Family History Score as a Predictor of Breast Cancer Mortality: Prospective data from the Cancer Prevention Study II, United States, 1982-1991

by Quanhe Yang, Muin Khoury, Carmen Rodriguez, Eugenia E. Calle, Lilith M. Tatham, and W. Dana Flanders

Abstract

A consistent predictor of a woman's risk for breast cancer is a family history of the disease. Most studies of family history and breast cancer have used the number of affected relatives in the family to calculate relative risk, but have not considered the heterogeneity of the familial risk for breast cancer in a systematic way. Using data from a large prospective mortality study of US adults, the authors compared simple classification of family history of breast cancer (Yes/No) to the method of using a family history score (FHS), which takes into account the effects of family structure, age, and birth-cohort as predictors of breast cancer mortality. After 9 years of follow-up 1,428 cases of fatal breast cancer were observed among 453,073 women with complete information on number and age of siblings and family history. Using the FHS, about one third of women with a positive family history of breast cancer were at no higher risk for breast cancer mortality than those without a family history of the disease. As a quantitative measure of relative risk for each family, the FHS gave a better fit to the data, and it provided an incremental improvement of predictive accuracy of developing fatal breast cancer. FHS also can be used as a categorical variable to stratify families, allowing researchers to focus on which risk groups would benefit from conducting further genetic analysis and to test the effects of genetic factors, environmental exposure and gene-environment interactions on the etiology of development of breast cancer.

Studies have consistently shown that women with a positive family history of breast cancer in a first-degree relative have a higher risk of developing breast cancer than women without a family history of the disease (1-7). The increased risk is further elevated if two first-degree relatives have the disease (1, 8-11). Highest risk is experienced by those whose relatives developed breast cancer early in their lives (2, 8, 9, 12), but risk may be elevated even if the relative was affected when she was more than 70 years of age (3). Some reports have suggested a steady decrease in the association between a positive family history and the risk for breast cancer with increasing age (6, 12), whereas others found no variation in risk across age groups (13-16). Other studies have argued that, for most women with a family history of breast cancer, particularly those in whose mothers breast cancer was diagnosed when they were older, the excess risk is not large (3, 13, 14). Several studies have shown that there is clear evidence for the role of three genes (BRCA1, BRCA2, and p53) as the etiologic basis for the greatly increased risk for breast cancer observed in some families (17-21).

Clearly, studies of family history and breast cancer indicate a high degree of heterogeneity of familial risk for breast cancer. Breast cancer is likely to be multifactorial, involving genetic factors, environmental exposure, and the gene-environment interaction. Many studies of family history and breast cancer define exposure as the number of affected first-degree or second-degree relatives (coded as 0, 1, or 2 or more), sometimes stratified by the relative's age at onset of the disease. This method of using family-history information does not take into account the possible effects of family structure, age, or racial differences among families with a positive family history in assessing familial risk for breast cancer. It also fails to consider systematically the heterogeneity of the familial risk for breast cancer, which may vary considerably among families.

In this study, we compare the method of using simple categories of observed numbers of

cases of breast cancer in a family (0, 1, 2, or more) with using a quantitative family history score (FHS) method to predict breast cancer mortality. We also examine the power of different methods of defining a positive family history of breast cancer in predicting fatal breast cancer. The study uses data from a large prospective mortality study, the Cancer Prevention Study II (CPS-II), United States, 1982-1991. Although FHS had been used to study other chronic diseases (22-24), there is no prospective study, to our knowledge, which used FHS as a predictor of fatal breast cancer.

Materials and methods

Cancer Prevention Study II

Women in this study were selected from the 676,526 female participants of the Cancer Prevention Study II, a prospective mortality study of about 1.2 million American men and women begun by the American Cancer Society in 1982. Participants were identified and enrolled by more than 77,000 American Cancer Society volunteers in all 50 States, the District of Columbia, and Puerto Rico (25). In 1982, participants completed a confidential questionnaire that included personal identifiers; demographic characteristics; personal and family history of breast cancer and other diseases; and various behavioral, environmental, occupational, and dietary exposures. The median age of female study participants in 1982 was 56 years; 75 percent of the women were from 45 to 70 years of age, and none was younger than 30 years old.

Every 2 years, from 1982 through 1988, volunteers made personal inquiries to determine whether or not study participants were living and to record the date and place of all deaths. Automated linkage using the National Death Index (NDI) was used to extend follow-up through December 31, 1991 (26) and to identify women who had died among the 13,219 women lost to

follow-up between 1982 and 1988. Mortality follow-up was completed through December 31, 1991; at that time 615,009 women (90.9 percent) were still living, 59,439 (8.8 percent) had died, and for 2,078 (0.3 percent), follow-up stopped in September 1988 because of insufficient data for NDI linkage.

Population for analysis

From the initial group of 676,526 participants, we excluded from the analysis 3,275 women (12 breast cancer deaths) with incomplete information on race, 56,861 women (3,048 breast cancer deaths) who had prevalent cancer (except nonmelanoma skin cancer) at study entry in 1982, and 163,317 women (581 breast cancer deaths) with incomplete family history information. After nine years of follow-up, 1,428 eligible cases of fatal breast cancer were observed among 453,073 women. We treated these women as 453,073 families and assumed that no two women were from the same family. Among 1,428 eligible cases, 170 (11.9%) had family history of breast cancer. We used only mother's and sisters' information for the analysis and excluded all males from the family in order to calculate FHS.

Data on family history of cancer

Information about history of cancer in the parents and siblings (up to six siblings) was elicited on the 1982 questionnaire. This information included whether the parents and siblings were still living; the age and sex of siblings and if deceased, their age at death; and whether a diagnosis of cancer had been made in family members, the type of cancer, and the age at which the diagnosis was made.

Derivation of family history score

Breast cancer risk among families was measured using a statistic that describes deviations from expected risk for each family, and which takes into account family structure and the risk covariates of family members (age, sex, race, and birth-cohort) (22, 27, 28). We calculated standardized statistic for the *i*th family as follows:

$$T_i = \frac{\sum_j O_{ij} - \sum_j E_{ij}}{\sqrt{\sum_j E_{ij}(1 - E_{ij})}} \qquad (1)$$

where O_{ij} is the observed breast cancer status for *j*th member in family *I* (0 or 1), and E_{ij} is the expected risk for breast cancer for the *j*th member in family *I*. T_i is the family history score for family *I*.

The expected risk for breast cancer for each person in a family was obtained by multiplying age, race, and time-specific U.S. cancer-incidence rates from the Surveillance, Epidemiology, and End Results (SEER) program by age, race, and birth cohort-specific person-years at risk. Person-years at risk were accumulated until age at interview or age at death for people without breast cancer or age when a diagnosis of breast cancer was made for people with breast cancer. We estimated the cumulative incidence (risk) for breast cancer and the risk of a person developing breast cancer by a specific age as follows (29):

$$E_{ij} = 1 - Exp[-\sum_{kj} ID_k(\alpha t] \qquad (2)$$

where *k* is the age group, *ID* is the incidence density in the *k*th age group, and αt is the age interval. For people aged 65 years and older in 1982, we used 1973 SEER breast cancer-incidence rates, and for people younger than 65 years of age at interview, we used 1982 SEER breast

cancer-incidence rates. The calculation of the expected occurrence of breast cancer takes into account age, race, and birth-cohort differences in breast cancer incidence.

A negative value of FHS indicates that a family contains fewer women with breast cancer than would be expected, and a positive value of FHS indicates that a family contains more women with breast cancer cases than expected. The magnitude of the negative value of FHS was mainly determined by the family size and average age of the family members. The larger families with an older average age among family members without family history of breast cancer had a larger negative FHS because the expected value of breast cancer for those families was high (formula 1). Similarly, the smaller families with younger average age and with family history of breast cancer had larger positive value of FHS. In CPS-II data set, all families with negative FHS had no family history of breast cancer. The relative risks of developing fatal breast cancer among families with a negative FHS was calculated by classifying FHS into 10 groups. This analysis showed no trend and risks remained stable among these groups (results not shown). In the present study, therefore, we set the FHS for the families with negative FHS equal to zero since the main purpose of the study was to compare different methods of defining positive family history of breast cancer to predict future cases of breast cancer mortality (22, 23).

There are primarily two methods used to calculate FHS; each version is based on a comparison of the observed number of cases in a family with the expected number during the observation period, taking into account some covariates of family (22, 27). We presented our results using the method of calculating FHS developed by Chakroborty et al. (27). Williams et al. (22) proposed a slightly different method to calculate FHS which used a correction term in the formula to approximate the normal distribution. The usage of correction term set the FHS of all families with expected values <0.5 to zero. Williams et al. (22) compared these two methods and

concluded that their outcomes were similar. For our analysis, both methods gave similar rankings of positive FHS with different magnitudes of the score. Regardless of which method we used, the results (the relative risk for breast cancer obtained by FHS) remained unchanged.

Statistical analysis

To assess the association between a family history of breast cancer and fatal breast cancer, we used two different approaches. First, we classified families by the observed number of cases of breast cancer in first degree relatives as follows: no family history of breast cancer, one first-degree relative with cancer, and two or more first-degree relatives with breast cancer (only 0.4 percent of families with a positive family history of breast cancer had three or more first-degree relatives with breast cancer). Second, we classified all families by FHS into four groups: zero FHS, and three equal groups (33 percent each group) with a low, medium, or high FHS among families with positive FHS. We compared the relative risk of developing breast cancer as measured by these two different methods of measuring family history. We used the proportional hazards model to calculate hazard ratios (30) while adjusting for the effects of multiple risk factors, (menopausal status; age at menarche; age when first living child was born; history of breast cysts; oral contraceptive use; other estrogen use; body-mass index; diethystilbestrol (DES); education; religion; race; alcohol use; smoking status; and among postmenopausal women, the age at which periods stopped) and were stratified on the basis of age at enrollment.

Because the estimates of the adjusted relative risks (odds ratios) obtained by logistic regression were virtually identical to those obtained by proportional hazard model, we used logistic regression to examine the goodness-of-fit and the accuracy of different methods of defining a family history of breast cancer in predicting the development of fatal breast cancer.

Cox and Snell (31) and Nagelkerke (32) proposed a generalized coefficient of determination (R^2) and an adjusted generalized coefficient of determination (R^2_{adj}) to give an objective measure of how well each model fits the data. The R^2 and R^2_{adj} are analogous in interpretation to R^2 in linear regression analysis (32). The receiver operating characteristic (ROC) curve displays graphically the discriminatory ability of a logistic model. The ROC curve rises quickly with high predictive accuracy. Thus, the area under the curve (C statistic) approaches one for a model with higher predictive accuracy (33). The value of C statistics range from 0 to 1.

Results

As of 1982, 32,937 (7.3 percent) women reported that breast cancer had been diagnosed in either their mothers or sisters. Figure 1 shows the frequency distribution of breast cancer FHS across all families. The FHS ranged from a minimum of 0 to a maximum of 46.2. The distribution of positive FHS was skewed to the right with a mean FHS of 5.2 (standard deviation 2.2). To examine the relationship between the observed number of relatives with breast cancer in the families and the FHS, we tabulated the observed number of relatives with breast cancer in each of four FHS categories: zero, low (lower 1/3 of positive FHSs), medium (middle 1/3 of positive FHSs), and high (highest 1/3 of positive FHSs)(Table 1). As expected, all families without a family history of breast cancer had a zero FHS. Among families with a family history of breast cancer, the higher FHSs were associated with an increased number of relatives who had breast cancer. The kappa statistics of *?* = 0.65 (se = 0.0015) and *z* = 430 indicate that there is a good degree of agreement between the two approaches of classifying family history of breast cancer (34).

The families in which more women had more breast cancer tended to be larger (more

sisters) and older on average than those families in which fewer women had breast cancer, and a larger proportion of the families were white (Table 2). The average expected number of breast cancer cases calculated by cancer incidence rates was also higher for the families with a positive history of breast cancer than for the families without a history of breast cancer (Table 2). On the other hand, the families with higher positive FHSs tended to be smaller (there were fewer sisters in the families) and on average, younger, and a larger proportion of the families were from other racial groups. Among families with positive FHSs, FHSs were negatively correlated with expected number of breast cancer cases (Table 2).

Table 3 compares the effects of using different methods of defining a positive family history on the relative hazard for breast cancer mortality by age groups (we derived the adjusted relative hazard by using a proportional hazard model and controlling for other covariates). Using the usual definition of "any affected mother or sisters" resulted in calculated risks which were positively associated with a positive family history of breast cancer, with the magnitude of the relative hazard increasing with the age of the participants. Using FHS, about one third of the families with a positive family history (lower 1/3 FHS families) were not at significantly higher risk for death from breast cancer than those families with no family history of the disease. As shown before, compared with women whose families had high FHSs, women whose families had low FHSs were more likely to be from families with more and older family members; these families had a higher expected number of breast cancer cases. The simple method of classifying families by the number of the women with breast cancer would not be able to distinguish these families which had a positive family history of breast cancer but were not at higher risk of breast cancer mortality. In addition, the FHS also showed a dose-response relationship with the

increased risk for breast cancer mortality overall and in each age group. The adjusted relative hazard of breast cancer increased from 1.0 (95% CI, 0.8-1.4) for women from families with a low FHS to 2.3 (95% CI, 1.6-2.6) for women from families with a high FHS.

Both R^2 and adjusted R^2 were larger when we used the FHS in the model as a continuous variable than when we used the number of observed affected relatives in the family to predict breast cancer mortality. The C statistics also increased from 0.73 to 0.75, representing a 5 percent increase of the area under the ROC curve (Table 4). We found the same pattern when we compared using three equal positive FHS categories to predict breast cancer mortality with using categories of 0, 1, or 2+ for the observed number of cases of breast cancer in the family. For the 130 families in which 3 or more first degree relatives had breast cancer, none of the women participants developed fatal breast, so we were unable to calculate the odds ratios by logistic regression. Instead, we combined these families with those in which two or more relatives had breast cancer (Table 4). It is clear that by using FHS in the analysis the models fit the data better and gave an incremental improvement of predictive accuracy than would have resulted from classifying the family on the basis of the number of women in the family with breast cancer.

Discussion

In this study, we used relative hazard, which was virtually identical to the relative risk, to evaluate the FHS as a predictor of breast cancer mortality. Our analyses suggest that FHS gave a better fit to the data and provided a higher degree of predictive accuracy of breast cancer mortality than simply predicting that risk by designating family history as either positive or negative.

Because FHS is a continuous variable, every family was ranked by the FHS and used in

the analysis. Studies conducted for other diseases (23, 35) have suggested that, when estimating the relative risk of any specific risk factor in disease, using continuous FHS ranking to control for effects of family history should be more powerful than using either the observed number of breast cancer cases in a family or a two-or-three category grouping based on the number of affected individuals.

We calculated FHS on the basis of information about nuclear families. Presumably, the possible effects of family structure and age on familial risk of breast cancer would be greater among extended families than among nuclear families. The number of breast cancer cases would be greater among extended families than among nuclear families. The value of using FHS to assess the impact of a positive family history on breast cancer risk should increase with a greater number of relatives and with the complexity of the pedigrees. In our study, we assumed that those 1,428 eligible cases (170 cases had family history of breast cancer) were from different families. If those cases were not from different families, the FHS could overestimate the strength of association of FHS with fatal breast cancer mortality.

We excluded from our study 26 percent of CPS-II participants for whom family history information was incomplete; it is possible that the relative risk could differ between those with completed family information and those without it, and such a difference could bias the results. A previous study using a full data set of CPS-II on family history, age, and risk for breast cancer (6) showed similar relative risks for breast cancer by age groups such as those in our study. Such findings may suggest that the families for whom family information is incomplete were more likely to be not substantial different from other families. It should also be pointed out that the end point of CPS-II study was to determine breast cancer mortality. By deleting all breast cancer cases at baseline (1982), we limited fatal cases to those which were diagnosed and become fatal

within nine years of follow-up. This represented a group of more aggressive breast cancer cases. It is unclear how FHS would predict breast cancer mortality at a longer period of follow-up. In addition, the ability of FHS to predict breast cancer incidence remains unknown.

Other studies have shown that the family history of breast cancer, even among first degree relatives, is not perfectly reported (36-38). To assess the extent of possible underascertainment, we compared the reported prevalence of breast cancer among first degree relatives of CPS-II participants with estimates of expected prevalence based on breast cancer age-specific prevalence rates from the Connecticut Tumor Registry (39). The reported prevalence of a mother with breast cancer was 99% of the expected prevalence and the reported prevalence of sister(s) with breast cancer was 86% of the expected. If the probability of developing fatal breast cancer are nondifferential between reported families and not-reported families with family history of breast cancer, the estimates of relative risk using FHS will be diluted towards null. Differential probabilities of developing fatal breast cancer between reported and not-reported families would affect the estimation in either direction.

The FHS method has been shown to be capable of identifying "high-risk families" (families with positive family history of breast cancer) who may have no increased risk for breast cancer and of identifying an unaffected women's increased risk of developing breast cancer in the future. This knowledge would provide a better rationale for identifying the women who are most likely to be at high risk for fatal breast cancer. As a quantitative measure of relative risk for each family, the FHS can be used as a covariate in risk-factor analysis or as a categorical variable to stratify families by FHS. For example, if we classify the positive FHS into 20 groups (five percentiles), compared with women without family history of breast cancer, the adjusted risk of breast cancer increased from 1.0 [95% CI, 0.7-1.5] for women of first five percentile group to 2.1

(95% CI, 1.3-3.1] for women tenth five percentile group to 5.4 [95% CI, 1.7-17.2) for women of last five percentile group. This method of cross-classifying families may allow researchers to better identify those who would benefit from further genetic analysis and to test the effects of genetic factors, environmental exposures and gene-environment interaction on breast cancer etiology.

Acknowledgments

The authors thank J. David Erickson and Jim Buehler for their helpful suggestions and Michael Atkinson for his technical assistance. They also thank two anonymous reviewers for their constructive comments.

References

1. Ottman R, Pike MC, King M-C, et al. Familial breast cancer in a population-based series. Am J Epidemiol 1986;123:15-21.

2. Claus EB, Risch NJ, Thompson WD. Age at onset as an indicator of familial risk of breast cancer. Am J Epidemiol 1990;131;961-972.

3. Colditz GA, Willett WC, Hunter DJ, et al. Family history, age, and risk of breast cancer: prospective data from the Nurses' Health Study. JAMA 1993;270:338-343.

4. Slattery ML, Kerber RA. A comprehensive evaluation of family history and breast cancer risk. The Utah Population Database. JAMA 1993;270:1563-1568.

5. Sellers TA, Potter JD, Rich SS, et al. Familial clustering of breast and prostate cancers and risk of postmenopausal breast cancer. J Natl Cancer Inst 1994;24:1860-1865.

6. Calle EE, Martin LM, Thun MJ, et al. Family history, age, and risk of fatal breast cancer. Am J

Epidemiol 1993;138:675-681.

7. Anderson DE. Genetic study of breast cancer: identification of a high risk group. Cancer 1974;34:1090-1097.

8. Bain C, Speizer FE, Rosner B et al. Family history of breast cancer as a risk indicator for the disease. Am J Epidemiol 1980;111(3):301-308.

9. Brinton LA, Hoover R, Fraumeni JF. Interaction of familial and hormonal risk factors for breast cancer. J Natl Cancer Inst 1982;69(4):817-822.

10. Sattin RW, Rubin GL, Webster LA, et al. Family history and the risk of breast cancer. JAMA 1985;253:1908-1913.

11. Byrne C, Brinton LA, Haile RW, et al. Heterogeneity of the effect of family history on breast cancer risk. Epidemiology 1991;2:276-284.

12. Roseman DL, Straus AK, Shorey W. A positive family history of breast cancer: does its effect diminish with age? Arch Intern Med 1990;150:191-194.

13. Adami HO, Hanssen J, Jung B, et al. Familiality and breast cancer: a case-control study in a Swedish population. Br J Cancer 1980;42:71-77.

14. Schneider NR, Chaganti SR, German J, et al. Familial predisposition to cancer and age at onset of disease in randomly selected cancer patients. Am J Hum Genet 1983;35:454-467.

15. Lubin JH, Burns PE, Blot WJ, et al. Risk factors for breast cancer in women in Northern Alberta, Canada, as related to age at diagnosis. J Natl Cancer Inst 1982;68:211-217.

16. Negri E, La Vecchia C, Bruzzi P, et al. Risk factors for breast cancer: pooled results from three Italian case-control studies. Am J Epidemiol 1988;128:1207-1215.

17. Hall JM, Lee MK, Newman B, et al. Linkage of early-onset familial breast cancer to chromosome 17q21. Science 1990;250:1684-1689.

18. Malkin D, Li FP, Strong LC, et al. Gemline p53 mutations in a familial syndrome of breast cancer, sarcomas and other neoplasms. Science 1990;250:1233-1938.

19. Miki Y, Swensen J, Shattuck-Eidens D, et al. A strong candidate for the breast and ovarian cancer susceptibility gene. BRCA1 Science 1994;266:66-71.

20. Wooster R, Neuhausen SL, Mangion J, et al. Localization of a breast cancer susceptibility gene, BRCA1, to chromosome 13q12-13. Science 1994;265:2088-2090.

21. Fitzgerald MG, MacDonald DJ, Krainer M, et al. Germ-line BRCA1 mutations in Jewish and non-Jewish women with early-onset breast cancer. New Engl J Med 1996;334:143-149.

22. Williams RR, Dadone MM, Hunt SC, et al. The genetic epidemiology of hypertension: a review of past studies and current results for 948 persons in 48 Utah pedigrees. In: Rao DC, Elston RC, Kuller LH, et al., eds. Genetic Epidemiology of Coronary Heart Disease: Past, Present and Future. New York, NY: Alan R. Liss; 1984;419-442.

23. Hunt SC, Williams RR, Barlow GK. A comparison of positive family history definitions for defining risk of future disease. J Chron Disease 1986;39:809-821.

24. Howard G, Evans GW, Toole JF, et al. Characteristics of stroke victims associated with early cardiovascular mortality in their children. J Clin Epidemiol 1990;43(1):49-54.

25. Garfinkel L. Selection, follow-up, and analysis in the American Cancer Society prospective studies. Monogr Natl Cancer Inst 1985;67:49-52.

26. Calle EE, Terrell DD. Utility of the National Death Index for ascertainment of mortality among Cancer Prevention Study II participants. Am J Epidemiol 1993;137:235-241.

27. Chakroborty R, Weiss KM, Majumder PP, et al. A method to detect excess risk of disease in structured data: cancer in relatives of retinoblastoma. Genet Epidemiol 1984;1:229-244.

28. Fain PR, Goldgar DE. A nonparametric test of heterogeneity of family risk. Genet Epidemiol

1986;[Suppl]:61-66.

29. Rothman KJ. *Modern Epidemiology*. Boston, Mass: Little Brown & Co Inc; 1986.

30. Cox DR. Regression models and life tables. J R Stat Soc (B) 1972;34:187-220.

31. Cox DR, Snell EJ. *The Analysis of Binary Data*. Second ed. London: Chapman and Hall; 1989.

32. Nagelkerke NJD. A note on general definition of the coefficient of determination. Biometrika 1991;78:691-692.

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

34. Fleiss JL. Statistical methods for raates and proportions. Second ed. New York: John Wiley & Sons; 1981.

35. Reed T, Wagener DK, Donahue RP, et al. Family history of cancer related to cholesterol level in young adults. Genet Epidemiol 1986;3:63-71.

36. Breuer B, Kash KM, Rosenthal G, et al. Reporting bilaterality status in first-degree relatives with breast cancer: A validity study. Genet Epidemiol 1993;10:245-256.

37. Theis B, Boyd N, Lockwood G, et al. Accuracy of family cancer history in breast cancer patients. Eur J Cancer Prev 1994;3:321-327.

38. Anton-Culver H, Kurosaki T, Taylor TH, et al. Validation of family history of breast cancer and identification of the BRCA1 And other syndromes using a population-based cancer registry. Genet Epidemiol 1996;13:193-205.

39. Feldman AR, Kessker L, Myers MH, et al. The prevalence of cancer. Estimates based on the Connecticut Tumor Registry. N Engl J Med 1986;315:1394-7.

TABLE 1. Distribution of families by the observed number of first degree relatives with breast cancer and by family history score (FHS): Cancer Prevention Study II, United States, 1982-91

	Observed number of relatives with breast cancer				Kappa
FHS*	0	1	2	3+	statistics
0	420,136	0	0	0	k=0.65 (SE‖=.0015)
1	0	10,979	0	0	z = 430
2	0	10,871	104	0	
3	0	9,268	1,585	130	
Total	420,136	31,118	1,689	130	

* All families with positive FHS were divided into three equal groups, I.e.,1 = low, first 33%; 2 = medium, second 33%; and 3 = high FHS, third 33%.

‖ SE, standard error.

TABLE 2. Family characteristics by different methods of classifying family history of breast cancer: Cancer Prevention Study II, United States, 1982-1991

Family history of breast cancer	Number of families	Average family size* (SD)	Mean age of family members (SD)	Percent white	Number of affected relatives (SD)	Expected number of cases (SD)	Mean FHS[H] (SD)
Observed cases							
0	420136	3.3 (1.3)	60.4 (9.9)	93.8	0.0	0.048 (.027)	0.0
1	31118	3.6 (1.3)	61.1 (9.1)	96.4	1.0	0.049 (.030)	5.0 (2.0)
2	1689	4.5 (1.2)	62.6 (8.0)	97.3	2.0	0.063 (.030)	8.6 (2.8)
3+	130	5.5 (1.1)	62.2 (7.7)	97.7	3.0	0.073 (.034)	12.5 (4.1)
FHS[I]							
0	420136	3.3 (1.3)	60.4 (9.9)	93.8	0.0	0.048 (.027)	0.0
1	10979	4.6 (1.2)	66.4 (7.0)	97.6	1.0	0.082 (.025)	3.3 (0.5)
2	10975	3.3 (1.0)	61.4 (7.7)	96.3	1.1 (0.1)	0.042 (.011)	4.8 (0.4)
3	10983	2.9 (1.1)	55.8 (9.1)	95.5	1.2 (0.4)	0.026 (.018)	7.6 (2.2)
All families	453073	3.3 (1.3)	60.5 (9.9)	94.0	0.08 (0.28)	0.049 (.027)	0.38 (1.5)

* The family includes mothers and sisters only.

[H] SD = Standard deviation; FHS, family history score.

[I] We divided all families with positive family history score (FHS) into three equal groups (1 = low, first 33%; 2 = medium, second 33%; and 3 = high FHS, third 33%)

TABLE 3. Comparison of the effect of different family history definitions on the risk of fatal breast cancer: Cancer Prevention Study II, United States, 1982-1991

Age of participants (years)	Observed no. of breast cancer cases in family				FHS*,H					
	1		2		1		2		3	
	Relative hazard	95% CI*	Relative hazard	95% CI	Relative hazard	95% CI	Relative hazard	95% CI	Relative hazard	95% CI
Unadjusted										
< 40	4.9	2.0-11.8	I	I	I	I	I	I	5.1	2.1-12.4
40-49	2.3	1.6-3.4	0.99	0.26-4.2	0.75	0.11-5.4	1.7	1.3-2.8	3.3	2.2-4.9
50-59	1.8	1.4-2.4	1.06	0.25-4.2	1.4	0.80-2.3	1.9	1.3-2.8	1.9	1.3-3.0
>= 60	1.3	0.96-1.6	1.07	0.40-2.9	1.0	0.68-1.5	1.5	0.97-2.2	1.7	0.96-2.9
All age	1.7	1.4-2.0	2.0	1.1-3.6	1.3	0.97-1.7	1.8	1.4-2.3	2.0	1.6-2.6
Fully adjustedξ										
All age	1.6	1.4-1.9	1.7	0.94-3.1	1.0	0.8-1.4	1.6	1.3-2.1	2.3	1.8-2.9

* FHS, family history score; CI, confidence interval.

H All families with positive FHS were divided into three equal groups, I.e., 1 = low, first 33%; 2 = medium, second 33%; and 3 = high FHS, third 33%.

I The relative hazard can not be calculated because there were no cases of breast cancer in the category.

ξ Estimates were adjusted for menopausal status; age at menarche; age when first living child was born; history of breast cysts; oral contraceptive use; other estrogen use; body-mass index; diethystilbestrol (DES); education; religion; race; alcohol use; smoking status; and among postmenopausal women, the age at which periods stopped.

TABLE 4. Summary of estimates from proportional hazard and logistic analysis by different family history classifications: Cancer Prevention Study II, United States, 1982-1991

	Proportional Hazard		Logistic regression				
Family history	Relative hazard*	95% CI[H]	Relative risk*	95% CI	R^2	R^2_{adj}	C statistics
Observed breast cancer cases[I]	1.49	1.30-1.73	1.50	1.30-1.73	0.033	0.085	0.73
FHS[HI]	1.10	1.07-1.13	1.10	1.07-1.12	0.036	0.091	0.75
Observed breast cancer cases[ξ]							
0	1.0		1.0				
1	1.59	1.36-1.27	1.59	1.36-1.28			
2+	1.71	0.94-3.09	1.72	0.95-3.10	0.035	0.087	0.73
FHS[ξ]							
0	1.0		1.0				
1	1.06	0.78-1.43	1.06	0.78-1.45			
2	1.64	1.27-2.14	1.66	1.29-2.19			
3	2.31	1.80-2.96	2.30	1.80-2.95	0.039	0.099	0.77

* Estimates were adjusted for menopausal status; age at menarche; age when first living child was born; history of breast cysts; oral contraceptive use; other estrogen use; body-mass index; diethystilbestrol (DES); education; religion; race; alcohol use; smoking status; and among postmenopausal women, the age at which periods stopped.

[H] CI, confidence interval; FHS, family history score.

[I] Observed breast cancer cases and FHS used as continuous variables.

[ξ] Observed breast cancer cases and FHS used as categorical variables. All families with positive FHS were divided into three equal groups, i.e., 1 = low, first 33%; 2 = medium, second 33%; and 3 = high FHS, third 33%.

FIGURE 1. Frequency distribution of family history score of breast cancer: Cancer Prevention Study II, United States, 1982-1991