementia (n = 132), mild cognitive impairment (MCI) (n = 44), Parkinson's disease (PD) (n = 30), dementia with Lewy bodies (DLB) (n = 17), and multisystem atrophy (MSA)/progressive supranuclear palsy (PSP) (n = 12).
Results: The APOE allele frequencies of this Belgian control population (ε2: 6.9%; ε3: 76.2%; ε4: 16.9%) did not differ from those reported for other white populations. AD, MCI, and mixed dementia patients had higher APOE ε4 (32.9%, 38.6%, and 28.4% respectively) and lower APOE ε3 (62.2%, 53.4%, and 66.3%) frequencies compared with controls, whereas only AD and mixed dementia patients had lower APOE ε2 frequencies (4.9% and 5.3%). Apart from a borderline significant different distribution of APOE allele frequencies in VaD patients compared with controls, no other differences were detected. The influence of APOE ε4 on clinical features of dementia was limited to lower age at onset in AD patients and a less pronounced negative correlation between age at onset and number of ε4 alleles in MCI and mixed dementia patients.
Conclusions: This study confirmed the risk association between APOE ε4 and AD. The observation that APOE ε4 is associated with mixed dementia reflected the role of AD in the aetiopathogenesis of this condition. Although MCI is an aetiologically heterogeneous syndrome, the increased APOE ε4 frequencies indicated that a large proportion of the MCI patients included in the study might be predisposed to develop AD.

▲ *Journal of Neurology, Neurosurgery, and Psychiatry* 2003;**74**:1148–1151.