Recalibration of the Gail model for predicting invasive breast cancer risk in Spanish women: a population-based cohort study

Roberto Pastor-Barriuso · Nieves Ascunce · María Ederra · Nieves Erdozáin · Alberto Murillo · José E. Alés-Martínez · Marina Pollán

Abstract The Gail model for predicting the absolute risk of invasive breast cancer has been validated extensively in US populations, but its performance in the international setting remains uncertain. We evaluated the predictive accuracy of the Gail model in 54,649 Spanish women aged 45–68 years who were free of breast cancer at the 1996–1998 baseline mammographic examination in the population-based Navarre Breast Cancer Screening Program. Incident cases of invasive breast cancer and competing deaths were ascertained until the end of 2005 (average follow-up of 7.7 years) through linkage with population-based cancer and mortality registries. The Gail model was tested for calibration and discrimination in its original form and after recalibration to the lower breast cancer incidence and risk factor prevalence in the study cohort, and compared through cross-validation with a Navarre model fully developed from this cohort. The original Gail model overpredicted significantly the 835 cases of invasive breast cancer observed in the cohort (ratio of expected to observed cases 1.46, 95 % CI 1.36–1.56). The recalibrated Gail model was well calibrated overall (expected-to-observed ratio 1.00, 95 % CI 0.94–1.07), but it tended to underestimate risk for women in low-risk quintiles and to overestimate risk in high-risk quintiles ($P = 0.01$). The Navarre model showed good cross-validated calibration overall (expected-to-observed ratio 0.98, 95 % CI 0.92–1.05) and in different cohort subsets. The Navarre and Gail models had modest cross-validated discrimination indexes of 0.542 (95 % CI 0.521–0.564) and 0.544 (95 % CI 0.523–0.565), respectively. Although the original Gail model cannot be applied directly to populations with different underlying rates of invasive breast cancer, it can readily be recalibrated to provide unbiased estimates of absolute risk in such populations. Nevertheless, its limited discrimination ability at the individual level highlights the need to develop extended models with additional strong risk factors.

Keywords Risk prediction model · Invasive breast cancer · Spanish cohort · Calibration and discrimination accuracy · Model recalibration · Screening applications

R. Pastor-Barriuso (✉) · M. Pollán
National Center for Epidemiology, Carlos III Institute of Health, Monforte de Lemos 5, 28029 Madrid, Spain
e-mail: rpastor@isciii.es

N. Ascunce · M. Ederra · N. Erdozáin · A. Murillo
Navarre Breast Cancer Screening Program, Navarre Institute of Public Health, Leyre 15, 31003 Pamplona, Spain

J. E. Alés-Martínez
Medical Oncology Unit, Hospital Nuestra Señora de Sonsoles, Juan Carlos I, 05004 Ávila, Spain

R. Pastor-Barriuso · N. Ascunce · M. Ederra · N. Erdozáin · M. Pollán
Consortium for Biomedical Research in Epidemiology and Public Health (CIBERESP), Madrid, Spain

Introduction

The Gail model for predicting the absolute risk of invasive breast cancer in white women combined relative risks associated to four traditional risk factors (age at menarche, number of breast biopsies, age at first live birth, and number of first-degree relatives with breast cancer) derived from a case–control study conducted in the Breast Cancer

Detection Demonstration Project [1] with baseline age-specific incidence rates of invasive breast cancer from population-based US cancer registries in the Surveillance, Epidemiology, and End Results Program [2]. This prediction model has been validated in several cohorts from the United States, including large general populations [3–5], regularly screened subpopulations at elevated risk [2–4], and small studies in high-risk clinics [6, 7]. The Gail model showed heterogeneous but generally acceptable calibration with modest discrimination ability among white US women [8], and it has been widely used to design international prevention trials [9, 10] and to counsel women about their individual risk [11].

Few and relatively small validation studies have been conducted in Western non-US populations, and none of them used a population-based cohort design. In the United Kingdom, the Gail model underestimated significantly breast cancer risk in 3,150 women attending a family history clinic [12]. In Italy, the Gail model showed good overall calibration but modest individual discrimination in 5,383 hysterectomized women enrolled in a breast cancer chemoprevention trial [13] and, more recently, in 10,031 female volunteers with high prevalence of risk factors who participated in the Florence cohort of the European Prospective Investigation into Cancer and Nutrition study [14].

The Gail model could be useful to predict the risk of developing invasive breast cancer in Spain, where all women aged 50–69 years are currently covered by population-based mammographic screening programs [15]. However, age-standardized breast cancer incidence rates in Spain (61 cases per 100,000 women in 2008) are substantially lower than those in the United States (76) and most countries in Northern (84), Western (90), and Southern Europe (69) [16]. Thus, to avoid a systematic overestimation of breast cancer risk among Spanish women, it may be necessary to recalibrate the Gail model for the different incidence rates of invasive breast cancer and prevalences of risk factors in the Spanish population. In this study, we evaluated the predictive accuracy of the Gail model in its original form and after recalibration in a large population-based cohort of women who participated in the Navarre Breast Cancer Screening Program (NBCSP), and compared its performance with that of a similar prediction model fully developed from this Spanish cohort.

Methods

Navarre Breast Cancer Screening Program

The NBCSP belongs to the European Breast Cancer Network and was the first population-based mammographic screening program implemented in Spain in September 1990. The initial target population covered all women aged 45–65 years residing in the northern Spanish province of Navarre, but this age range was extended to 69 years in 1998 (77,455 female inhabitants aged 45–69 years in 2001). The program achieved full coverage of the target population in 2 years, the period established as the screening interval. All performance indicators of the NBCSP during the period 1990–2004, including a participation rate for first invitation of 84 % and an adherence to successive invitations of 97 % [17], have consistently exceeded the reference levels set by European guidelines [18]. The population impact of the NBCSP on breast cancer incidence and mortality rates in Navarre has recently been reported [19, 20].

Study cohort, baseline assessment, and follow-up

A total of 62,909 women with no history of invasive or in situ breast cancer who resided in Navarre and were born between January 1, 1931 and December 31, 1952 were invited to participate in the fourth screening round of the fully consolidated NBCSP. Of these, 54,995 women agreed to participate and were mammographically screened between September 1996 and July 1998 (participation rate 87.4 %).

Baseline information on age at menarche, previous breast biopsy, number of births, age at first live birth, and number of first-degree relatives (mother or sisters) with breast cancer was obtained from structured questionnaires administered by trained interviewers in the fourth screening round. Most women who reported ever having a breast biopsy referred to tests performed outside the NBCSP, and hence we were unable to determine the precise number of previous breast biopsies. Also, atypical hyperplasia was only ascertained in a small subset of women with biopsies performed within the NBCSP and, therefore, not considered in risk predictions.

For the present study, we excluded 168 women with prevalent breast cancer at their baseline mammographic examination in the fourth screening round, as well as 3 women who developed breast cancer and 35 women who died within 180 days from baseline. We also excluded 113 women lost to follow-up after baseline examination and 27 women with missing baseline information on the required risk factors. Thus, the starting cohort consisted of 54,649 women aged 45–68 years who were followed for the period beginning 180 days after the 1996–1998 baseline examination through December 31, 2005. Breast cancer cases were ascertained through linkage with the population-based Navarre Cancer Registry [21], which records all incident cases of invasive or in situ breast cancer diagnosed since 1973 in women residing in Navarre. Case ascertainment during follow-up was likely to be complete, since the

registry searched all relevant case sources in addition to the NBCSP, with 99 % of breast cancer cases histologically verified and 0.8 % registered solely on the basis of death certificates in 1998–2002 [22]. Deaths from other causes were identified through the Navarre Mortality Registry, which includes all deaths registered in Spain among residents in Navarre. The municipal register of inhabitants and the regional health system were consulted to confirm that disease-free women were still living in Navarre at the end of follow-up. Only 292 women were lost to follow-up and censored at the time of their last visit to the NBCSP, while the remaining women were followed disease free through December 31, 2005.

During an average follow-up of 7.7 years, 835 cases of invasive breast cancer, 150 cases of ductal carcinoma in situ, and 2 cases of non-epithelial breast tumor were diagnosed. In addition, 1,218 other women died from causes not related to breast cancer. Hormone receptor status could be determined from pathology reports in 767 of the 835 invasive breast cancers (91.9 %), with 653 tumors positive for either estrogen (634) or progesterone receptors (486) and 114 tumors negative for both receptors.

Statistical analysis

The baseline hazards and hazard ratios of invasive breast cancer in the NBCSP cohort were estimated from a piecewise exponential model [23] with constant baseline hazards in each 5-year age interval from 45 to 74 years and the same ordinal risk factors as the original Gail model [1], except for the simpler never/ever classification for previous breast biopsy. In particular, the risk factors included in this model were age at menarche (coded as 0, 1, or 2 for ≥14, 12–13, or <12 years, respectively), previous breast biopsy (coded as 0 if no and 1 if yes), age at first live birth (coded as 0, 1, 2, or 3 for <20, 20–24, 25–29 or nulliparous, or ≥30 years, respectively), and number of first-degree relatives with breast cancer (coded as 0, 1, or 2 for 0, 1, or ≥2 affected relatives, respectively). The model also included interaction terms between age at first birth and number of affected relatives and between breast biopsy and age (coded as 0 if <50 and 1 if ≥50 years), so that the hazard ratio for breast biopsy was allowed to vary from age intervals below to those above 50 years. A detailed justification and specification of this model is provided in the statistical appendix (Supplementary Material 1). The composite hazards of death from other causes were calculated by dividing the observed number of deaths in the NBCSP cohort by the woman-years at risk in each 5-year age interval.

Following standard competing risk methods [1, 23], three alternative models were used to predict the absolute risk of invasive breast cancer for each NBCSP woman according to their own risk factor profile. The Navarre model was based on baseline hazards and hazard ratios of invasive breast cancer estimated from the above piecewise exponential model in the NBCSP cohort, as well as on composite hazards of competing death among NBCSP women. The original Gail model used Gail relative risk estimates [1] and invasive breast cancer and mortality rates for white US women [2], whereas the recalibrated Gail model combined the original relative risk estimates [1] with composite incidence rates of invasive breast cancer, composite mortality rates, and risk factor prevalences among cases from the NBCSP cohort. The Gail relative risk for women with any previous breast biopsy was calculated as a weighted average of the reported relative risks [1] for one and two or more biopsies. Further details on the development of these prediction models are provided in the statistical appendix (Supplementary Material 1).

Calibration and discrimination of the three prediction models among NBCSP women were evaluated through a 10-fold cross-validation to correct for the optimistic bias induced by testing the Navarre model on the same training NBCSP data [24]. Calibration was assessed by comparing the observed cases of invasive breast cancer in the NBCSP cohort by age interval, risk factor category, and quintile of predicted 5-year risk with those expected under the Navarre, original Gail, and recalibrated Gail models [25]. Discrimination was evaluated using overall and age-specific C indexes [26], which are extensions of the area under the receiver-operating curve to censored time-to-event data. The discrimination ability of the Gail model remained unchanged after recalibration. Further details on cross-validated calibration and discrimination statistics are provided in the statistical appendix (Supplementary Material 1).

Results

Cause-specific hazards and hazard ratios from NBCSP cohort

The hazard of invasive breast cancer was higher in NBCSP women with previous benign breast biopsies, and it increased with decreasing age at menarche and with increasing age at first live birth and number of affected first-degree relatives. These hazard ratios were similar in direction but lower in magnitude than those from the Gail model, particularly for the strata of age at first birth by number of affected relatives (Table 1). Contrary to the Gail model, there was no significant interaction between breast biopsy and age ($P = 0.97$) or between age at first birth and number of affected relatives ($P = 0.23$). The population attributable risk for all four factors was 0.280 and varied little with age.

Table 1 Hazard ratios of invasive breast cancer by risk factor category in the Navarre Breast Cancer Screening Program cohort, 1996–1998 to 2005

Risk factor	No. of women	No. of woman-years	No. of invasive breast cancers	HR, NBCSP (95 % CI)[a]	OR, Gail (95 % CI)[b]
Age at menarche (years)					
≥14	23,530	181,394	335	1.00 (reference)	1.00 (reference)
12–13	25,198	193,970	413	1.07 (0.97–1.19)	1.10 (1.02–1.19)
<12	5,921	45,436	87	1.15 (0.94–1.41)	1.21 (1.03–1.41)
Previous breast biopsy					
Age <50 years					
No	12,289	29,166	42	1.00 (reference)	1.00 (reference)
Yes	1,221	2,950	7	1.65 (0.74–3.67)	1.89 (1.50–2.38)[c]
Age ≥50 years					
No	49,562	353,410	673	1.00 (reference)	1.00 (reference)
Yes	4,966	35,274	113	1.67 (1.37–2.04)	1.36 (1.19–1.56)[c]
No. of affected first-degree relatives					
Age at first live birth <20 years					
0	983	7,593	10	1.00 (reference)	1.00 (reference)
1	47	370	1	0.92 (0.45–1.87)	2.61 (1.99–3.42)
≥2	1	8	0	0.84 (0.20–3.51)	6.80 (3.96–11.68)
Age at first live birth 20–24 years					
0	15,377	118,331	204	1.11 (1.01–1.23)	1.24 (1.16–1.33)
1	879	6,746	12	1.26 (0.81–1.95)	2.68 (2.23–3.22)
≥2	36	264	1	1.42 (0.62–3.28)	5.78 (4.14–8.06)
Age at first live birth 25–29 years or nulliparous					
0	27,437	211,445	434	1.24 (1.01–1.52)	1.55 (1.35–1.78)
1	1,696	12,975	39	1.72 (1.28–2.32)	2.76 (2.32–3.27)
≥2	108	810	2	2.39 (1.48–3.88)	4.91 (3.76–6.41)
Age at first live birth ≥30 years					
0	7,575	58,382	116	1.38 (1.02–1.88)	1.93 (1.56–2.38)
1	491	3,726	16	2.36 (1.51–3.67)	2.83 (2.22–3.62)
≥2	19	150	0	4.02 (1.82–8.88)	4.17 (2.75–6.31)

[a] Hazard ratios (HRs) and 95 % confidence intervals (CIs) of invasive breast cancer estimated from the Navarre Breast Cancer Screening Program (NBCSP) cohort by fitting a piecewise exponential model with the same risk factors and ordinal codes as the original Gail model

[b] Odds ratios (ORs) and 95 % confidence intervals (CIs) of invasive or in situ breast cancer derived from the original Gail logistic model in the Breast Cancer Detection Demonstration Project case–control study [1]

[c] The age-specific odds ratios of breast cancer for women with any previous breast biopsy were calculated by combining the age-specific odds ratios for women with one and two or more biopsies reported in the original Gail model [1] (see statistical appendix in Supplementary Material 1)

The baseline hazards of invasive breast cancer from the NBCSP cohort increased steadily in screened women aged 45–64 years and declined in unscreened older women. These baseline incidence rates were similar to those derived from the Navarre Cancer Registry, except that the latter also included prevalent cases aged 45–49 years detected at their first participation in the NBCSP (Table 2). The composite mortality rates from other causes in the NBCSP cohort increased sharply with age but were 18.8 % [standardized mortality ratio 0.812, 95 % confidence interval (CI) 0.768–0.859] lower than those registered in the entire female population of Navarre (Table 2), suggesting that self-selected women in the NBCSP cohort were somewhat healthier than the general female population.

Calibration of prediction models

The Navarre model showed good cross-validated calibration overall (ratio of expected to observed cases 820.1/835 = 0.98, 95 % CI 0.92–1.05), as well as across categories of age at menarche (goodness-of-fit $P = 0.42$), breast biopsy by age ($P = 0.99$), and age at first birth by number of affected relatives ($P = 0.95$). The original Gail model overestimated significantly the absolute risk of invasive

Table 2 Age-specific incidence rates of invasive breast cancer and mortality rates from other causes (per 100,000 woman-years) in the Navarre Breast Cancer Screening Program cohort, 1996–1998 to 2005

Age (years)	No. of woman-years	Invasive breast cancer			Death from other causes		
		No. of cases	Baseline rate, NBCSP (95 % CI)[a]	External baseline rate, Navarre[b]	No. of deaths	Composite rate, NBCSP (95 % CI)[c]	External composite rate, Navarre[d]
45–49	32,116	49	111.3 (77.4–160.1)	156.6	34	105.9 (75.6–148.2)	115.2
50–54	96,663	185	138.9 (108.1–178.4)	134.8	151	156.2 (133.2–183.2)	166.6
55–59	100,953	202	144.0 (112.2–184.9)	146.4	204	202.1 (176.2–231.8)	239.7
60–64	86,085	193	160.4 (124.4–206.9)	175.9	241	280.0 (246.8–317.6)	327.8
65–69	77,621	166	152.8 (117.3–199.1)	146.2	373	480.5 (434.2–531.9)	619.5
70–74	27,362	40	104.3 (71.3–152.4)	114.1	215	785.8 (687.4–898.1)	1085.2

[a] Baseline incidence rates of invasive breast cancer and 95 % confidence intervals (CIs) for a woman at the reference level of all risk factors (age at menarche ≥14 years, no previous breast biopsy, age at first live birth <20 years, and no affected first-degree relatives) estimated from a piecewise exponential model in the Navarre Breast Cancer Screening Program (NBCSP) cohort

[b] External baseline incidence rates of invasive breast cancer for a woman at the reference level of all risk factors (age at menarche ≥14 years, no previous breast biopsy, age at first live birth <20 years, and no affected first-degree relatives) calculated as the composite incidence rates of invasive breast cancer for the period 2000–2004 obtained from the Navarre Cancer Registry multiplied by one minus the overall attributable risk estimated from the Navarre Breast Cancer Screening Program cohort

[c] Composite mortality rates from other causes and 95 % confidence intervals (CIs) among women in the Navarre Breast Cancer Screening Program (NBCSP) cohort

[d] External composite mortality rates from other causes in the entire female population of Navarre during the period 2000–2004 obtained from the Spanish National Institute of Statistics

breast cancer in the NBCSP cohort by 46 % (expected-to-observed ratio 1215.5/835 = 1.46, 95 % CI 1.36–1.56), with greater overprediction in the older age intervals (Table 3). This systematic overestimation disappeared after recalibrating the Gail model (expected-to-observed ratio 836.4/835 = 1.00, 95 % CI 0.94–1.07), with no significant lack of fit across the three risk factor categorizations (P = 0.48, 0.36, and 0.15, respectively).

The median predicted 5-year risks of invasive breast cancer were 0.93, 1.31, and 0.95 % under the Navarre, original Gail, and recalibrated Gail models, respectively, with 2.9, 25.6, and 4.1 % of NBCSP women above the standard risk threshold of 1.67 %. The Navarre model showed good agreement between observed and expected cases by quintile of predicted 5-year risk (goodness-of-fit P = 0.36). The original Gail model overpredicted significantly invasive breast cancer cases in all quintiles of risk (Table 4). The recalibrated Gail model corrected this systematic overprediction (goodness-of-fit P = 0.25), but due to the larger Gail relative risks, it still showed a significant positive trend in the expected-to-observed ratios across quintiles of risk (P for linear trend = 0.01).

Discrimination of prediction models

Overall, the cross-validated discrimination indexes among NBCSP women were modest and equal to 0.542 (95 % CI 0.521–0.564) for the Navarre model and 0.544 (95 % CI 0.523–0.565) for the Gail model, with no significant

difference between models (P = 0.67). Discrimination remained similar in age intervals below 70 years and increased marginally to 0.628 for the Navarre model and 0.626 for the Gail model among women aged 70–74 years (P for deviation from overall discrimination = 0.09 and 0.08, respectively; Table 5).

The cross-validated discrimination indexes were somewhat better for hormone receptor-positive invasive breast cancers (0.545, 95 % CI 0.521–0.569, for the Navarre model and 0.543, 95 % CI 0.519–0.567, for the Gail model) than for hormone receptor-negative cancers (0.508, 95 % CI 0.446–0.571, and 0.530, 95 % CI 0.469–0.591, respectively).

Discussion

The original Gail model overestimated the actual invasive breast cancer incidence by 46 % in a large population-based cohort of biennially screened Spanish women aged 45–68 years who were followed for an average of 7.7 years. The recalibrated Gail model was well calibrated overall, but it still underestimated breast cancer risk for women with a low risk-factor profile and overestimated risk for women with a high risk-factor profile. The Navarre model showed good cross-validated calibration overall and in different cohort subsets. Nevertheless, both the Navarre and Gail models had limited discrimination ability of 0.54 in this cohort.

Table 3 Ratios of the expected cases of invasive breast cancer under the Navarre, original Gail, and recalibrated Gail prediction models to the observed cases in the Navarre Breast Cancer Screening Program cohort by age interval and risk factor category, 1996–1998 to 2005

Stratum	Observed cases of invasive breast cancer	Navarre model		Original Gail model[a]		Recalibrated Gail model[a]	
		Expected cases of invasive breast cancer[b]	Ratio of expected to observed cases (95 % CI)[c]	Expected cases of invasive breast cancer[b]	Ratio of expected to observed cases (95 % CI)[c]	Expected cases of invasive breast cancer[b]	Ratio of expected to observed cases (95 % CI)[c]
Age interval (years)							
45–49	49	48.87	1.00 (0.75–1.32)	61.60	1.26 (0.95–1.66)	50.81	1.04 (0.78–1.37)
50–54	185	182.97	0.99 (0.86–1.14)	208.61	1.13 (0.98–1.30)	183.01	0.99 (0.86–1.14)
55–59	202	198.90	0.98 (0.86–1.13)	268.34	1.33 (1.16–1.52)	201.40	1.00 (0.87–1.14)
60–64	193	189.49	0.98 (0.85–1.13)	281.55	1.46 (1.27–1.68)	194.40	1.01 (0.87–1.16)
65–69	166	161.72	0.97 (0.84–1.13)	292.15	1.76 (1.51–2.05)	167.14	1.01 (0.86–1.17)
70–74	40	38.14	0.95 (0.70–1.30)	103.24	2.58 (1.89–3.52)	39.61	0.99 (0.73–1.35)
Age at menarche (years)							
≥14	335	339.14	1.01 (0.91–1.13)	508.56	1.52 (1.36–1.69)	341.86	1.02 (0.92–1.14)
12–13	413	385.28	0.93 (0.85–1.03)	567.08	1.37 (1.25–1.51)	395.11	0.96 (0.87–1.05)
<12	87	95.68	1.10 (0.89–1.36)	139.86	1.61 (1.30–1.98)	99.41	1.14 (0.93–1.41)
Previous breast biopsy							
Age <50 years							
No	42	41.84	1.00 (0.74–1.35)	51.60	1.23 (0.91–1.66)	42.56	1.01 (0.75–1.37)
Yes	7	7.03	1.00 (0.48–2.11)	10.00	1.43 (0.68–3.00)	8.25	1.18 (0.56–2.47)
Age ≥50 years							
No	673	660.74	0.98 (0.91–1.06)	1016.75	1.51 (1.40–1.63)	691.11	1.03 (0.95–1.11)
Yes	113	110.48	0.98 (0.81–1.18)	137.14	1.21 (1.01–1.46)	94.45	0.84 (0.70–1.01)
No. of affected first-degree relatives							
Age at first live birth <20 years							
0	10	11.85	1.18 (0.64–2.20)	13.60	1.36 (0.73–2.53)	9.60	0.96 (0.52–1.78)
1	1	0.54	0.54 (0.10–21.35)	1.72	1.72 (0.31–67.95)	1.25	1.25 (0.22–49.28)
≥2	0	0.01	–	0.12	–	0.06	–
Age at first live birth 20–24 years							
0	204	206.10	1.01 (0.88–1.16)	259.50	1.27 (1.11–1.46)	185.37	0.91 (0.79–1.04)
1	12	13.47	1.12 (0.64–1.98)	31.98	2.67 (1.51–4.69)	22.84	1.90 (1.08–3.35)
≥2	1	0.61	0.61 (0.11–23.98)	2.60	2.60 (0.47–102.68)	1.87	1.87 (0.34–73.89)
Age at first live birth 25–29 years or nulliparous							
0	434	409.97	0.94 (0.86–1.04)	600.26	1.38 (1.26–1.52)	411.27	0.95 (0.86–1.04)
1	39	35.33	0.91 (0.66–1.24)	66.61	1.71 (1.25–2.34)	45.08	1.16 (0.84–1.58)
≥2	2	3.11	1.55 (0.43–12.83)	7.50	3.75 (0.94–14.99)	5.03	2.52 (0.63–10.06)
Age at first live birth ≥30 years							
0	116	124.45	1.07 (0.89–1.29)	210.19	1.81 (1.51–2.17)	139.99	1.21 (1.01–1.45)
1	16	13.77	0.86 (0.53–1.41)	20.11	1.26 (0.77–2.05)	13.23	0.83 (0.51–1.35)
≥2	0	0.90	–	1.30	–	0.78	–
Overall	835	820.10	0.98 (0.92–1.05)	1215.49	1.46 (1.36–1.56)	836.37	1.00 (0.94–1.07)

To correct for the optimistic bias induced by assessing calibration of the Navarre prediction model on the same data used to fit the model, a 10-fold cross-validation was used in which the absolute risk of invasive breast cancer for women in each 10 % random subcohort was calculated based on cause-specific hazards and hazard ratios estimated from the remaining 90 % of women in the Navarre Breast Cancer Screening Program cohort (see statistical appendix in Supplementary Material 1)

[a] The Gail prediction model was tested for calibration in its original form, which combined the original relative risk estimates [1] with invasive breast cancer and mortality rates for white women in the United States [2], and after recalibration, which combined the original relative risk estimates [1] with cross-validated estimates of composite invasive breast cancer and mortality rates and risk factor prevalences among cases from the Navarre Breast Cancer Screening Program cohort (see statistical appendix in Supplementary Material 1)

[b] The expected number of invasive breast cancer cases for a given age interval or risk factor category was calculated as the sum of the individual absolute risks of invasive breast cancer predicted by the models over that age interval or risk factor category

[c] Ratios of expected to observed cases and 95 % confidence intervals (CIs) assuming a negligible variance for the expected number of cases and a Poisson variance for the observed number of cases. If the expected number of cases was below 5, exact 95 % CIs were calculated based on a Poisson distribution

Comparison with other studies

Model calibration is strongly affected by temporal and geographical variations in disease incidence. The Gail model used invasive breast cancer rates among white US women for the period 1983–1987 [2]. Since breast cancer incidence increased steadily during the 1990 s in the United States [27], subsequent validation studies of the Gail model resulted in overall underestimations of invasive breast cancer risk by 6 % in the Nurses' Health Study [3], by 21 % in the Women's Health Initiative [4], and by 13–14 % in two other recent US cohorts [5]. Thus, claims have been raised about the need to update invasive breast cancer rates used in the Gail model to ensure a good overall calibration in recent US cohorts [3, 5, 28]. Our results further highlight that, due to large worldwide variations in breast cancer incidence [16], the Gail model should also be recalibrated when applied to the international setting [29]. The lower breast cancer incidence rates in Spain compared with the United States caused the Gail model to overestimate breast cancer risk by 46 % in this Spanish cohort. This systematic overprediction was corrected after recalibrating the Gail model to the lower incidence rates and risk factor prevalences in the study cohort.

The lower incidence of breast cancer in Spain can hardly be explained by differences in regular mammography use since its prevalence is similar in Spain (59 % of women aged 45 years or older in 2006) [30] and the United States (67 % of women aged 40 years or older in 2005) [31]. The distribution of Gail risk factors could better account for part of the observed differences in countrywide rates, as women younger than 12 years at menarche, with biopsy examinations, and with affected first-degree relatives were half as prevalent in the 1996–1998 baseline assessment of this Spanish cohort as in concurrent assessments of large representative US cohorts [4, 5]. Nevertheless, the baseline incidence rates of invasive breast cancer for NBCSP women aged 45–74 years were still 16 % lower than those used in the Gail model [2], suggesting that other factors may contribute to these differences. Obesity is more prevalent among adult white US women [32] than their counterparts in Spain [33]. Moreover, more than one-third of postmenopausal women in the United States were taking hormone replacement therapy between 1995 and 2001 [34], whereas this therapy was rarely used in Spain [35].

The relative risks estimated from this Spanish cohort were lower than those reported in the Gail model [1] which, combined with the smaller risk factor prevalences, resulted in an attributable risk of 0.28, substantially lower than the value of 0.42 found in white US women [2]. The lower relative risks for age at menarche, age at first birth, and number of affected relatives may be explained by the later age at diagnosis of breast cancer cases: only 6 % of cases in our cohort were diagnosed before 50 years of age, as opposed to the 29 % enrolled in the Gail analysis [1]. There is compelling evidence that reproductive [36] and familial factors [37] have stronger effects on the risk of early-onset than late-onset breast tumors. In fact, these risk factors showed consistently weaker associations in three large US cohorts of postmenopausal women [4, 5] than in the Gail model.

Clinical and public health implications and future research

The less pronounced relative risks observed in this Spanish population resulted in a modest discrimination of 0.54 for both the Navarre and Gail models, somewhat lower than the values of 0.58–0.59 reported for the Gail model among white US [3–5] and Italian women [14]. Well-calibrated prediction models with limited discrimination ability, such as the Navarre and recalibrated Gail models, may be useful in clinical practice for counseling individual patients on the risks and benefits of a preventive treatment [38], as well as for designing adequately powered intervention trials. However, higher discrimination is required for implementing an effective prevention strategy in high-risk subsets of the general population, in order to achieve large reductions in disease incidence [39]. The inclusion of 7–18 common genetic variants for breast cancer has been shown to increase discrimination of the Gail model by 0.03–0.07 [40–42]. Apart from the substantial costs of obtaining genetic information, this modest improvement in discrimination was similar to the increase of 0.05 obtained from adding only mammographic density, a strong and highly prevalent risk factor [43]. A nested case–control study is currently being conducted within the NBCSP to obtain mammographic density measurements in nearly 1,000 incident cases of invasive breast cancer and 4,000 disease-free women. This case–control study might provide valuable data to improve the discrimination accuracy of the Navarre model among Spanish women by including mammographic density and enhanced family history information on breast cancer.

Strengths and limitations of the study

The strengths of this study include the use of a large representative cohort of regularly screened Spanish women, the high participation rate, and the relatively long follow-up period with negligible losses to follow-up, virtually complete case ascertainment, and information on tumor receptor status.

The study has several limitations. First, nearly all breast cancer cases were diagnosed in women aged 50 years or older, so our findings may not apply to younger premenopausal women in regular screening. Second, information on atypical hyperplasia was not available in 4,462

Table 4 Ratios of expected to observed cases of invasive breast cancer in the Navarre Breast Cancer Screening Program cohort by quintile of predicted 5-year risk based on the Navarre, original Gail, and recalibrated Gail prediction models, 1996–1998 to 2005

Predicted 5-year risk (%)	No. of women	No. of woman-years	Observed cases of invasive breast cancer	Expected cases of invasive breast cancer[b]	Ratio of expected to observed cases (95 % CI)[c]
Navarre prediction model					
0.40–0.81	10,930	84,233	154	129.95	0.84 (0.72–0.99)
0.82–0.88	10,930	83,860	144	143.72	1.00 (0.85–1.18)
0.89–0.95	10,929	84,274	155	154.16	0.99 (0.85–1.16)
0.96–1.05	10,930	84,457	170	165.91	0.98 (0.84–1.13)
1.06–5.59	10,930	83,976	212	226.36	1.07 (0.93–1.22)
Original Gail prediction model[a]					
0.54–0.96	10,930	84,783	126	151.09	1.20 (1.01–1.43)
0.97–1.18	10,931	84,260	156	190.94	1.22 (1.05–1.43)
1.19–1.44	10,927	84,287	173	229.06	1.32 (1.14–1.54)
1.45–1.74	10,931	84,041	175	273.13	1.56 (1.35–1.81)
1.75–7.70	10,930	83,429	205	371.28	1.81 (1.58–2.08)
Recalibrated Gail prediction model[a]					
0.39–0.77	10,930	84,296	138	120.15	0.87 (0.74–1.03)
0.78–0.88	10,929	83,973	151	141.16	0.93 (0.80–1.10)
0.89–0.98	10,930	84,138	158	157.62	1.00 (0.85–1.17)
0.99–1.16	10,930	84,409	175	176.26	1.01 (0.87–1.17)
1.17–4.96	10,930	83,984	213	241.18	1.13 (0.99–1.30)

To correct for the optimistic bias induced by assessing calibration of the Navarre prediction model on the same data used to fit the model, a 10-fold cross-validation was used in which the absolute risk of invasive breast cancer for women in each 10 % random subcohort was calculated based on cause-specific hazards and hazard ratios estimated from the remaining 90 % of women in the Navarre Breast Cancer Screening Program cohort (see statistical appendix in Supplementary Material 1)

[a] The Gail prediction model was tested for calibration in its original form, which combined the original relative risk estimates [1] with invasive breast cancer and mortality rates for white women in the United States [2], and after recalibration, which combined the original relative risk estimates [1] with cross-validated estimates of composite invasive breast cancer and mortality rates and risk factor prevalences among cases from the Navarre Breast Cancer Screening Program cohort (see statistical appendix in Supplementary Material 1)

[b] The expected number of invasive breast cancer cases in each quintile of predicted 5-year risk was calculated as the sum of the individual absolute risks of invasive breast cancer predicted by the model over all women in that quintile

[c] Ratios of expected to observed cases and 95 % confidence intervals (CIs) assuming a negligible variance for the expected number of cases and a Poisson variance for the observed number of cases

of the 4,983 women with previous biopsy because they referred to tests performed outside the NBCSP. Of the remaining 521 women with biopsies performed within the program, 16 had atypical hyperplasia. Thus, we can infer that roughly 0.3 % of the entire NBCSP cohort had atypical hyperplasia (3.1 % with atypia out of 9.1 % with biopsy) and that the overall performance of the Navarre and Gail models was little affected by knowledge of atypical hyperplasia status. However, atypical hyperplasia is a strong risk factor for breast cancer [44] and these models will substantially underestimate breast cancer risk in women with atypia, as has already been reported in other cohorts [7]. Third, nondifferential misclassification of baseline exposure [45] might have partially accounted for the low relative risks and discrimination ability of the Gail model in this Spanish cohort. Nevertheless, data were collected from structured personal interviews and self-reported Gail model variables, including family history of breast cancer in first-degree relatives [46], are typically accurate in this setting. Finally, cross-validation was used to obtain overfitting-corrected estimates of the expected internal validity of the Navarre model in new subjects from the same population, but a more stringent external validation would be required in related but different populations.

Conclusions

The Gail model cannot be applied directly to populations with different underlying rates of invasive breast cancer, but it can readily be recalibrated to provide unbiased estimates of absolute risk in these populations. In our study, the original Gail model showed a substantial overestimation of breast cancer risk that was corrected after recalibrating the model to

Table 5 Overall and age-specific discrimination of the Navarre and Gail prediction models among women in the Navarre Breast Cancer Screening Program cohort, 1996–1998 to 2005

Age (years)	No. of invasive breast cancers	No. of comparable pairs[a]	Navarre prediction model		Gail prediction model		Discrimination difference (95 % CI)[c]
			No. of concordant pairs[b]	Discrimination index (95 % CI)[c]	No. of concordant pairs[b]	Discrimination index (95 % CI)[c]	
45–49	49	44,932	25748.5	0.573 (0.497–0.649)	25879.5	0.576 (0.501–0.651)	−0.003 (−0.015 to 0.009)
50–54	185	378,725	206573.5	0.545 (0.502–0.589)	209957.5	0.554 (0.512–0.597)	−0.009 (−0.031 to 0.013)
55–59	202	447,830	240137.5	0.536 (0.492–0.580)	243176.5	0.543 (0.499–0.587)	−0.007 (−0.029 to 0.016)
60–64	193	368,818	199167.5	0.540 (0.496–0.584)	195754.5	0.531 (0.488–0.574)	0.009 (−0.001 to 0.019)
65–69	166	271,740	145556.5	0.536 (0.490–0.581)	145508.5	0.535 (0.490–0.581)	0.000 (−0.011 to 0.011)
70–74	40	31,467	19753.0	0.628 (0.531–0.725)	19683.0	0.626 (0.534–0.717)	0.002 (−0.038 to 0.042)
Overall				0.542 (0.521–0.564)		0.544 (0.523–0.565)	−0.002 (−0.011 to 0.007)

To correct for the optimistic bias arising from assessing discrimination of the Navarre prediction model on the same data used to fit the model, a 10-fold cross-validation was used in which separate discrimination indexes were estimated for each 10 % random subcohort based on hazard ratios estimated from the remaining 90 % of women in the Navarre Breast Cancer Screening Program cohort and the resulting estimates were combined over the 10 subcohorts (see statistical appendix in Supplementary Material 1)

[a] Number of woman pairs in which their actual times to invasive breast cancer in the corresponding age interval can be ranked (two different event times or a censoring time equal to or longer than an event time)

[b] Number of comparable pairs in which the woman with shorter event time had a higher relative risk of developing breast cancer in the corresponding age interval according to the Navarre or Gail prediction model. When predicted risks were identical for a woman pair (the same risk factor pattern), a value of 0.5 was added to the number of concordant pairs

[c] Age-specific discrimination indexes for the Navarre and Gail prediction models were estimated as the proportion of concordant pairs over the total number of comparable pairs in the corresponding age interval. Jackknife methods were used to compute 95 % confidence intervals (CIs) for each model's age-specific discrimination as well as for the difference in age-specific discrimination between models. Overall discrimination indexes were calculated as the average of age-specific discrimination indexes with weights proportional to the number of comparable pairs in each age interval

the lower breast cancer incidence rates and risk factor prevalences in this Spanish cohort. Nevertheless, the limited discrimination ability of the Navarre and Gail models among Spanish women precludes their use for screening applications and highlights the need to develop extended models with additional strong risk factors, such as mammographic density and detailed family history.

Conflict of interest The authors have no conflicts of interest to disclose.

References

1. Gail MH, Brinton LA, Byar DP, Corle DK, Green SB, Schairer C, Mulvihill JJ (1989) Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. J Natl Cancer Inst 81:1879–1886

2. Costantino JP, Gail MH, Pee D, Anderson S, Redmond CK, Benichou J, Wieand HS (1999) Validation studies for models projecting the risk of invasive and total breast cancer incidence. J Natl Cancer Inst 91:1541–1548

3. Rockhill B, Spiegelman D, Byrne C, Hunter DJ, Colditz GA (2001) Validation of the Gail et al. model of breast cancer risk prediction and implications for chemoprevention. J Natl Cancer Inst 93:358–366

4. Chlebowski RT, Anderson GL, Lane DS, Aragaki AK, Rohan T, Yasmeen S, Sarto G, Rosenberg CA, Hubbell FA (2007) Predicting risk of breast cancer in postmenopausal women by hormone receptor status. J Natl Cancer Inst 99:1695–1705

5. Schonfeld SJ, Pee D, Greenlee RT, Hartge P, Lacey JV Jr, Park Y, Schatzkin A, Visvanathan K, Pfeiffer RM (2010) Effect of changing breast cancer incidence rates on the calibration of the Gail model. J Clin Oncol 28:2411–2417

6. Bondy ML, Lustbader ED, Halabi S, Ross E, Vogel VG (1994) Validation of a breast cancer risk assessment model in women with a positive family history. J Natl Cancer Inst 86:620–625

7. Pankratz VS, Hartmann LC, Degnim AC, Vierkant RA, Ghosh K, Vachon CM, Frost MH, Maloney SD, Reynolds C, Boughey JC (2008) Assessment of the accuracy of the Gail model in women with atypical hyperplasia. J Clin Oncol 26:5374–5379

8. Meads C, Ahmed I, Riley RD (2012) A systematic review of breast cancer incidence risk prediction models with meta-analysis of their performance. Breast Cancer Res Treat 132:365–377

9. Fisher B, Costantino JP, Wickerham DL, Redmond CK, Kavanah M, Cronin WM, Vogel V, Robidoux A, Dimitrov N, Atkins J, Daly M, Wieand S, Tan-Chiu E, Ford L, Wolmark N (1998) Tamoxifen for prevention of breast cancer: report of the National Surgical Adjuvant Breast and Bowel Project P-1 Study. J Natl Cancer Inst 90:1371–1388

10. Goss PE, Ingle JN, Alés-Martínez JE, Cheung AM, Chlebowski RT, Wactawski-Wende J, McTiernan A, Robbins J, Johnson KC, Martin LW, Winquist E, Sarto GE, Garber JE, Fabian CJ, Pujol P, Maunsell E, Farmer P, Gelmon KA, Tu D, Richardson H (2011) Exemestane for breast-cancer prevention in postmenopausal women. N Engl J Med 364:2381–2391

11. Visvanathan K, Chlebowski RT, Hurley P, Col NF, Ropka M, Collyar D, Morrow M, Runowicz C, Pritchard KI, Hagerty K, Arun B, Garber J, Vogel VG, Wade JL, Brown P, Cuzick J, Kramer BS, Lippman SM (2009) American society of clinical oncology clinical practice guideline update on the use of pharmacologic interventions including tamoxifen, raloxifene, and aromatase inhibition for breast cancer risk reduction. J Clin Oncol 27:3235–3258

12. Amir E, Evans DG, Shenton A, Lalloo F, Moran A, Boggis C, Wilson M, Howell A (2003) Evaluation of breast cancer risk assessment packages in the family history evaluation and screening programme. J Med Genet 40:807–814

13. Boyle P, Mezzetti M, La Vecchia C, Franceschi S, Decarli A, Robertson C (2004) Contribution of three components to individual cancer risk predicting breast cancer risk in Italy. Eur J Cancer Prev 13:183–191

14. Decarli A, Calza S, Masala G, Specchia C, Palli D, Gail MH (2006) Gail model for prediction of absolute risk of invasive breast cancer: independent evaluation in the Florence-European Prospective Investigation into Cancer and Nutrition cohort. J Natl Cancer Inst 98:1686–1693

15. Ascunce N, Salas D, Zubizarreta R, Almazán R, Ibáñez J, Ederra M (2010) Cancer screening in Spain. Ann Oncol 21(Suppl 3):iii43–iii51

16. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM (2010) GLOBOCAN 2008, cancer incidence and mortality worldwide: IARC CancerBase No. 10. International Agency for Research on Cancer, Lyon. http://globocan.iarc.fr. Accessed Dec 2012

17. Ascunce N, Ederra M, Barcos A, Zubizarreta R, Fernández AB, Casamitjana M (2007) Situación del cribado de cáncer de mama en España: características y principales resultados de los programas existentes. In: Castells X, Sala M, Ascunce N, Salas D, Zubizarreta R, Casamitjana M (eds) Descripción del cribado del cáncer en España. Proyecto DESCRIC. Ministerio de Sanidad y Consumo, Madrid, pp 31–73

18. Perry N, Broeders M, de Wolf C, Törnberg S, Holland R, von Karsa L (2006) European guidelines for quality assurance in breast cancer screening and diagnosis, 4th edn. Office for Official Publications of the European Communities, Luxembourg

19. Pollán M, Pastor-Barriuso R, Ardanaz E, Argüelles M, Martos C, Galcerán J, Sánchez-Pérez MJ, Chirlaque MD, Larrañaga N, Martínez-Cobo R, Tobalina MC, Vidal E, Marcos-Gragera R, Mateos A, Garau I, Rojas-Martín MD, Jiménez R, Torrella-Ramos A, Perucha J, Pérez-de-Rada ME, González S, Rabanaque MJ, Borràs J, Navarro C, Hernández E, Izquierdo A, López-Abente G, Martínez C (2009) Recent changes in breast cancer incidence in Spain, 1980–2004. J Natl Cancer Inst 101:1584–1591

20. Ascunce N, Moreno-Iribas C, Barcos-Urtiaga A, Ardanaz E, Ederra-Sanz M, Castilla J, Egüés N (2007) Changes in breast cancer mortality in Navarre (Spain) after introduction of a screening programme. J Med Screen 14:14–20

21. Registro de Cáncer de Navarra, Instituto de Salud Pública de Navarra. http://www.navarra.es/home_es/Temas/Portal+de+la+Salud/Profesionales/Planes+y+programas/Registro+de+cancer+de+Navarra. Accessed Dec 2012

22. Curado MP, Edwards B, Shin HR, Storm H, Ferlay J, Heanue M, Boyle P (2007) Cancer incidence in five continents, vol IX. International Agency for Research on Cancer, Lyon

23. Benichou J, Gail MH (1990) Estimates of absolute cause-specific risk in cohort studies. Biometrics 46:813–826

24. Efron B, Tibshirani RJ (1993) An introduction to the bootstrap. Chapman and Hall, London

25. Gail MH, Pfeiffer RM (2005) On criteria for evaluating models of absolute risk. Biostatistics 6:227–239

26. Pencina MJ, D'Agostino RB (2004) Overall C as a measure of discrimination in survival analysis: model specific population value and confidence interval estimation. Stat Med 23:2109–2123

27. Ravdin PM, Cronin KA, Howlader N, Berg CD, Chlebowski RT, Feuer EJ, Edwards BK, Berry DA (2007) The decrease in breast-cancer incidence in 2003 in the United States. N Engl J Med 356:1670–1674

28. Banegas MP, Gail MH, LaCroix A, Thompson B, Martinez ME, Wactawski-Wende J, John EM, Hubbell FA, Yasmeen S, Katki HA (2012) Evaluating breast cancer risk projections for Hispanic women. Breast Cancer Res Treat 132:347–353

29. Chay WY, Ong WS, Tan PH, Jie Leo NQ, Ho GH, Wong CS, Chia KS, Chow KY, Tan M Sr, Ang P Sr (2012) Validation of the Gail model for predicting individual breast cancer risk in a prospective nationwide study of 28,104 Singapore women. Breast Cancer Res 14:R19

30. Ministerio de Sanidad y Consumo (2007) Spanish National Health Survey 2006. Ministerio de Sanidad y Consumo, Madrid. http://www.msps.es/en/estadEstudios/estadisticas/encuestaNacional/encuesta2006.htm. Accessed Dec 2012

31. Swan J, Breen N, Graubard BI, McNeel TS, Blackman D, Tangka FK, Ballard-Barbash R (2010) Data and trends in cancer screening in the United States: results from the 2005 National Health Interview Survey. Cancer 116:4872–4881

32. Flegal KM, Carroll MD, Ogden CL, Curtin LR (2010) Prevalence and trends in obesity among US adults, 1999–2008. JAMA 303:235–241

33. Ortiz-Moncada R, Álvarez-Dardet C, Miralles-Bueno JJ, Ruíz-Cantero MT, dal Re-Saavedra MA, Villar-Villalba C, Pérez-Farinós N, Serra-Majem L (2011) Social determinants of overweight and obesity in Spain in 2006. Med Clin (Barc) 137:678–684

34. Hersh AL, Stefanick ML, Stafford RS (2004) National use of postmenopausal hormone therapy: annual trends and response to recent evidence. JAMA 291:47–53

35. Benet Rodríguez M, Carvajal García-Pando A, García del Pozo J, Álvarez Requejo A, Vega Alonso T (2002) Hormonal replacement therapy in Spain. Med Clin (Barc) 119:4–8

36. Clavel-Chapelon F, Gerber M (2002) Reproductive factors and breast cancer risk. Do they differ according to age at diagnosis? Breast Cancer Res Treat 72:107–115

37. Collaborative Group on Hormonal Factors in Breast Cancer (2001) 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 358:1389–1399

38. Gail MH, Costantino JP, Bryant J, Croyle R, Freedman L, Helzlsouer K, Vogel V (1999) Weighing the risks and benefits of tamoxifen treatment for preventing breast cancer. J Natl Cancer Inst 91:1829–1846

39. Gail MH (2011) Personalized estimates of breast cancer risk in clinical practice and public health. Stat Med 30:1090–1104

40. Gail MH (2008) Discriminatory accuracy from single-nucleotide polymorphisms in models to predict breast cancer risk. J Natl Cancer Inst 100:1037–1041

41. Wacholder S, Hartge P, Prentice R, Garcia-Closas M, Feigelson HS, Diver WR, Thun MJ, Cox DG, Hankinson SE, Kraft P, Rosner B, Berg CD, Brinton LA, Lissowska J, Sherman ME, Chlebowski R, Kooperberg C, Jackson RD, Buckman DW, Hui P, Pfeiffer R, Jacobs KB, Thomas GD, Hoover RN, Gail MH, Chanock SJ, Hunter DJ (2010) Performance of common genetic variants in breast-cancer risk models. N Engl J Med 362:986–993

42. Darabi H, Czene K, Zhao W, Liu J, Hall P, Humphreys K (2012) Breast cancer risk prediction and individualised screening based

on common genetic variation and breast density measurement. Breast Cancer Res 14:R25

43. Chen J, Pee D, Ayyagari R, Graubard B, Schairer C, Byrne C, Benichou J, Gail MH (2006) Projecting absolute invasive breast cancer risk in white women with a model that includes mammographic density. J Natl Cancer Inst 98:1215–1226

44. Hartmann LC, Sellers TA, Frost MH, Lingle WL, Degnim AC, Ghosh K, Vierkant RA, Maloney SD, Pankratz VS, Hillman DW, Suman VJ, Johnson J, Blake C, Tlsty T, Vachon CM, Melton LJ III, Visscher DW (2005) Benign breast disease and the risk of breast cancer. N Engl J Med 353:229–237

45. Schisterman EF, Faraggi D, Reiser B, Trevisan M (2001) Statistical inference for the area under the receiver operating characteristic curve in the presence of random measurement error. Am J Epidemiol 154:174–179

46. Murff HJ, Spigel DR, Syngal S (2004) Does this patient have a family history of cancer? An evidence-based analysis of the accuracy of family cancer history. JAMA 292:1480–1489