Estimates of the Gene Frequency of BRCA^I and Its Contribution to Breast and Ovarian Cancer Incidence

D. Ford, D. F. Easton,* and J. Peto

Section of Epidemiology, Institute of Cancer Research, Belmont, Surrey

Summary

The majority of multiple-case families that segregate both breast and ovarian cancer in a dominant fashion are due to mutations in the BRCA1 gene on chromosome 17q. In this paper, we have combined penetrance estimates for BRCA1 with the results of two populationbased genetic epidemiological studies to estimate the gene frequency of BRCA1. On the assumption that the excess risk of ovarian cancer in first degree relatives of breast cancer patients and the breast cancer excess in relatives of ovarian cancer patients are both entirely accounted for by BRCA1, we estimate that the BRCA1 gene frequency is 0.0006 (95% confidence interval $[0.0002 - 0.001]$ and that the proportion of breast cancer cases in the general population due to BRCA1 is 5.3% below age 40 years, 2.2% between ages 40 and 49 years, and 1.1% between ages 50 and 70 years. The corresponding estimates for ovarian cancer are 5.7%, 4.6%, and 2.1%, respectively. Our results suggest that the majority of breast cancer families with less than four cases and no ovarian cancer are not due to rare highly penetrant genes such as BRCA1 but are more likely to be due either to chance or to more common genes of lower penetrance.

Introduction

BRCA1, which increases susceptibility to both breast and ovarian cancer, was localized to chromosome 17q by genetic linkage in 1990 (Hall et al. 1990; Narod et al. 1991) and has recently been cloned (Miki et al. 1994). Linkage studies have shown that BRCA1 accounts for almost all multiple case breast-ovarian cancer families (Easton et al. 1993; Narod et al. 1995). BRCA1 mutations confer a high lifetime risk of both cancers, but

*Current address: CRC Genetic Epidemiology Unit, Institute of Public Health, University of Cambridge, Cambridge.

© ¹⁹⁹⁵ by The American Society of Human Genetics. All rights reserved. 0002-9297/95/5706-0024\$02.00

there is evidence that the risk of ovarian cancer is heterogeneous (Easton et al. 1995). A second breast cancer gene, BRCA2, has recently been localized to chromosome 13q (Wooster et al. 1994). BRCA2 appears to cause predominantly breast cancer, and the age-specific breast-cancer penetrance of BRCA2 is similar to that of BRCA1 (D. F. Easton and D. E. Goldgar, personal communication).

In this paper, we use data from two population-based studies of cancer mortality in the relatives of breast and ovarian cancer patients (Easton et al., in press; Peto et al., in press) to estimate the gene frequency of BRCA1. We consider the contribution of BRCA1 to overall breast and ovarian cancer incidence and to multiple case families. The possible contribution of more common genes of lower penetrance is also discussed.

Methods

Age-specific incidence rates of breast and ovarian cancer in BRCA1 mutation carriers were estimated from the model described by Easton et al. (1995), which assumes two BRCA1 susceptibility alleles, "type 1" comprising 89% of mutations and conferring ^a 26% risk of ovarian cancer by age 70 years and "type 2" comprising 11% of mutations and conferring an 85% risk of ovarian cancer by age 70. Both alleles confer ^a 76% risk of breast cancer by age 70 years. Corresponding age-specific cumulative mortality rates for breast and ovarian cancer were estimated by adjusting these incidence rates by general population relative survival probabilities (supplied for age at diagnosis in 10-year intervals and time since diagnosis in 1-year intervals by the Thames Cancer Registry). We assume here that the survival rates for breast and ovarian cancer are similar in gene carriers and noncarriers. The estimated cumulative ovarian cancer mortality by age 70 years (in the absence of other causes of death) was 21% for the type ¹ allele and 72% for type 2. Both alleles were assumed to confer a breast cancer mortality of 42% by age 70 years. Population rates for England and Wales were used for incidence and mortality rates in noncarriers.

These rates were then used to estimate the BRCA1 gene frequency from population-based studies of cancer mortality risks in first-degree relatives of breast and ovarian cancer probands (Easton et al., in press; Peto et

Received April 14, 1995; accepted for publication August 22, 1995. Address for correspondence and reprints: Ms. Deborah Ford, Section of Epidemiology, Institute of Cancer Research, Block D, 15 Cotswold Road, Belmont, Surrey SM2 SNG, United Kingdom: E-mail: DEBBIE@ICR.AC.UK

al., in press). The probability that each breast cancer proband was ^a BRCA1 mutation carrier (either type ¹ or 2) was calculated from her age at diagnosis. The probability, q_{ik} , that a breast cancer proband diagnosed in age group j has genotype l is given by the formula:

$$
q_{jl} = \frac{a_l f_{lj} G_{lj}}{\sum_{m} a_m f_{mj} G_{mj}}, \qquad (1)
$$

where f_{li} is the probability of being diagnosed with breast cancer at the midpoint of age group i , in the absence of an ovarian cancer risk or other causes of death, for an individual with genotype $l (l = 0$ for noncarriers, 1 for BRCA1 type ¹ carriers, and 2 for BRCA1 type 2 carriers). G_{ij} is the probability that an individual with genotype *l* will not have died of ovarian cancer by the midpoint of age group j , in the absence of a breast cancer risk or other causes of death, and a_i is the proportion of individuals of genotype l in the general population at birth, so that $a_0 = (1 - p)^2$, $a_1 = (\alpha p)^2 + 2\alpha p(1 - \alpha p)$, and $a_2 = 2p(1 - \alpha)(1 - p)$, where p is the overall mutant BRCA1 gene frequency, and α is .89, the assumed proportion of BRCA1 mutations that are type 1. Similar probabilities were calculated for ovarian cancer probands.

The probabilities π_{ijk} that a relative in age group *i* of a proband in age group j had genotype k were then calculated from the formula:

$$
\pi_{ijk} = \left(\sum_{l} q_{jl} \tau_{lk}\right) F_{ki} G_{ki} , \qquad (2)
$$

where τ_{lk} is the probability that a first-degree relative of an individual with genotype l has genotype k at birth, and F_{ki} is the probability that an individual with genotype k will not have died of breast cancer by the midpoint of age group i , in the absence of an ovarian cancer risk and other causes of death. (Note that π_{ijk} decreases for $k = 1$ and $k = 2$ as the age of the relative increases.) Expected numbers of breast and ovarian cancer deaths in female first-degree relatives were then computed in the standard way using the Person Years program (Coleman et al. 1986), assuming national age- and periodspecific mortality rates for noncarriers $(k = 0)$ and the estimated mortality rates described above for carriers (k $= 1, 2$).

The total predicted number of breast (ovarian) cancer deaths in each cohort was then computed as:

$$
E = \sum E_{ijk} \pi_{ijk} \,, \tag{3}
$$

where E_{ijk} is the expected number of breast (ovarian) cancer deaths for relatives of genotype k in age group i of a proband in age group *j*. The gene frequency p of BRCA1 was estimated from the study of cancer mortality in relatives of breast cancer probands as the frequency required to make the total predicted number of ovarian cancer deaths equal the observed number. An independent estimate of the gene frequency was obtained similarly from the study of cancer mortality in relatives of ovarian cancer probands by requiring BRCA1 to explain the observed excess of breast cancer deaths.

The resulting estimate of the gene frequency for BRCA1 enabled us to compute predicted numbers of multiple case families due to BRCA1 or occurring sporadically in the population-based studies. For each study, exit rates for all causes except breast or ovarian cancer mortality were computed by summing age-specific cohort censoring rates and population death rates. For an individual of genotype k, the probability d_{ik} of dying of breast cancer (or ovarian cancer) in age interval, i, was calculated as:

$$
d_{ik} = (S_{ik} - S_{(i+1)k}) b_{ik} / r_{ik} \t{, \t(4)}
$$

where S_{ik} is the probability of being alive (and not censored) at the beginning of interval *i*, and r_{ik} is the exit rate in interval *i*. $r_{ik} = e_i + b_{ik} + o_{ik}$, where e_i is the exit rate for all causes except breast and ovarian cancer mortality and is independent of genotype, and b_{ik} and o_{ik} are the genotype-specific death rates in interval *i* for breast and ovarian cancer.

For each female first-degree relative of an individual with genotype *l*, the probability of dying of breast cancer (or ovarian cancer) in interval i was then computed as:

$$
u_l = \sum_i \sum_k \tau_{lk} d_{ik} , \qquad (5)
$$

where τ_{lk} is the probability that a first-degree relative of an individual with genotype l has genotype k . Probabilities of death were summed over all intervals between 1939, the commencement of follow-up, and age 70 years or the end of 1992 when follow-up ended. (Individual deaths and censorings that occurred during follow-up were ignored.) Probabilities v_i of dying of ovarian cancer were computed similarly. By combining the probabilities u_1 and v_1 for relatives in a family the probability w_{bol} of each possible permutation of numbers of breast cancers (b) and numbers of ovarian cancers (o), conditional on the probands genotype was computed. The expected number of families with each combination of numbers of breast and ovarian cancers where the proband was genotype l was then computed by summing w_{bol} over all families, weighting each family by the proband's probability q_{il} .

Estimated Proportion of Cases due to BRCAI

Results

In the study of cancer mortality in first-degree relatives of breast cancer probands, 49 ovarian cancer deaths were observed under age 70 years, compared with 33.71 expected at national rates, an excess of 15.29 deaths (Peto et al., in press). For this excess to be entirely due to BRCA1 the overall gene frequency of BRCA1 must be .00064. Similarly, in the companion study of cancer mortality in first-degree relatives of ovarian cancer probands, 45 breast cancer deaths in females were observed under age 70 years, compared with 33.63 expected, an excess of 11.37 deaths (Easton et al., in press), giving an estimated BRCA1 gene frequency of .00052. Since the gene-frequency estimate is directly proportional to the size of the excess, the variance of each estimate can be computed directly. Thus, the best estimate of the BRCA1 gene frequency, taken as the inverse variance weighted mean, is .0006 with 95% confidence limits 0.0002-0.001. Table ¹ shows the proportions of breast and ovarian cancers at different ages in the general population that are due to BRCA1 assuming ^a gene frequency of .0006 and the estimated risks associated with BRCA1.

In the study of cancer mortality in relatives of breast cancer patients, 185 breast cancer deaths were observed under age 70 years in females, compared with 95.72 expected at national rates (Peto et al., in press). If the familial effect were entirely due to BRCA1, and assuming a gene frequency of .0006, then 122.74 deaths would be predicted. Thus, 30% of the excess breast cancer risk is explained by BRCA1. Table 2 shows predicted numbers of multiple case families, assuming that BRCA1 is the only familial effect. There were 165 families with 2 cases of breast cancer (113.2 predicted), of which 90.1 would have been expected at national rates, and only 23.1 can be attributed to BRCA1. There were 10 families with ≥ 3 cases of breast cancer (3.5 predicted), of which 1.0 would have been expected at national rates and 2.4 were predicted due to BRCA1.

In the study of cancer mortality in relatives of ovarian cancer patients, 30 ovarian cancer deaths were observed under age 70 years, compared with 11.80 expected (Easton et al., in press). If only the effect of BRCA1 is allowed for, 22.25 deaths are predicted. Thus, 57% of the excess ovarian cancer risk is explained by BRCA1. Table 2 shows the corresponding predicted numbers of multiple case families.

Discussion

This study provides convincing evidence that the frequency of BRCA1 mutation carriers in the general population is low. Our estimate that the gene frequency is .0006 implies that only 1/833 women carries ^a mutation, and even at the upper 95% confidence limit the carrier frequency would only be 1/500 women. The proportions of breast and ovarian cancer due to BRCA1 in the general population are correspondingly low. We estimate that only 1.7% of all breast cancer cases diagnosed below age 70 years are due to BRCA1, the proportion rises to 5.3% of cases below age 40 years and to 7.5% below age 30 years. We similarly estimate that 2.8% of all ovarian cancer cases diagnosed below age 70 years are attributable to BRCA1; the proportions for the under 40-year and under 30-year age groups are 5.7% and 5.9%, respectively.

It is interesting to note that, in a targeted screening study for the 5382 ins C mutation in an unselected series of ovarian cancers in the UK, one mutation was found in 250 cases (0.4%) (Shattuck-Eidens et al. 1995). This mutation occurs in ¹¹ % of families in which mutations have been found. Some studies of mutations in series of breast cancer patients were also reported by Shattuck-Eidens et al. (1995), but it is unclear to what extent these studies are biased toward patients with a family history or early onset disease.

Our estimate of the BRCA1 gene frequency is based on the assumption that all excess ovarian cancer in relatives of breast cancer patients was due to BRCA1 and, conversely, that all excess breast cancer in relatives of ovarian cancer patients was due to BRCA1. This is a reasonable approximation, in light of the high proportion of breast-ovarian cancer families that are linked to BRCA1. In a study of 132 breast-ovarian cancer families, Narod et al (1995) estimated that 82% of such families were linked (95% confidence interval on proportion linked 74%-97%). BRCA2 mutations also cause ovarian cancer, but the risk is probably too low to have a measurable effect in population-based studies. It is possible that a proportion of the excess familial risk observed in these population-based studies may be due to other genes of lower penetrance, in which case the gene-frequency estimate for BRCA1 would be proportionately lower. We have also assumed that there are no BRCA1 alleles that cause only breast cancer or only ovarian cancer. If such mutations exist, the gene fre-

Table 2

Predicted Numbers of Multiple Case Families Obtained by Modeling the Effect of BRCA ^I within Population-Based Nuclear-Family Studies

^a Excluding index case.

^b Sporadic rate assumed to be equal to population rate.

^c Estimated BRCA1 penetrance function from Easton et al. (1995), gene frequency .0006.

quency could be higher. However, few large, apparently linked families have been reported with only breast or only ovarian cancers. Thus, it seems likely that such mutations, if they exist, are rare in comparison with breast-ovary mutations. The existence of BRCA1 mutations conferring substantially lower overall cancer risks would also imply that we have underestimated the gene frequency, but to date there is no evidence for such low penetrance mutations.

Our estimate also depends critically on the relative risk of ovarian cancer in relatives of breast cancer patients (1.45) and of breast cancer in relatives of ovarian cancer patients (1.34). At any given age, these relative risks should be identical. The majority of previous cohort and case-control studies that reported the relative risk of ovarian cancer in first-degree relatives of breast cancer cases (Schildkraut et al. 1989; Anderson et al. 1992; Tulinius et al. 1992; Parazzini et al. 1993; Goldgar et al. 1994; Teare et al. 1994) or vice-versa (Cramer et al. 1983; Tzonou et al. 1984; Koch et al. 1989; Schildkraut et al. 1989) included breast and ovarian cancer cases of all ages. If we pool the estimates from these studies, we obtain a relative risk of 1.32 (95% confidence interval [1.17-1.48]). This is expected to be somewhat lower than our estimates, since in the current studies probands were all <60 years of age at diagnosis and the majority were <50 years of age. If we combine the estimates from studies that reported the risk of ovarian cancer in relatives of young breast cancer probands (Schildkraut et al. 1989; Anderson et al. 1992; Goldgar et al. 1994), we do obtain ^a higher estimate (1.60 [1.24-2.07]). A relative risk of 1.6 corresponds to an excess risk \sim 50% higher than the one used in the current study, and on this basis the gene frequency estimate would be \sim 0.0009, just below the upper 95% confidence limit on our estimate.

We allowed, in our analyses, for heterogeneity of ovarian cancer risk between mutations by assuming two mutations with differing risks. Precise mutation specific risks cannot yet be determined, because of the large number of possible mutations; however, preliminary evidence suggests that mutations in the ³' third of the gene are associated with a lower risk of ovarian cancer (Shattuck-Eidens et al. 1995).

We estimated the proportion of the excess familial ovarian cancer risk attributable to BRCA1 to be 57%. The unexplained excess in this study is confined to sisters and may therefore be the result of ^a recessive gene(s) or shared environmental factors. The discrepancy between observed and predicted ovarian cancer deaths (30 versus 22.25) is however barely statistically significant (1-sided $P = .05$). Given the additional uncertainty in the gene-frequency estimate and the ovarian cancer penetrance, we cannot exclude the possibility that BRCA1 accounts for all excess familial ovarian cancer risk.

The estimated proportion of the excess familial breast

^a Excluding index case.

 b Sporadic rate = 0.6 of population rate.</sup>

^c Estimated BRCA1 penetrance function from Easton et al. (1995), gene frequency .0006.

^d Breast cancer risk as for BRCA1, ovarian cancer risk as in general population, gene frequency .0002.

^e Carrier:sporadic rate ratio 8.2, gene frequency .046. Derived from Scott et al. (1994).

 f Sporadic rate = population rate.

⁸ Penetrance function and gene frequency, 0.003, from CASH segregation analysis (Claus et al. 1991).

cancer risk attributable to BRCA1 was only 30%. Even if the BRCA1 gene frequency were .001, only 49% of the excess familial risk would be explained. There were marked deficits in the predicted numbers of breast-cancer only families with 2 or more cases of breast cancer (110.5 predicted and 163 observed with 2 cases of breast cancer; 3.3 predicted and 10 observed with 3 or more cases of breast cancer). These observations, while imprecise, are consistent with Breast Cancer Linkage Consortium results, which suggest that only a minority of families with 2 or 3 cases of breast cancer diagnosed ≤ 60 years of age are linked to BRCA1 (Easton et al. 1993).

An important unresolved issue is the genetic basis underlying familial breast cancer, which is not explained by BRCA1 or BRCA2. In the Breast Cancer Linkage Consortium data set (Easton et al. 1993), approximately three-quarters of families with four or more breast cancers are linked to BRCA1. This may be an overestimate, since the families analyzed by the Consortium may be biased by the selective inclusion of breast-ovarian cancer families, hence inflating the proportion linked. We have, however, on this basis, assumed that the combined gene frequency of other high-penetrance genes (BRCAX) is .0002. The breast cancer penetrance of BRCA2 (included within BRCAX) is similar to that of BRCA1 (D. F. Easton and D. E. Goldgar, personal communication), and other high-penetrance genes (if they exist) are likely to be rare in comparison to BRCA2, which accounts for the majority of large families unlinked to BRCA1 (Wooster et al. 1994). We therefore assumed the BRCA1 breast cancer penetrance for BRCAX but no excess ovarian cancer risk. On these assumptions, the numbers of multiple case breast cancer families that

are due to other high-penetrance genes would be approximately one-third of the numbers due to BRCA1. The marked discrepancy between observed and predicted numbers of families with two or three breast cancers but no ovarian cancer shown in the left-hand part of table 2 thus cannot be explained entirely by this model. One plausible explanation for this discrepancy is that the residual familial effect is due to ^a common gene(s) of low penetrance. For example Scott et al. (1994) recently reported that 42% of breast cancer patients have a radiation sensitivity phenotype that is present in 9% of the population. The simplest model consistent with these figures is a dominant gene with frequency ⁵% and ^a constant carrier:sporadic breast cancer incidence ratio of 8. If we allow for this effect in addition to the effects of BRCA1 and BRCAX (left-hand part of table 3), we predict very similar numbers of multiple case families to those observed. The combined effect of such genes would cause \sim 45% of all breast cancers before 70 years of age. Whether this speculative model is correct can be resolved only by the identification and characterization of such a hypothetical gene(s). The right-hand part of table 3 shows that the distribution of multiple case families we observed could also be accounted for almost exactly by the model derived by segregation analysis of the Cancer and Steroid Hormone Study data involving highly penetrant genes with combined frequency .003, which would cause $\sim 8\%$ of all breast cancers under age 70 years (Claus et al. 1991). This model is, however, inconsistent with our estimate of the BRCA1 gene frequency and the observation that BRCA1 and BRCA2 account for most high risk families.

Note added in proof.—Analysis of recent Breast Cancer Linkage Consortium data suggests that BRCA1 may account for less than three-quarters of families with four or more breast cancers, and, thus, our estimate of the combined gene frequency of other high-penetrance breast cancer genes could be too low. However, our estimate of the gene frequency of BRCA1 remains unchanged, and, even if one assumes that BRCAX has the same gene frequency as BRCA1, \sim 40% of the excess familial breast cancer risk would not be explained by rare, highly penetrant genes.

Acknowledgments

The Institute of Cancer Research receives financial support from the Cancer Research Campaign and the Medical Research Council. We thank Karen Linklater and staff of the Thames Cancer Registry for providing breast and ovarian cancer survival probabilities.

References

- Anderson KE, Easton DF, Matthews FE, Peto J (1992) Cancer mortality in the first degree relatives of young breast cancer patients. Br J Cancer 66:599-602
- Claus EB, Risch N. Thompson WD (1991) Genetic analysis of breast cancer in the cancer and steroid hormone study. Am ^J Hum Genet:48:232-242.
- Coleman M, Douglas A, Hermon C, Peto ^J (1986) Cohort study analysis with a Fortran computer program. Int J Epidemiol 15:134-137
- Cramer DW, Hutchinson GB, Welch WR, Scully RE, Ryan KJ (1983) Determinants of ovarian cancer risk. I. Reproductive experiences and family history. J Natl Cancer Inst 71:711- 716
- Easton DF, Bishop DT, Ford D, Crockford GP, and the 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, and the Breast Cancer Linkage Consortium (1995) Breast and ovarian cancer incidence in BRCA1 mutation carriers. Am ^J Hum Genet 56:265-271
- Easton DF, Matthews FE, Ford D, Swerdlow AJ, Peto J. Cancer risks to relatives of women with ovarian cancer. Int ^J Cancer (in press)
- Goldgar DE, Easton DF, Cannon-Albright LA, Skolnick MH (1994) Systematic population-based assessment of cancer risk in first-degree relatives of cancer probands. ^J Natl Cancer Inst 86:1600-1608
- Hall JM, Lee MK, Newman B, Morrow J, Anderson LA, Huey B. King M-C (1990) Linkage analysis of early onset familial breast cancer to chromosome 17q21. Science 250:1684- 1689
- Koch M, Gaedke H, Jenkins H (1989) Family history of ovarian cancer patients: A case-control study. Int ^J Epidemiol 18:782-785
- Miki Y. Swensen J, Shattuck-Eidens D, Futreal PA, Harshman K, Tavtigian S, Liu QY, et al (1994) Isolation of BRCA1, the 17q-linked breast and ovarian cancer susceptibility gene. Science 266:66-71
- Narod SA, Feunteun J, Lynch HT, Watson P. Conway T. Lynch J, Lenoir GM (1991) Familial breast-ovarian cancer locus on chromosome 17q12-23. Lancet 338:82-83
- Narod SA, Ford D, Devilee P. Barkardottir RB, Lynch HT, Smith SA, Ponder BAJ, et al (1995) An evaluation of genetic heterogeneity in ¹⁴⁵ breast-ovarian cancer families. Am ^J Hum Genet 56:254-264
- Parazzini F. La Vecchia C, Negri E, Franceschi S, Tozzi L (1993) Family history of breast, ovarian and endometrial cancer and risk of breast cancer. Int J Epidemiol 22:614- 618
- Peto J, Easton DF, Matthews F, Ford D, Swerdlow AJ. Cancer mortality in relatives of women with breast cancer. Int ^J Cancer (in press)
- Shattuck-Eidens D, McClure M, Simard J. Labrie F, Narod S, Couch F. Hoskins K, et al (1995) A collaborative survey of 80 mutations in the BRCA1 breast and ovarian cancer susceptibility gene. JAMA 273:535-541
- Schildkraut JM, Risch N. Thompson WD (1989) Evaluating genetic association among ovarian, breast, and endometrial cancer: evidence for a breast/ovarian cancer relationship. Am ^J Hum Genet 45:521-529
- Scott D, Spreadborough A, Levine E, Roberts SA (1994) Genetic predisposition in breast cancer. Lancet 344:1444
- Teare MD, Wallace SA, Harris M, Howell A, Birch JM (1994) Cancer experience in the relatives of an unselected series of breast cancer patients. Br J Cancer 70:102-111
- Tulinius H. Egilsson V, Olafsdottir GH, Sigvaldson H (1992) Risk of prostate, ovarian, and endometrial cancer among relatives of women with breast cancer. Br Med ^J 305:855- 857
- Tzonou A, Day NE, Trichopoulos D, Walker A, Saliaraki M, Papapostolou M, Polychronopoulou A (1984) The epidemiology of ovarian cancer in Greece: a case control study. Eur J Cancer Clin Oncol 20:1045-1052
- 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