Am. J. Hum. Genet. 61:678–690, 1997

Interactions between Genetic and Reproductive Factors in Breast Cancer Risk in a French Family Sample

N. Andrieu[1] and F. Demenais[2]

[1]INSERM U. 351, Institut Gustave-Roussy, Villejuif; and [2]INSERM U. 358, Hôpital Saint-Louis, Paris

Summary

Considerable progress has been made in the characterization of the genetic component of breast cancer (BC). However, BC still remains a complex disease involving a genetic component and many other risk factors essentially linked to reproductive-life factors. To search for interactions between genetic and reproductive-life factors in the etiology of BC, a systematic family study was performed in two French hospitals from December 1987 to January 1990 and led to recruitment of 288 families, the IGRC data ("IGRC" refers to the Institut Gustave Roussy and Institut Curie, where the data were obtained). Detailed information on reproductive factors was recorded for probands and female first-degree relatives. Segregation analysis of BC was conducted by taking into account a variable age at onset of disease, by use of the class D regressive logistic model, as implemented in the REGRESS computer program. Segregation analyses of BC in IGRC data showed evidence for the segregation of a dominant gene and additional sister-sister dependence, both when reproductive factors were ignored and when they were included. A significant interaction was detected between the dominant gene and age when reproductive factors were taken into account. Among the reproductive factors included in segregation analysis, parity was found to interact with the dominant-gene effect, and there was an indication of an interaction, albeit not significant, between the dominant gene and age at menarche. Whereas the usual protective effect conferred on breast-cancer risk by high parity remained in nonsusceptible women, it disappeared in susceptible women. The increased BC risk associated with a late age at menarche was higher in susceptible women than in nonsusceptible women. Interactions between inherited predisposition to BC and reproductive factors were detected here for the first time by segregation analysis. It would be of major interest to confirm these results by family studies in other populations.

Received January 24, 1997; accepted for publication June 16, 1997.

Address for correspondence and reprints: Dr. Nadine Andrieu, Unité de Recherche en Epidémiologie des Cancers, Institut de la Santé et de la Recherche Médicale, Institut Gustave-Roussy, 94805 Villejuif Cedex, France.

0002-9297/97/6103-0027$02.00

Introduction

During the past 10 years, considerable progress has been made in the characterization of the genetic component of breast cancer (BC). For many years, a family history of breast cancer was consistently reported to be one of the most important risk factors for the disease (for a review, see Kelsey and Horm-Ross 1993). The risk of BC in relatives was found to vary with age at diagnosis of BC (Claus et al. 1990, 1991; Mettlin et al. 1990), the number of affected relatives, and the unilaterality versus bilaterality of the tumor (Ottman et al. 1986). Segregation analyses of large population-based family samples have shown that familial aggregation of BC could be due to the transmission of a dominant gene with a high lifetime penetrance, accounting for a minority (5%–10%) of cases, with the remaining cases occurring sporadically (Williams and Anderson 1984; Newman et al. 1988; Claus et al. 1991; Iselius et al. 1991). However, more-complex mechanisms have also been suggested, and several family studies have indicated genetic heterogeneity of BC, according to clinical and/or epidemiological characteristics of the probands, the histologic type of the tumor, or the presence, among family members, of cancers other than BC (Demenais et al. 1986; Gilligan and Borecki 1986; Goldstein et al. 1987, 1988; Andrieu et al. 1988; Goldstein and Amos 1990). Linkage analyses of multiple breast and breast-ovarian cancer families led to the localization of a first BC gene, BRCA1, on 17q21 (Hall et al. 1990; Narod et al. 1991). Pooled data on 214 families collected worldwide confirmed the genetic heterogeneity of BC, with BRCA1 being found in 45% of BC families and in $\geq$76% of breast-ovarian cancer families (Easton et al. 1993; Narod et al. 1995*a*). A second BC gene (BRCA2), mapped to chromosome 13q12-13 (Wooster et al. 1994), was found to be mainly responsible for BC alone and for male BC (J. Feunteun and G. M. Lenoir, personal communication). Other genes are likely to be involved, with possibly a third BC gene on chromosome 8q (Sobol et al. 1994; Kerangueven et al. 1995). The identification of the BRCA1 and BRCA2 genes (Miki et al. 1994; Wooster et al. 1995) has led to an extensive search for mutations worldwide. Estimates of the age-specific risks of breast and ovarian cancers in women carrying BRCA1-linked markers have shown evidence of heterogeneity of risk

between families (Easton et al. 1995). Moreover, within a family, there may be major variations in the expression of the BRCA1 mutation (Goldgar et al. 1994). Rare alleles on the HRAS1 VNTR locus have recently been found to modify the risk of ovarian cancer in women carrying BRCA1 mutations (Phelan et al. 1996). These observations suggest that other genetic and nongenetic factors may play a role in BC development.

Besides genetic factors, many other risk factors for BC have been reported (for a review, see Kelsey and Horm-Ross 1993). Among them, reproductive variables are well-established risk factors, including an early age at menarche, a late age at menopause, a late age at first full-term pregnancy, and nulliparity when a BC is diagnosed at age >40 years. Reproductive factors such as spontaneous and induced abortions, certain characteristics of the menstrual cycle, and infertility are still controversial (Brind et al. 1996; Michels and Willett 1996; Rookus and van Leeuven 1996; Weed and Kramer 1996; Michels-Blanck et al. 1996; Wu et al. 1996; Hartge 1997; Melbye et al. 1997). Overall relative risks associated with reproductive factors are $\sim \leqslant 2.0$, and mechanisms underlying their effects are still obscure. The difficulty in detecting relevant risk factors and in understanding their role in the etiology of BC may be due to the heterogeneity of the population of cases studied. Case-control studies have found that reproductive factors have a different effect on the occurrence of BC, according to the presence or absence of a family history of BC (Byrne et al. 1991; Parazzini et al. 1992; Sellers et al. 1992, 1993; Andrieu et al. 1993, 1995; Colditz et al. 1993, 1996). A study of 333 North American carriers of BRCA1-linked markers has suggested that reproductive factors may modify BC risk (Narod et al. 1995*b*).

To search for interactions between genetic and reproductive factors in the etiology of BC, we conducted a systematic family study of BC in two French cancer hospitals, which led to the collection of 288 families selected through 288 BC probands (IGRC data). Before conducting segregation analysis, we first identified risk factors associated with BC in this sample, by comparing the BC probands to two types of controls: (1) blood-related controls, the unaffected sisters of the probands, and (2) unrelated hospital controls (Andrieu and Demenais 1994). With either control group, the estimates of BC risk associated with reproductive factors were similar to those commonly reported, with the exceptions of age at menarche and number of abortions. With increasing age at menarche, the risk of BC increased when sister controls were used but decreased when hospital controls were used. Two or more abortions increased BC risk to a higher extent with sister controls than with hospital controls. Such differences in risk could signify interactions between genetic susceptibility and these two reproductive factors, as recently shown theoretically (Andrieu and Goldstein 1996). The goal of the present paper is to estimate the role of genetic and reproductive factors in BC causation and to test for interactions between these factors, by segregation analysis of the IGRC data.

Subjects and Methods

Recruitment of Families and Data Collection

From December 1987 to January 1990, a systematic family study was conducted in two French hospitals (Institut Gustave Roussy [IGR], in Villejuif, and Institut Curie [IC], in Paris). Eligible probands were defined as Caucasians living in France who had a recently diagnosed and histologically confirmed BC. These probands were asked to participate in the study, during the first follow-up visits after surgery. Clinical and histological data on the probands were obtained from medical records. Family data, collected from these patients, on their first-degree (parents and siblings) and second-degree (uncles, aunts, and grandparents) relatives, included demographic characteristics (gender, date of birth, and, if deceased, age at death and cause of death) and the occurrence of BC and any other cancer, along with the age at diagnosis. Epidemiological data were obtained from the probands and their female relatives who were given a questionnaire via the probands. The recorded reproductive factors were age at menarche, length of the menstrual cycle, age at first pregnancy, number of children, number of abortions (no differentiation between induced and spontaneous), and menopausal characteristics (menopausal status, age, and cause of menopause). The questionnaire also included information on cancer occurrence, with age at diagnosis and places of medical care.

Three hundred eighty-five patients were contacted, and 288 of them were recruited into the study. Eighty-three patients refused to participate for the following reasons: unknown information on their family, unknown information on the affection status of their relatives, or refusal to contact their relatives. Thirteen were excluded because insufficient information was supplied, and one was excluded because BC could not be verified. The 288 enrolled probands included 174 cases from IC and 114 from IGR, with an age range at diagnosis being 20–80 years (mean age 51.4 ± 10.0 years). All information was gathered <2 years after diagnosis, for IC probands, and <6 years after diagnosis, for IGR probands. Comparisons of IGR and IC probands did not show any difference for mean age at diagnosis, histological type of BC, stage and inflammatory status of the tumor, number of relatives, family history of BC, and distribution of reproductive factors. Questionnaires distributed to the relatives were returned by <30% of second-degree relatives and by 95% of first-degree relatives. Analyses

were thus restricted to first-degree relatives. Information on a few dead first-degree relatives was obtained by interview of the probands and contact persons in the family. More than 50% of breast malignancies reported in first-degree relatives could be confirmed by pathological records, and, in all of those but one, there was a complete agreement between the case report and the pathological record. Note that information on BC occurrence was provided by both the proband and the affected first-degree relative and that BC has been found to be reported with great accuracy (concordance rate 99%, between case report and pathological record; Theis et al. 1994). Information on reproductive factors was obtained in ⩾80% of female first-degree relatives, except for age at menarche, which was known for 55% of mothers and 78% of sisters. The proportion of missing data on reproductive factors among relatives was similar in the IC and IGR data. Segregation analysis was thus performed on the pooled set of 288 nuclear families, including the probands, mothers, and sisters, with males considered as unknown.

Our previous case-control study (Andrieu and Demenais 1994), using probands of the present study as cases and two sets of controls—probands'unaffected sisters and unrelated hospital controls—indicated that the most relevant risk factors were age at menarche, number of children, number of abortions, and menopausal status (table 1). Age at menarche and number of abortions led to different odds ratios, according to the set of controls used, suggesting interactions between these factors and genetic/familial factors. A high number of children (three or more) had a significant protective effect on BC, and menopausal status was taken into account because it is often described as a possible confounder. In order to limit to a reasonable number the parameters that were to be estimated in segregation analysis, we dichotomized all covariates, by using as the risk category the category leading to the highest or smallest odds ratio and by pooling the others in the baseline category. On the basis of data in table 1, the baseline and at-risk categories for each covariate were defined as follows: age at menarche, <15 years of age versus ⩾15 years of age; number of children, fewer than three versus three or more; number of abortions, less than two versus two or more; and menopausal status, premenopausal versus postmenopausal.

Table 1

Adjusted Odds Ratios of BC Associated with Reproductive Factors, with Use of Unaffected Sister Controls and Hospital Controls

Reproductive Factors	Odds Ratio[a] (95% Confidence Interval)	
	Sister Control	Hospital Controls
Age at menarche:		
<12 years	1	1
13–14 years	1.7 (.9–3.4)	1.1 (.7–1.6)
⩾15 years	2.6 (.9–7.6)[b]	.8 (.5–1.4)[c]
No. of children:		
None	1	1
One or two	.8 (.4–1.6)	.9 (.6–1.5)
Three or more	.5 (.2–1.0)[d]	.4 (.2–.6)[e]
No. of abortions (induced or spontaneous):		
None	1	1
One	1.6 (.8–3.4)	1.0 (.6–1.6)
Two or more	2.1 (.9–5.0)[f]	1.4 (.9–2.2)[g]
Menopausal status:		
Premenopausal	1	1
Postmenopausal	.7 (.3–1.5)	.4 (.2–.6)

[a] Adjusted on age at interview.
[b] Dichotomized results, for age at menarche ⩾15 years/⩽14 years, were 1.70 (.79–3.65).
[c] Dichotomized results, for age at menarche ⩾15 years/⩽14 years, were .80 (.48–1.33).
[d] Dichotomized results, for number of children three or more/two or fewer, were .50 (.30–.84).
[e] Dichotomized results, for number of children three or more/two or fewer, were .38 (.35–.58).
[f] Dichotomized results, for number of abortions two or more/one or fewer, were 2.60 (1.27–5.31).
[g] Dichotomized results, for number of abortions two or more/one or fewer, were 1.40 (.89–2.20).

Methods

Regressive models.—Segregation analysis of BC was conducted by use of the class D regressive logistic model (Bonney 1986) extended to allow for variable age at onset of disease (Abel and Bonney 1990). The regressive models are constructed by specification of a regression relationship between each person's phenotype (affected/unaffected with BC) and a set of explanatory variables, including the person's major genotype, the phenotype of older relatives (to take into account residual family dependences [FD] of unspecified origin [genetic and/or environmental]), and measured covariates. Abel and Bonney (1990) introduced survival-analysis concepts into the regressive models to model age-dependent penetrance functions. Age at onset is considered as a failure time, and age at examination for unaffected subjects is considered as a censored failure time, where the measurement scale is age. The period of follow-up (taken, for BC, as the period from age 20 years [provided that BC risk is negligible at age <20 years] to either age at onset [for affected women], age at examination [for unaffected women], or age at death [for deceased subjects]) is partitioned into K mutually exclusive intervals. The conditional probability that a woman will be affected within the kth interval if she is not affected before is the hazard function, defined as $\lambda(k)$. From the hazard function

are derived $f(k)$, the probability of being affected at an age at onset included in the kth interval,

$$f(k) = \lambda(k) \prod_{h=1}^{k-1} [1 - \lambda(h)] ;$$

$S(k)$, the probability of being unaffected at an age at examination included in the kth interval,

$$S(k) = \prod_{h=1}^{k} [1 - \lambda(h)] ;$$

and $F(k)$, the probability of being affected at an age at examination included in the kth interval when age at onset is unknown $F(k) = 1 - S(k)$. The quantities $f(k)$, $S(k)$, and $F(k)$ are the penetrance functions used in the likelihood formulation explained below. The hazard function is modeled like a logistic function in which the explanatory variables include the major gene (MG) effect, the antecedents' phenotypes, and the measured covariates that can be time dependent. These different effects can be estimated simultaneously or one at a time and can be tested successively, as in classical regression analyses. If we let g_i be the genotype of the ith individual, Y_j be the phenotype of the jth antecedent of i, and $X_i(k)$ be the vector of covariates of i within the kth interval, then the hazard function for the ith individual in the kth interval is $\lambda(k) = \exp[\theta_i(k)]/(1 + \exp[\theta_i(k)])$, where $\theta_i(k)$, the logit of the hazard function, is

$$\theta_i(k) = \alpha_{g_i} + \sum_{j=1}^{i-1} \Gamma_{ji} Y_{ji} + \beta_{g_i} X_i(k) ,$$

where (1) α_g is the genotype-specific baseline parameter (g = AA, Aa, or aa, for a diallelic autosomal locus); (2) Γ_j is a vector of regression coefficients on j antecedents' phenotypes of the ith person (i.e., mother and sisters); and (3) β_g is a vector of genotype-specific regression coefficients of covariates (here including age and reproductive factors). The antecedents' phenotypes were coded as proposed by Demenais (1991), by use of two dummy variables, so that Y_j is a column vector, where $Y_j = (1\ 0)'$ (the prime denotes "transpose") if j is affected by age at examination, $Y_j = (0\ 1)'$ if j is unaffected by age at examination, and $Y_j = (0\ 0)'$ if j has unknown affection status. Each Γ_{ji} parameter is a vector of two coefficients (γ_{j1} and γ_{j2}), so that the logit is modified by γ_{j1} if the antecedent j of i is affected, is modified by γ_{j2} if the antecedent is unaffected, and remains unchanged if j has an unknown affection status. The class D model specifies four types of FD of the ith person on his on her antecedents: spouse (Γ_S), father (Γ_F), mother (Γ_M), and preceding siblings (Γ_C). These dependences were reduced to two—the mother-daughter dependence (Γ_M) and the sister-sister dependence (Γ_C). Since there was no sister with unknown affection status, the γ_{C2} parameter was not needed. Interpretations of these γ parameters, which are not directly interpretable in terms of odds ratios, can be found in the study by Abel et al. (1993). With respect to the covariates, age was the only time-dependent covariate, whereas the effects of the others were assumed to be constant over time, since chronological information on reproductive variables was not recorded. The length of the interval in the hazard function was 1 year, and different functions of age were considered—polynomial of age and logarithm of age. The logarithm function was found to fit the data better and was subsequently used. Let us recall that, since BC risk is negligible at age <20 years, age was taken as (age − 20 years). For example, if we consider a susceptible woman with genotype Aa who has an unaffected mother, one affected sister, and the following reproductive characteristics (age at menarche <15 years of age, more than three children, no abortion, and premenopausal), the logit of her hazard function at age 40, that is within her 21st year of follow-up is $\theta(21) = \alpha_{Aa} + \gamma_{M2} + \gamma_{C1} + \beta_{Aa}^{age} \times \mathrm{Ln}(21) + \beta_{Aa}^{children}$.

Likelihood formulation.—The likelihood of a family of n individuals can be written as

$$\mathrm{L(family)} = \prod_{i=1}^{n} \sum_{g} P_{g_i} \mathrm{L}_i(g_i, Y_{ji}, X_i) ,$$

where P_{gi} is the probability of the unobserved genotypes at the major locus, $\mathrm{L}_i(g_i, Y_{ji}, X_i)$ is the penetrance function of the ith individual, given g_i, Y_{ji}, X_i, and the sum is over all possible genotypes. The probability P_{gi} for individuals with no parents in the nuclear family is a function of the allelic frequency of the deleterious allele A, q, under the assumption of Hardy-Weinberg equilibrium. For children with parents in the family, P_{gi} for a given parental mating type is specified by Mendelian transmission probabilities and Elston and Stewart's (1971) general transmission probabilities. These transmission parameters, denoted as τ_{AAA}, τ_{AaA}, and τ_{aaA}, are the conditional probabilities of transmitting to offspring allele A, for parental genotypes AA, Aa, and aa, respectively. They are equal to 1, .5, and 0 according to the Mendelian hypothesis, whereas under the general transmission model they can take any value in the range 0–1. The penetrance function $\mathrm{L}_i(g_i, Y_{ji}, X_i)$ is equal to $f(k; g_i, Y_{ji}, X_i)$ if i is affected in interval k, $S(k; g_i, Y_{ji}, X_i)$ if i is unaffected in interval k, or $1 - S(k; g_i, Y_{ji}, X_i)$ if i is affected in interval k and age at onset is unknown. The penetrance is equal to 1 if the affection status is unknown. Under the general class D regressive model, the likelihood is a function of the following parameters: allele A frequency (q), three genotype-specific baseline

parameters (α_{AA}, α_{Aa}, and α_{aa}), three transmission probabilities (τ_{AAA}, τ_{AaA}, and τ_{aaA}), three parameters specifying mother-daughter (γ_{M1} and γ_{M2}) and sister-sister (γ_{C1}) dependences, and β_g (g = AA, Aa, and aa) regression coefficients for covariates including age and reproductive factors.

The likelihood function was corrected for the ascertainment bias by use of the approach proposed by Elston and Sobel (1979). Given the ascertainment scheme for a nuclear family through one BC offspring, the probability π of an affected woman being a proband was set at .01.

Hypothesis testing.—Parameter estimates and tests of hypotheses were performed by use of maximum-likelihood methods. The likelihood of the IGRC data was maximized under different models, always including covariates effects and, according to the hypotheses tested, an MG effect and/or residual FD. The first class of models (I) is a sporadic model with no FD and no MG effect; the second class of models (II) includes FD but no MG effect; the third class of models (III) is an MG model without residual FD; and the fourth class of models (IV) includes both an MG effect and residual FD. Two additional models, including a major factor and residual FD, are used to test transmission of this major effect: (1) a model with no parent-offspring transmission of the major factor, in which the three transmission probabilities are equal (model V), and (2) the general transmission model, in which the three transmission probabilities are estimated (model VI). Nested models were compared by use of likelihood-ratio tests. Segregation of the MG can be inferred if three consecutive tests lead to the following conclusions: (1) rejection of model II as compared with model IV; (2) failure to reject the Mendelian transmission hypothesis when compared with the general transmission model (i.e., model IV vs. model VI); and (3) rejection of the hypothesis of no transmission of the major factor when compared with the general model (i.e., model V vs. model VI). Gene-covariate interactions were tested within model IV (or model III) by comparison of submodels in which the β_g's were set equal to the same estimate of β, whatever g (no interaction), versus models in which three (or two) β_g's were estimated (interaction). Likelihood computations were performed with the REGRESS computer program (Bonney et al. 1988; Demenais and Lathrop 1994), which incorporates the regressive approach in the ILINK program of the LINKAGE package (Lathrop and Lalouel 1984) and uses the GEMINI optimization routine (Lalouel and Yee 1980).

Strategy of analysis.—Two types of segregation analyses were conducted, by (1) considering only age, as a time-dependent covariate, and (2) considering age, as a time-dependent covariate, plus the four dichotomized reproductive factors (age at menarche, number of children, number of abortions, and menopausal status). In the latter case, two different approaches were used to deal with missing covariates. The first one excludes subjects with missing covariates (the "complete-subject method," or "CS method"). The second one (the "missing-indicator method," or "MI method") creates two dummy variables for each covariate: a missing-value indicator, which is equal to 1 for missing and 0 for a known value, and a second variable, which is equal to 1 for the exposed subjects and 0 for the others (nonexposed subjects and subjects with a missing value).

Results

The results of segregation analyses are presented in tables 2 and 3, in which reproductive factors are ignored and taken into account, respectively.

Segregation Analysis of BC When the Effects of Reproductive Factors Are Ignored

There is strong evidence for FD (model II-5 vs. model I; $\chi^2_3 = 63.0$, $P < 10^{-5}$). A model including both mother-daughter and sister-sister dependences fits significantly better than a model with either sister-sister dependence only (II-2 vs. II-5; $\chi^2_2 = 12.5$, $P = .002$) or mother-daughter dependence only (II-1 vs. II-5; $\chi^2_1 = 20.0$, $P < 10^{-5}$). Moreover, a model assuming no change in risk when the mother is unaffected ($\gamma_{M2} = 0$) fits the data (II-4 vs. II-5; $\chi^2_1 = 1.8$), whereas a model assuming an equal change in risk regardless of whether the mother is affected or unaffected ($\gamma_{M1} = \gamma_{M2}$) is rejected (II-3 vs. II-5; $\chi^2_1 = 11.7$, $P = .0006$). Thus, the model that best fits FD includes the γ_{M1} parameter estimated at 1.63 ± 0.51 and the γ_{C1} parameter estimated at 2.96 ± 0.34. A dominant-gene effect was detected (II-4 vs. IV-2; $\chi^2_2 = 30.2$, $P < 10^{-5}$) in the presence of FD. A dominant mode of inheritance for this gene fitted as well as a more general codominant model, whereas a recessive mode of inheritance was rejected ($P < .02$) (results not shown). In addition to this major effect, a residual sister-sister dependence was significant (III vs. IV-2; $\chi^2_{1-2} = 22.9$, $P < 10^{-5}$), whereas the mother-daughter dependence converged to 0. An interaction between this major effect and age was not significant (IV-2 vs. IV-3; $\chi^2_1 = 1.8$). When compared with the general transmission–probability model, Mendelian transmission of this major effect fitted the data (IV-2 vs. VI; $\chi^2_3 = 1.4$), whereas the hypothesis of no parent-offspring transmission of the major effect was rejected (V vs. VI; $\chi^2_2 = 26.0$, $P < 10^{-5}$). Thus, familial transmission of BC can be accounted for by segregation of a dominant gene plus residual sister-sister dependence. The estimated frequency of the deleterious allele is .0006. The cumulative BC risk for genetically susceptible women is .25 by age 55 years and .98 by age 75 years. The proportion of nonsusceptible women

Table 2

Segregation Analysis of BC, by Use of Class D Regressive Model When Reproductive Factors Are Ignored

Model	q	α_{Aa} [a]	α_{aa}	γ_{M_2}	γ_{M_1}	γ_{C_1}	β_{Aa}^{age} [b]	β_{aa}^{age}	τ_{AAA}	τ_{AaA}	τ_{aaA}	$-2LnL+c$ [c]
I [Sporadic]	[0]	−20.76	[=α_{Aa}]	[0]	[0]	[0]	4.07	[=β_{Aa}^{age}]	...	...	...	92.8
II [no MG, FD]:												
1. γ_{M_2}, γ_{M_1}, $\gamma_{C_1} = 0$	[0]	−24.33	[=α_{Aa}]	2.70	4.55	[0]	4.52	[=β_{Aa}^{age}]	...	...	...	49.8
2. $\gamma_{M_2} = \gamma_{M_1} = 0$, γ_{C_1}	[0]	−23.02	[=α_{Aa}]	[0]	[0]	2.10	4.53	[=β_{Aa}^{age}]	...	...	...	42.3
3. $\gamma_{M_2} = \gamma_{M_1}$, γ_{C_1}	[0]	−22.91	[=α_{Aa}]	−.49	[=γ_{M_2}]	2.40	4.51	[=β_{Aa}^{age}]	...	...	...	41.5
4. $\gamma_{M_2} = 0$, γ_{M_1}, γ_{C_1}	[0]	−24.44	[=α_{Aa}]	[0]	1.63	2.96	4.52	[=β_{Aa}^{age}]	...	...	...	31.6
5. γ_{M_2}, γ_{M_1}, γ_{C_1}	[0]	−25.28	[=α_{Aa}]	1.18	3.37	2.59	4.57	[=β_{Aa}^{age}]	...	...	...	29.8
III [dominant MG, no FD]	.0006	−20.95	−26.86	[0]	[0]	[0]	5.07	[=β_{Aa}^{age}]	[1]	[.5]	[0]	24.3
IV [dominant MG + FD]:												
1. $\gamma_{M_2} = \gamma_{M_1} = 0$, γ_{C_1}	.0006	−21.91	−26.97	[0]	[0]	2.33	5.31	[=β_{Aa}^{age}]	[1]	[.5]	[0]	1.4
2. $\gamma_{M_2} = 0$, γ_{M_1}, γ_{C_1}	.0006	−21.90	−26.97	[0]	→.0	2.33	5.31	[=β_{Aa}^{age}]	[1]	[.5]	[0]	1.4
3. MG × age, γ_{C_1}	.0006	−19.56	−28.11	[0]	[0]	2.30	4.56	5.58	[1]	[.5]	[0]	−.4
V [nontransmitted dominant major effect + FD]	.0006	−16.40	−24.00	[0]	[0]	2.11	4.81	[=β_{Aa}^{age}]	→0.0	[=τ_{AAA}]	[=τ_{AAA}]	26.0
VI [general transmission of dominant major effect + FD]	.0006	−22.10	−27.41	[0]	[0]	2.42	5.40	[=β_{Aa}^{age}]	.70	.40	.0	.0

NOTE.—Parameters in square brackets were fixed at the value indicated.

[a] $\alpha_{AA} = \alpha_{Aa}$.

[b] $\beta_{AA}^{age} = \beta_{Aa}^{age}$.

[c] $c = -2{,}231.1$.

among affecteds (phenocopies) reaches 98% by age 75 years.

Segregation Analysis of BC When Reproductive Factors Are Taken into Account

The results presented in table 3 correspond to the CS strategy for missing covariates—that is, the person's affection status is coded as unknown if a covariate is unknown. These results are similar to those previously found. Tests of the different patterns of FD showed that the best-fitting model included a change in risk when a mother is affected (γ_{M1}) and a change in risk when a sister is affected (γ_{C1}), with γ_{M1} estimated at 1.79 ± 0.74 and γ_{C1} estimated at 3.52 ± 0.61. This model is significant compared with the sporadic model (I vs. II; $\chi_2^2 = 39.9$, $P < 10^{-5}$). Again, when an MG effect was included in the model, the mother-daughter dependence converged to 0. There was significant evidence of a dominant effect (II vs. IV-1; $\chi_2^2 = 15.1$, $P = .0005$) plus residual sister-sister dependence (III vs. IV-1; $\chi_{1-2}^2 = 14.3$, $P = .0008$). Interaction between age and the dominant major effect became significant in this analysis (IV-1 vs. IV-2; $\chi_1^2 = 5.9$, $P = .015$). The effects of reproductive factors were globally significant under the model of FD, both with no MG ($\chi_4^2 = 18.9$, $P = .0008$) and when the MG effect was included ($\chi_4^2 = 17.5$, $P = .0015$) (results not shown). Interactions between the dominant major effect and each of the reproductive factors were not significant, except for the number of children (IV-2 vs. IV-4; $\chi_1^2 = 5.7$, $P = .017$). However, whereas estimates of regression coefficients (β) in susceptible and nonsusceptible women were similar for number of abortions and menopausal status, they differed for age at menarche, although not signicantly ($\beta_{Aa} = 1.75 \pm 0.68$ and $\beta_{aa} = .36 \pm .48$, for an age at menarche $\geqslant$15 years). As before, with respect to the general transmission model, Mendelian transmission of the dominant major fitted the data well (IV-4 vs. VI; $\chi_3^2 = 1.3$), and the hypothesis of no parent-offspring transmission was rejected (V vs. VI; $\chi_2^2 = 29.1$, $P < 10^{-5}$). When the MI method for missing covariates was used, the conclusions of segregation analysis were similar, except for the interaction between age and MG effect, which was not significant (results not shown).

Comparison of Segregation Analyses

Thus, the general conclusion was the same when reproductive factors were taken into account and when they were ignored: there was evidence for segregation of a dominant gene plus sister-sister dependence. Moreover, analyses repeated by including reproductive factors one at a time in the model led to conclusions similar to those obtained when all of them were considered together in the model. The estimate of the deleterious-allele frequency was similar in all analyses ($q = .0006$). However, inclusion of reproductive factors led to a vari-

Table 3

Segregation Analysis of BC, by Use of Class D Model When Reproductive Factors Are Taken into Account

Model	α_{Aa}[a]	α_{aa}	γ_{M_1}	γ_{C_1}	β_{Aa}^{age}[b]	β_{aa}^{age}	Age at Menarche: ≥15 Years vs. ≤14 Years		No. of Children: Three or More vs. Two or Fewer		No. of Abortions: Two or More vs. One or Fewer		Menopausal Status: Post- vs. Pre-		τ_{AAA}	τ_{AaA}	τ_{aaA}	$-2LnL+c$[c]
							$\beta_{Aa}^{menarche}$[b]	$\beta_{aa}^{menarche}$	$\beta_{Aa}^{children}$[b]	$\beta_{aa}^{children}$	$\beta_{Aa}^{abortion}$[b]	$\beta_{aa}^{abortion}$	$\beta_{Aa}^{meno.}$[b]	$\beta_{aa}^{meno.}$				
I [sporadic]	−24.01	[=α_{Aa}]	[0]	[0]	5.15	[=β_{Aa}^{age}]	.515	[=$\beta_{Aa}^{menarche}$]	−.975	[=$\beta_{Aa}^{children}$]	.504	[=$\beta_{Aa}^{abortion}$]	−.517	[=$\beta_{Aa}^{meno.}$]	. . .	. . .	. . .	67.9
II [FD], $\gamma_{M_2}=0$, γ_{M_1}, γ_{C_1}	−28.09	[=α_{Aa}]	1.79	3.52	5.48	[=β_{Aa}^{age}]	.638	[=$\beta_{Aa}^{menarche}$]	−1.088	[=$\beta_{Aa}^{children}$]	.685	[=$\beta_{Aa}^{abortion}$]	.087	[=$\beta_{Aa}^{meno.}$]	. . .	. . .	. . .	28.0
III [dominant MG, no FD]	−23.78	−29.73	[0]	[0]	5.95	[=β_{Aa}^{age}]	1.238	[=$\beta_{Aa}^{menarche}$]	−1.402	[=$\beta_{Aa}^{children}$]	.308	[=$\beta_{Aa}^{abortion}$]	−.089	[=$\beta_{Aa}^{meno.}$]	[1]	[.5]	[0]	27.2
IV [dominant MG + FD]:																		
1. No interaction between MG and age, γ_{M_1}, γ_{C_1}	−24.01	−28.97	→0	2.41	5.95	[=β_{Aa}^{age}]	.661	[=$\beta_{Aa}^{menarche}$]	−1.182	[=$\beta_{Aa}^{children}$]	.540	[=$\beta_{Aa}^{abortion}$]	.349	[=$\beta_{Aa}^{meno.}$]	[1]	[.5]	[0]	12.9
2. Interaction between MG and age, $\gamma_{M_1}=0$, γ_{C_1}	−18.92	−32.06	[0]	2.57	4.21	6.64	.684	[=$\beta_{Aa}^{menarche}$]	−1.115	[=$\beta_{Aa}^{children}$]	.580	[=$\beta_{Aa}^{abortion}$]	.230	[=$\beta_{Aa}^{meno.}$]	[1]	[.5]	[0]	7.0
3. Interaction between MG and menarche, γ_{C_1}	−19.79	−31.78	[0]	2.28	4.42	6.57	1.75	.36	−1.129	[=$\beta_{Aa}^{children}$]	.598	[=$\beta_{Aa}^{abortion}$]	.385	[=$\beta_{Aa}^{meno.}$]	[1]	[.5]	[0]	4.2
4. Interaction between MG and children, γ_{C_1}	−18.78	−31.62	[0]	2.79	4.04	6.54	.722	[=$\beta_{Aa}^{menarche}$]	.05	−1.5	.672	[=$\beta_{Aa}^{abortion}$]	.263	[=$\beta_{Aa}^{meno.}$]	[1]	[.5]	[0]	1.3
5. Interaction between MG and abortion, γ_{C_1}	−18.92	−32.05	[0]	2.58	4.20	6.63	.684	[=$\beta_{AA}^{menarche}$]	−1.111	[=$\beta_{Aa}^{children}$]	.63	.56	.233	[=$\beta_{Aa}^{meno.}$]	[1]	[.5]	[0]	7.0
6. Interaction between MG and menopause, γ_{C_1}	−18.75	−32.13	[0]	2.56	4.13	6.68	.696	[=$\beta_{Aa}^{menarche}$]	−1.111	[=$\beta_{Aa}^{children}$]	.584	[=$\beta_{Aa}^{abortion}$]	.35	.14	[1]	[.5]	[0]	7.0
V [no transmitted dominant major effect + FD]	−27.27	−27.35	[0]	2.49	5.72	7.59	.622	[=$\beta_{Aa}^{menarche}$]	.91	−1.2	.761	[=$\beta_{Aa}^{abortion}$]	−.106	[=$\beta_{Aa}^{meno.}$]	→.0	[=τ_{AAA}]		29.1
VI [transmission of dominant major effect + FD]	−19.42	−31.34	[0]	2.97	4.37	6.48	.752	[=$\beta_{Aa}^{menarche}$]	.26	−1.6	.566	[=$\beta_{Aa}^{abortion}$]	.293	[=$\beta_{Aa}^{meno.}$]	.9	.3	.0	0

NOTE.—Parameters in square brackets were fixed at the value indicated.

[a] $\alpha_{AA}=\alpha_{Aa}$.

[b] $\beta_{AA}^{x}=\beta_{Aa}^{x}$.

[c] $c=-1{,}822.5$.

ation in BC risk with age, a BC risk that differed significantly between susceptible and nonsusceptible women ($\beta_{Aa}^{age} = 4.21 \pm 0.49$ and $\beta_{aa}^{age} = 6.64 \pm 0.72$, respectively). The ratio of the hazard functions, calculated in susceptible and nonsusceptible women, was higher at younger ages than at older ages (this ratio was 124 at 45 years of age, 48 at 55 years, and 20 at 65 years), showing that the gene has a greater effect in younger women. Calculations of the lowest BC cumulative risks (values of dichotomous reproductive covariates set at 0 for those increasing BC risk and set at 1 for those decreasing BC risk) and highest BC cumulative risks (values of dichotomous reproductive covariates set at 0 for those decreasing BC risk and set at 1 for those increasing BC risk) indicate that the risk of developing BC in susceptible women increases from .03 by age 40 years to .56 by age 80, in the lowest-risk group, and from .02 to .99, in the highest-risk group. This increase in risk, by the same ages, in nonsusceptible women varies from 7.10^{-6} to .03, in the lowest-risk group, and from 5.10^{-5} to .16, in the highest-risk group (fig. 1).

Figure 2*A–D* illustrates the effect of each reproductive factor on BC risk, with and without an interaction between the BC gene and the reproductive factor, in susceptible and nonsusceptible women. Cumulative risks were similar for menopausal status and number of abortions (fig. 2*A* and *B*), whether an interaction with the BC gene was taken into account or was ignored. The hazard function was multiplied by 1.3 in postmenopausal women, compared with premenopausal women; and it was multiplied by 1.8 in women who had undergone two or more abortions, compared with those who had had fewer than two abortions. As seen in figure 2*C*, although interaction between age at menarche and the BC gene was not significant, the increase in risk associated with an age at menarche $\geq$15 years was higher in susceptible women with interaction than it was in susceptible women without interaction, whereas it was slightly lower in nonsusceptible women with interaction than in nonsusceptible women without interaction. The multiplicative factor of the hazard function for an age at menarche $\geq$15 years versus $\leq$14 years was 2.0 without interaction, in both susceptible and nonsusceptible women, whereas when the interaction was taken into account it was 5.7 in susceptible women and 1.4 in nonsusceptible women. The change in risk associated with the number of children is shown in figure 2*D*. When the interaction between the number of children and the BC gene was taken into account, the protective effect of having at least three children disappeared in susceptible women: without interaction, the hazard function decreased by a factor of .3 in high-parity women compared with low-parity women, whereas with interaction it was practically unchanged (multiplicative factor 1.05). However, in nonsusceptible women, the protective effect of high parity remained similar with and without interaction, the hazard function being decreased by .2 and .3, respectively, compared with that in low-parity women.

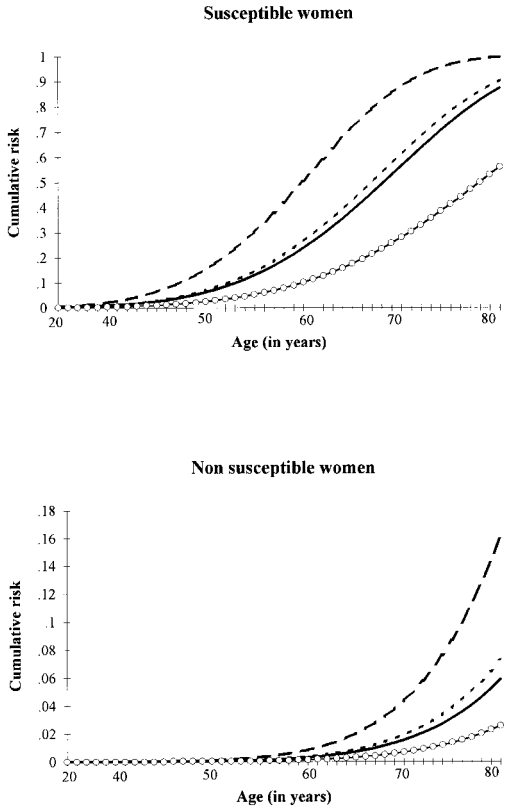
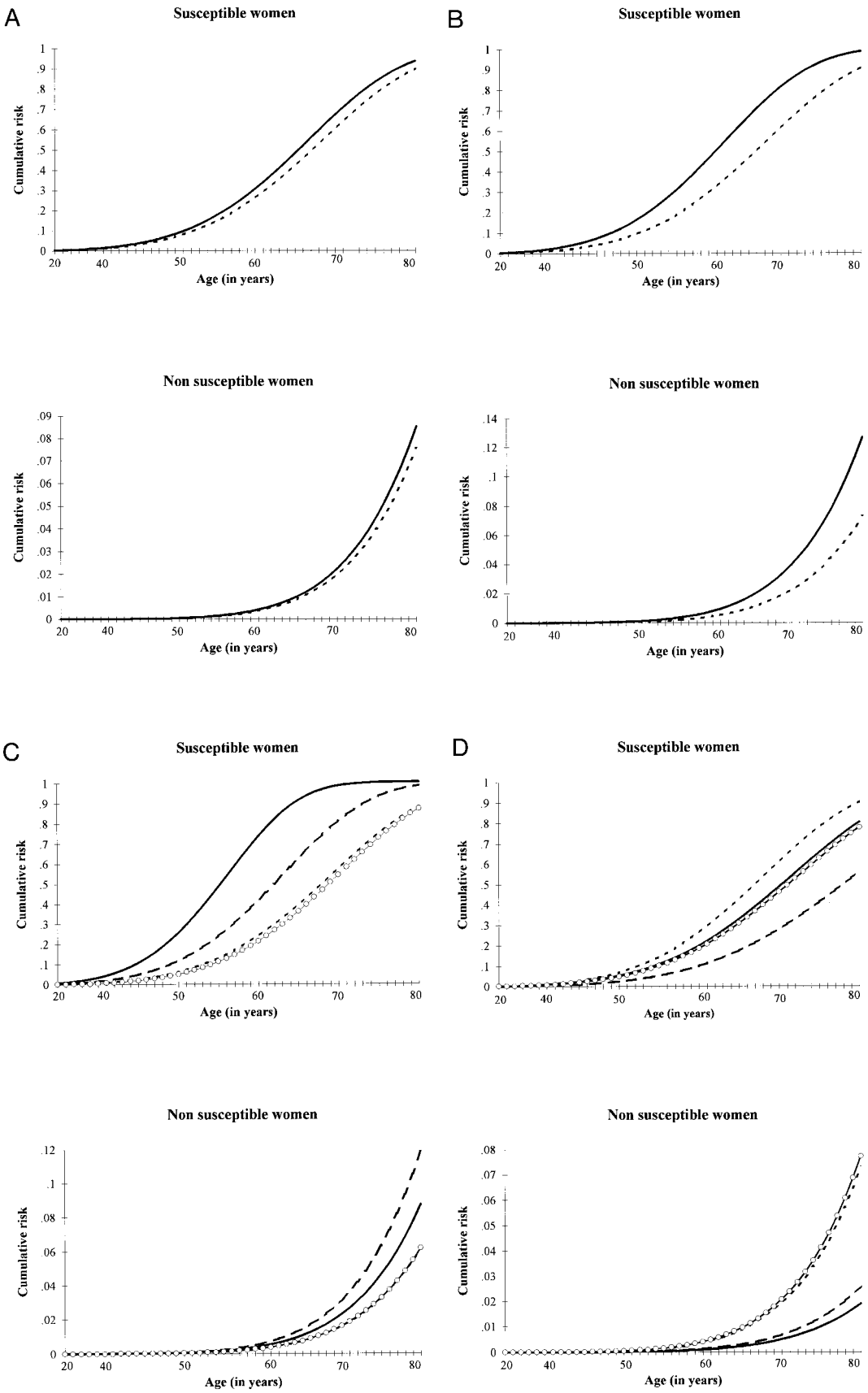


Figure 1 Effect of covariates on BC cumulative risk. Cumulative risks of BC as a function of age in susceptible (i.e., Aa) and nonsusceptible (i.e., aa) women, with and without the effects of reproductive factors. Unbroken lines (——) denote BC risks when reproductive factors are not included in the analysis [$\theta(k) = \alpha_g + \beta^{age}(k)$]; dotted lines (· · ·) denote baseline BC risks when reproductive factors are included [$\theta(k) = \alpha_g + \beta_g^{age}(k)$]; lines with circles (○—○) denote the lowest BC risks when reproductive factors are included (values of reproductive covariates increasing BC risk are set at 0 and those decreasing BC risk are set at 1) [$\theta(k) = \alpha_g + \beta_g^{age}(k) + \beta^{children}$]; and broken lines (– – –) denote the highest BC risk when reproductive factors are included (values of reproductive covariates decreasing BC risk are set at 0 and those increasing BC risk set at 1) [$\theta(k) = \alpha_g + \beta_g^{age}(k) + \beta^{menarche} + \beta^{abortion} + \beta^{menopause}$].

Discussion

Segregation analyses of BC in the IGRC data showed evidence for the segregation of a dominant gene and additional sister-sister dependence, both when reproductive factors were ignored and when they were included in the regressive models. A significant gene-age interaction was detected when reproductive factors were taken into account. These results are in agreement with previous segregation analyses of BC that have been conducted


A
Susceptible women
Cumulative risk
Age (in years)
B
Susceptible women
Cumulative risk
Age (in years)
Non susceptible women
Cumulative risk
Age (in years)
Non susceptible women
Cumulative risk
Age (in years)
C
Susceptible women
Cumulative risk
Age (in years)
D
Susceptible women
Cumulative risk
Age (in years)
Non susceptible women
Cumulative risk
Age (in years)
Non susceptible women
Cumulative risk
Age (in years)

in other populations by use of the mixed model (Lalouel and Morton 1981; Lalouel et al. 1983), except that residual FD modeled by a polygenic component was generally not significant (e.g., see Claus et al. 1991). As was observed in real and simulated data (Abel et al. 1995; Essioux et al. 1995), we found that Mendelian transmission of the major effect fitted the data of our 288 families when regressive models were used, whereas it was rejected when a preliminary analysis was conducted with the unified-mixed model (results not shown). This discrepancy may be due to a difference in the way in which the age of affected women is taken into account by these models: age at onset is taken into account by regressive models, and age at examination is taken into account by mixed models.

In terms of parameter estimation, the susceptibility-allele estimate of .0006 is in the lowest range of the values reported by previous segregation analyses of population-based samples (.0006–.003) and is in good agreement with a recent estimate for BRCA1. The frequency of BRCA1 was estimated at .0006, and it was suggested that the overall frequency of genes with a similar effect (i.e., BRCA1, BRCA2, and other putative genes) might be close to .0008 (Ford et al. 1995). At any rate, increasing the allele frequency by a factor of 10 in our analyses did not change our conclusions. Estimates of cumulative risks in susceptible women from our sample are lower at younger ages (2-fold lower by age 55 years) but slightly higher at older ages (1.2-fold higher by age 75 years) than those obtained in the largest series of 4,730 North American families analyzed by use of the mixed model (Claus et al. 1991). The cumulative incidences predicted by our estimates are lower than those estimated from the French tumor registries (lifetime incidence of 5% vs. 8.8%, respectively) (Benhamou et al. 1990). Differences in risk estimates may be due to (*a*) between-sample differences in age distribution of probands (e.g., Cancer and Steroid Hormone Study probands are 20–54 years of age, and IGRC probands are 20–80 years of age) and to (*b*) the methodology used for estimation. Mixed models assign liability classes to affected and unaffected subjects, according to their age at examination, with a class-specific morbid risk calculated from population data, whereas the regressive models (Abel and Bonney 1990) use survival-analysis concepts and consider age at onset for affected subjects and age at examination for unaffected subjects. Thus, the parameters of mixed models are constrained to fit the observed cumulative incidences of BC in the general population, whereas the parameters of the regressive models are not constrained. Statistical properties of different formulations of the regressive models, with respect to the use of these constraints for different ascertainment schemes, are being investigated. Comparisons of these models also can be found in the work of Demenais et al. (1992) and Abel et al. (1995).

When the covariates were taken into account in the analysis, similar results were obtained, whatever the method (CS or MI) used to deal with missing covariates, except that interaction between the major effect and age was detected only with the CS method. The problems raised by missing data have been widely studied in the statistical literature (Greenland and Finkle 1995). In general, the MI method yields estimates with smaller standard errors than does the CS method, but it is biased when the assumption of random distribution of missing

Figure 2 Effect of menopausal status, number of abortions, age at menarche, and number of children on BC cumulative risk. *A,* Effect of menopausal status on cumulative risks of BC in susceptible (i.e., Aa) and nonsusceptible (i.e., aa) women. Cumulative-risk curves computed under a model with interaction between a BC gene and menopausal status [$\theta(k) = \alpha_g + \beta_g^{age}(k) + \beta_{menopause}^{menopause}$] are superimposed on cumulative-risk curves computed under a model without such interaction [$\theta(k) = \alpha_g + \beta_g^{age}(k) + \beta^{menopause}$]. Dotted lines (· · ·) denote BC risks in premenopausal women (i.e., the baseline category); and unbroken lines (——) denote BC risks in postmenopausal women (i.e., the at-risk category). *B,* Effect of number of abortions on cumulative risks of BC in susceptible (i.e., Aa) and nonsusceptible (i.e., aa) women. Cumulative-risk curves computed under a model with interaction between a BC gene and number of abortions [$\theta(k) = \alpha_g + \beta_g^{age}(k) + \beta_g^{abortion}$] are superimposed on cumulative-risk curves computed under a model without such interaction [$\theta(k) = \alpha_g + \beta_g^{age}(k) + \beta^{abortion}$]. Dotted lines (· · ·) denote BC risks when number of abortions is fewer than two (i.e., the baseline category); and unbroken lines (——) denote BC risks when number of abortions is two or more (i.e., the at-risk category). *C,* Effect of age at menarche on cumulative risks of BC in susceptible (i.e., Aa) and nonsusceptible (i.e., aa) women. Dotted and broken lines denote BC risks computed under a model without interaction between a BC gene and age at menarche [$\theta(k) = \alpha_g + \beta_g^{age}(k) + \beta^{menarche}$]: dotted lines (· · ·) denote BC risks for age at menarche <15 years of age (i.e., the baseline category); and broken lines (– – –) denote BC risks for age at menarche ⩾15 years of age (i.e., the at-risk category). Lines with circles and unbroken lines denote BC risks computed under a model including an interaction between a BC gene and age at menarche [$\theta(k) = \alpha_g + \beta_g^{age}(k) + \beta_g^{menarche}$]: lines with circles (○–○) denote BC risks for age at menarche <15 years of age (i.e., the baseline category); and unbroken lines (——) denote BC risks for age at menarche ⩾15 years of age (i.e., the at-risk category). For nonsusceptible women the line with circles is superimposed on the dotted line. *D,* Effect of number of children on cumulative risks of BC in susceptible (i.e., Aa) and nonsusceptible (i.e., aa) women. Dotted and broken lines denote BC risks computed under a model without interaction between a BC gene and number of children [$\theta(k) = \alpha_g + \beta_g^{age}(k) + \beta^{children}$]: dotted lines (· · ·) denote BC risks when number of children is fewer than three (baseline category); and broken lines (– – –) denote BC risks when number of children is three or more (i.e., the exposed category). Unbroken lines and lines with circles denote BC risks computed under a model including an interaction between a BC gene and number of children [$\theta(k) = \alpha_g + \beta_g^{age}(k) + \beta_g^{children}$]: lines with circles (○–○) denote BC risks when number of children is fewer than three (i.e., the baseline category); and unbroken lines (——) denote BC risks when number of children is three or more (i.e., the exposed category).

covariates is violated, and it requires estimation of twice as many parameters for covariates. We have therefore chosen to show results for the CS method. More-complex methods have been proposed (e.g., Gibbs sampling) to allow for missing data, but they rarely have been applied to family data.

Among the reproductive factors included in segregation analysis, parity was found to interact with the BC-gene effect, and there was an indication of a gene–age at menarche interaction, albeit not a significant one. Whereas the usual protective effect that high parity (at least three children) has on BC risk remained in nonsusceptible women, it disappeared in susceptible women. The increase in the BC risk associated with a late age at menarche ($\geq$15 years of age) was higher in susceptible women than in nonsusceptible women.

A few case-control studies have searched for an interaction, in BC risk, between parity and family history. Five of eight studies found that the effect of parity did not vary according to family history of BC (Bain et al. 1980; Sellers et al. 1992 1993; Andrieu et al. 1993; Colditz et al. 1993). The other three found no protection from multiple births in women with a family history of BC (Negri et al. 1988; Parazzini et al. 1992; Colditz et al. 1996), which is consistent with our finding. Surprisingly, in women carrying BRCA1-linked markers, high parity had an inverse effect on breast versus ovarian cancers, decreasing BC risk and increasing the risk of ovarian cancer (Narod et al. 1995*b*). These BRCA1 carriers come from highly selected pedigrees with at least one ovarian cancer case and multiple BC cases and were analyzed as independent observations.

In epidemiological literature, a late age at menarche has often been described as protective against BC (Kelsey and Horm-Ross 1993). However, among the case-control studies investigating variations in BC risk associated with age at menarche according to a BC family history, three found that the risk associated with a late age at menarche increased for women with a family history of BC and decreased for women without a family history of BC (Bain et al. 1980; Malone and Daling 1992; Parazzini et al. 1992), whereas the others found no decreasing risk with a late age at menarche in women with a family history of BC (Brinton et al. 1982; Negri et al. 1988; Sellers et al. 1993; Colditz et al. 1996). In our previous case-control study (Andrieu and Demenais 1994), a late age at menarche conferred an increased BC risk in sister controls and a decreased risk in unrelated hospital controls, suggesting interaction with genetic factors, as has been demonstrated theoretically (Andrieu and Goldstein 1996).

Altogether, these results underline the complexity of the mechanisms involved in BC etiology. Interactions suggested by case-control studies are more likely to be detected if the genetic factors are common, whereas this segregation analysis shows evidence for interactions with a rare gene.

No obvious biological hypothesis explains the inverse effect of a late age at menarche in women genetically predisposed to BC. Recent studies on the BRCA1 gene and protein functions (Holt et al. 1996; Jensen et al. 1996) have shown that BRCA1 is a selective growth inhibitor of breast and ovarian cells (Holt et al. 1996), and BRCA1 mRNA has been found to be induced during mouse pregnancies (Marquis et al. 1995). As suggested by Jensen et al. (1996), BRCA1 may be mediating the protective effect of pregnancy by inhibiting the proliferation of breast epithelial cells, a function that is lost in genes bearing deleterious mutations. This hypothesis, which would offer an appealing explanation of our finding, merits further investigation.

This is the first study to detect interactions between inherited predisposition to BC and reproductive factors by segregation analysis. These results need to be confirmed by family studies in other populations. Collections of large family and population data of BRCA1 and BRCA2 carriers may also shed light on the biological mechanisms underlying these interactions.

Acknowledgments

This project was supported by INSERM, an Association pour la Recherche contre le Cancer grant, a fellowship from the Ligue Nationale Contre le Cancer, the Union Internationale contre le Cancer, Groupement de Recherches et d'Études sur les Génomes, and the French Ministry of Research. We would like to thank Dr. F. Clavel, for helping to perform this study; the physicians of the Institut Gustave Roussy and the Institut Curie, for allowing their patients to be interviewed; and Lorna Saint-Ange, for linguistic revision of the manuscript.

References

Abel L, Bonney G (1990) A time–dependant logistic hazard function for modeling variable age of onset in analysis of familial diseases. Genet Epidemiol 7:391–407

Abel L, Garcia A, Demenais F (1995) Complex segregation analysis of familial diseases with variable age of onset: comparison of different methods by a simulation study. Genet Epidemiol 12:231–249

Abel L, Golmard J-L, Mallet A (1993) An autologistic model for the genetic analysis of familial binary data. Am J Hum Genet 53:894–907

Andrieu N, Clavel F, Auquier A, Lê MG, Gairard B, Piana L, Brémond A, et al (1993) Variations in the risk of breast cancer associated with a family history of breast cancer according to age at onset and reproductive factors. J Clin Epidemiol 46:973–980

Andrieu N, Demenais F (1994) Role of genetic and reproductive factors in breast cancer. Genet Epidemiol 11:285

Andrieu N, Demenais F, Martinez M (1988) Genetic analysis

of human breast cancer: implications for family study designs. Genet Epidemiol 5:225–233

Andrieu N, Duffy S, Rohan T, Lê MG, Luporsi E, Gerber M, Renaud R, et al (1995) Familial risk, abortion and interactive effect on the risk of breast cancer—a combined analysis of six case–control studies. Br J Cancer 72:744–751

Andrieu N, Goldstein AM (1996) Use of relatives as controls to identify risk factors when an interaction between environmental and genetic factors exists. Int J Epidemiol 25:649–657

Bain C, Speizer FE, Rosner B, Belanger C, Hennehens CH (1980) Family history of breast cancer as a risk indicator for the disease. Am J Epidemiol 111:301–308

Benhamou E, Laplanche A, Wartelle M, Faivre J, Gignoux M, Menegoz F, Robillard J, et al (1990) Incidence des cancers en France in Statistiques de Santé. INSERM, Paris

Bonney GE (1986) Regression logistic models for familial disease and other binary traits. Biometrics 42:611–625

Bonney GE, Lathrop GM, Lalouel J-M (1988) Combined linkage and segregation analysis using regressive models. Am J Hum Genet 43:29–37

Brind J, Chichilli VM, Severs WB, Summy–Long J (1996) Induced abortion as an independant risk factor for breast cancer: a comprehensive review and meta–analysis. J Epidemiol Commun Health 50:481–496

Brinton L, Hoover R, Fraumeni J (1982) Interaction of familial and hormonal risk factors for breast cancer. J Natl Cancer Inst 69:817–822

Byrne C, Brinton LA, Haile RW, Schairer C (1991) Heterogeneity of the effect of family history on breast cancer risk. Epidemiology 2:276–284

Claus EB, Risch NJ, Thompson WD (1990) Age at onset as an indicator of familial risk of breast cancer. Am J Epidemiol 131:961–972

——— (1991) Genetic analysis of breast cancer in the Cancer and Steroid Hormone Study. Am J Hum Genet 48:232–242

Colditz GA, Rosner BA, Speizer FE (1996) Risk factors for breast cancer according to family history of breast cancer. J Natl Cancer Inst 88:365–371

Colditz GA, Willet WC, Hunter DJ, Stampfer MJ, Manson JE, Hennekens CH, Rosner BA, et al (1993) Family history, age, and risk of breast cancer: prospective data from the nurses' health study. JAMA 270:338–343

Demenais FM (1991) Regressive logistic models for familial diseases: a formulation assuming an underlying liability model. Am J Hum Genet 49:773–785

Demenais FM, Laing AE, Bonney GE (1992) Numerical comparisons of two formulations of the logistic regressive models with the mixed model in segregation analysis of discret traits. Genet Epidemiol 9:419–435

Demenais F, Lathrop M (1994) REGRESS: a computer program including the regressive approach into the LINKAGE programs. Genet Epidemiol 11:291

Demenais F, Martinez M, Bonaiti–Pellié C, Clerget–Darpoux F, Feingold N (1986) Segregation analysis of Jacobsen data. Genet Epidemiol Suppl 1:49–54

Easton DF, Bishop DT, Ford D, Crockford GP, Breast Cancer Linkage Consortium (1993) Genetic linkage analysis in familial breast and ovarian cancer: results from 214 families. Am J Hum Genet 52:678–701

Easton DF, Ford D, Bishop DT, Breast Cancer Linkage Consortium (1995) Breast and ovarian cancer incidence in BRCA1-mutation carriers. Am J Hum Genet 56:265–271

Elston RC, Sobel E (1979) Sampling considerations in the gathering and analysis of pedigree data. Am J Hum Genet 31:62–69

Elston RC, Steward J (1971) A general model for the analysis of human pedigree data. Hum Hered 21:523–542

Essioux L, Abel L, Bonaïti–Pellié C (1995) Genetic epidemiology of breast cancer: interest of survival analysis methods. Ann Hum Genet 59:271–282

Ford D, Easton DF, Peto J (1995) Estimates of the gene frequency of BRCA1 and its contribution to breast and ovarian cancer incidence. Am J Hum Genet 57:1457–1462

Gilligan SB, Borecki IB (1986) Examination of heterogeneity in 200 Danish breast cancer pedigrees. Genet Epidemiol Suppl 1:67–72

Goldgar DE, Fields P, Lewis CM, Tran TD, Cannon-Albright L, Ward JH, Swensen J, et al (1994) A large kindred with 17q linked breast and ovarian cancer genetic phenotypic genealogical analysis. J Natl Cancer Inst 86:200–209

Goldstein AM, Amos CI (1990) Segregation analysis of breast cancer from the Cancer and Steroid Hormone Study: histologic subtypes. J Natl Cancer Inst 82:1911–1917

Goldstein AM, Haile RWC, Hodge SE, Paganini–Hill A, Spence MA (1988) Possible heterogeneity in the segregation pattern of breast cancer in families with bilateral breast cancer. Genet Epidemiol 5:121–133

Goldstein AM, Haile RWC, Marazita ML, Paganini–Hill A (1987) A genetic epidemiologic investigation of breast cancer in families with bilateral breast cancer. I. Segregation analysis. J Natl Cancer Inst 78:911–918

Greenland S, Finkle WD (1995) A critical look at methods for handling missing covariates in epidemiologic regression analyses. Am J Epidemiol 142:1255–1264

Hall JM, Lee MK, Morrow J, Anderson L, King MC (1990) Linkage of early onset of familial breast cancer to chromosomes 17g 21. Science 250:1684–1689

Hartge P (1997) Abortion, breast cancer and epidemiology. N Engl J Med 336:127–128

Holt JT, Thompson ME, Szabo C, Robinson–Benion C, Arteaga CL, King M–C, Jensen RA (1996) Growth retardation and tumor inhibition by BRCA1. Nat Genet 12:298–302

Iselius L, Slack J, Littler M, Morton NE (1991) Genetic epidemiology of breast cancer in Britain. Ann Hum Genet 55: 151–159

Jensen RA, Thompson ME, Jetton TL, Szabo CI, Van der Meer R, Helou B, Tronick SR, et al (1996) BRCA1 is secreted and exhibits properties of a granin. Nat Genet 12:303–308

Kelsey JL, Horm–Ross PL (1993) Breast cancer: magnitude of the problem and descriptive epidemiology. Epidemiol Rev 15:7–16

Kerangueven F, Essioux L, Dib A, Noguchi T, Allione F, Geneix J, Longy M, et al (1995) Loss of heterozygosity and linkage analysis in breast carcinoma: indication for a putative third susceptibility gene on the short arm of chromosome 8. Oncogene 2:1023–1026

Lalouel JM, Morton NE (1981) Complex segregation analysis with pointers. Hum Hered 31:312–321

Lalouel JM, Rao DC, Morton NE, Elston RC (1983) A unified

model for complex segregation analysis. Am J Hum Genet 35:816–826

Lalouel JM, Yee S (1980) POINTER: a computer program for complex segregation analysis with pointers. Tech Rep, Population Genetics Laboratory, University of Hawaii, Honolulu

Lathrop GM, Lalouel JM (1984) Easy calculation of lod scores and genetic risks on small computers. Am J Hum Genet 36: 460–465

Malone KE, Daling JR (1992) Family history as a modifier of breast cancer risk factors. Am J Epidemiol 136:964

Marquis ST, Rajan JV, Wynshaw–Boris A, Xu J, Yin GY, Abel KJ, Weber BL, et al (1995) The developmental pattern of Brca1 expression implies a role in differentiation of the breast and other tissues. Nat Genet 11:17–26

Melbye M, Wohlfahrt J, Olsen J, Frisch M, Westergaard T, Helweg-Larsen K, Andersen PK (1997) Induced abortion and the risk of breast cancer. N Engl J Med 336:81–85

Mettlin C, Croghan I, Natarajan N, Lane W (1990) The association of age and familial risk in a case–control study of breast cancer. Am J Epidemiol 131:973–983

Michels KB, Willett WC (1996) Does induced or spontaneous abortion affect the risk of breast cancer? Epidemiology 7: 521–528

Michels-Blanck H, Byers T, Mokdad AH, Will JC, Calle EE (1996) Menstrual patterns and breast cancer mortality in a large US cohort. Epidemiology 7:543–546

Miki Y, Swensen J, Shattuck–Eidens D, Futreal PA, Harshman K, Tavtigian S, Liu Q, et al (1994) A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. Science 266:66–71

Narod SA, Feunteun J, Lynch HT, Watson P, Conway T, Lynch J, Lenoir G (1991) Familial breast–ovarian cancer chromosome 17q 12–23 Lancet 338:82–83

Narod SA, Ford D, Devilee D, Barkardottir RB, Lynch HT, Smith SA, Ponder BAJ (1995*a*) An evaluation of genetic heterogeneity in 145 breast–ovarian cancer families. Am J Hum Genet 56:254–264

Narod SA, Goldgar D, Cannon-Albright L, Weber B, Moslehi R, Ives E, Lenoir G, et al (1995*b*) Risk modifiers in carriers of BRCA1 mutations. Int J Cancer 64:394–398

Negri E, La Vecchia C, Bruzzi P, Dardanoni G, Decarli A, Palli D, Parazzini F, et al (1988) Risk factors for breast cancer: pooled results from three Italian case–control studies. Am J Epidemiol 128:1207–1215

Newman B, Austin MA, Lee M, King MC (1988) Inheritance of human breast cancer: evidence for autosomal dominant transmission in high–risk families. Proc Natl Acad Sci USA 85:3044–3048

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

Parazzini F, Negri E, La Vecchia C, Restelli C, Franceschi S (1992) Family history of reproductive cancers and ovarian cancer risk: an Italian case–control study. Am J Epidemiol 135:35–40

Phelan CM, Rebbeck TR, Weber BL, Devilee P, Ruttledge MH, Lynch HT, Lenoir GM, et al (1996) Ovarian cancer risk in BRCA1 carriers is modified by the HRAS1 variable number of tandem repeat (VNTR) locus. Nat Genet 12: 309–311

Rookus MA, van Leeuwen FE (1996) Induced abortion and risk for breast cancer: reporting (recall) bias in a Dutch case–control study. J Natl Cancer Inst 88:1759–1764

Sellers AS, Kushi LH, Potter JD, Kaye SA, Nelson CL, McGovern PG, Folsom A (1992) Effect of family history, body–fat distribution, and reproductive factors on the risk of postmenopausal breast cancer. N Engl J Med 326:1323–1329

Sellers TA, Potter JD, Severson RK, Bostick RM, Nelson CL, Kushi LH, Folsom AR (1993) Difficulty becoming pregnant and family history as interactive risk factors for postmenopausal breast cancer: the Iowa Women's Health Study. Cancer Causes Control 4:21–28

Sobol H, Birnbaum D, Eisinger F (1994) Evidence for a third breast–cancer susceptibility gene. Lancet 344:1151–1152

Theis B, Boyd N, Lockwood G, Trichler D (1994) Accuracy of family cancer history in breast cancer patients. Eur J Cancer Prev 3:321–327

Weed DL, Kramer BS (1996) Induced abortion, bias, and breast cancer: why epidemiology hasn't reached its limit. J Natl Cancer Inst 88:1698–1700

Williams WR, Anderson DE (1984) Genetic epidemiology of breast cancer: segregation analysis of 200 Danish pedigrees. Genet Epidemiol 1:7–20

Wooster R, Bignell G, Lancaster J, Swift S, Seal S, Mangion J, Collins N, et al (1995) Identification of the breast cancer susceptibility gene BRCA2. Nature 378:789–792

Wooster R, Neuhausen SL, Mangion J, Quirk Y, Ford D, Collins N, Nguyen K, et al (1994) Localization of a breast cancer susceptibility gene, BRCA2, to chromosome 13q12–13. Science 265:2088–2090

Wu AH, Ziegler RG, Pike MC, Nomura AMY, West DW, Kolonel LN, HornRoss PL, et al (1996) Menstrual and reproductive factors and risk of breast cancer in Asian–Americans. Br J Cancer 73:680–686