

# Why Are the Majority of Hereditary Cases of Early-Onset Breast Cancer Sporadic? A Simulation Study<sup>1</sup>

Jisheng Cui and John L. Hopper<sup>2</sup>

Centre for Genetic Epidemiology, The University of Melbourne, Carlton, Victoria 3053, Australia

## Abstract

Population-based studies, including those of Ashkenazi Jews, have observed that at least 50% of women with early-onset breast cancer who carry a germ line mutation in *BRCA1* or *BRCA2* do not report a family history of the disease. That is, the majority of “hereditary” cases are “sporadic.” Furthermore, the great majority of “familial breast cancers” are not hereditary. We conducted a simulation study to evaluate the probability that a woman with early-onset breast cancer is a mutation carrier, given the number of affected relatives, for a range of plausible values of allele frequency (0.001–0.01), and increased risk in mutation carriers (5–20, equivalent to cumulative risks to age 70 of 25–70%, respectively, for Australian women). Families consisted of a case proband and her mother, sisters, and maternal and paternal grandmothers, and aunts. The numbers of sisters and aunts were generated according to Poisson distributions, and ages were assigned according to a Weibull distribution. The simulated distributions of family history and of the prevalence of mutation carriers among case probands were in general similar to those observed in population-based studies, although there was a suggestion of heterogeneity of breast cancer risk in mutation carriers. As is being observed empirically in population-based samples, a family history of breast cancer was not a strong predictor of mutation status; each affected female relative increased the risk of being a mutation carrier by only 2- to 3-fold. The probability of being a mutation carrier was generally low, except in families with extreme histories of breast cancer.

## Introduction

The localization (1, 2) and isolation (3, 4) of the two breast cancer susceptibility genes, *BRCA1* and *BRCA2*, has brought about an increased awareness of genetic predisposition to breast cancer. Women with a strong family history of breast cancer, especially if their relatives were diagnosed at a young age, are

particularly concerned about their own risk, creating demand for risk assessment, mutation testing, and genetic counseling. The early mutation-detection studies were focused on such families, and the clinical expression of a typical “*BRCA1/2* family” has evolved from those investigations.

However, recent population-based studies have challenged this paradigm. As summarized in Table 1, most incident cases of early-onset breast cancer in women with a germ line mutation in *BRCA1* or *BRCA2* do not have a first- or second-degree relative with breast cancer (5–8). Even for women of Ashkenazi Jewish descent, for whom the probability of carrying a mutation is more than 2% and the risk of breast cancer in carriers is in excess of 50% (9), the majority of community-sampled cases found to be a mutation carrier do not have an affected first degree relative (10). In other words, the majority of these “hereditary” cases are “sporadic,” in that they do not have a family history of breast cancer.

Furthermore, as indicated in Table 2, these studies have shown that even for cases whose first- or second-degree relative has breast cancer, only about 1 in 15 carried a germ line mutation in *BRCA1* or *BRCA2*. This proportion increased to about 1 in 10 if two relatives were affected, and to 1 in 3 if more than two relatives were affected. In the context of Ashkenazi Jewish women, the probability of carrying a mutation in cases with a family history is typically low unless the case was diagnosed at a young age and there was either more than one relative affected or at least one relative affected before age 50 (11). That is, the great majority of “familial breast cancers” are not hereditary, where “familial” means having a family history (12).

These population-based findings do not fit the typical image of a *BRCA1/2* family, which is popularly characterized by early-onset disease and as “high-risk” through having multiple affected members. Also, these studies have raised doubts about the “high penetrance” of these genes. It is now apparent from population-based studies in the United Kingdom and Australia that the age-specific cumulative risk to age 70 of breast cancer, averaged over the mutations-causing early-onset diseases, is about 40% (7–8), only one-half of that estimated from the atypical multiple-case families used by the BCLC<sup>3</sup> to identify the genes (13–16). A similarly lower penetrance estimate has been calculated for the average of the three founder Ashkenazi mutations (9, 11, 17).

The purpose of this investigation was to conduct simulation studies to try to understand these issues. In particular, we have calculated the conditional probability that a case (proband) is a mutation carrier given the number of her affected relatives, and the percentages of cases, and of mutation-carrying cases, by family history, for genetic models that may be appropriate for

Received 1/17/00; revised 5/10/00; accepted 5/23/00.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

<sup>1</sup> This work is supported by the National Health and Medical Research Council of Australia and by NIH Grant U01-69638.

<sup>2</sup> To whom requests for reprints should be addressed, at The University of Melbourne, Centre for Genetic Epidemiology, 200 Berkeley Street, Carlton, Victoria 3053, Australia. E-mail: j.hopper@gpph.unimelb.edu.au.

<sup>3</sup> The abbreviations used are: BCLC, Breast Cancer Linkage Consortium; ABCFS, Australian Breast Cancer Family Study; HR, hazard ratio.

Table 1 Breast cancer cases with a *BRCA1* or *BRCA2* mutation by family history

Population-based study (age at onset)	No. of 1st or 2nd degree female relatives with breast cancer				Total
	0	1	2	>2	
Australia <sup>a</sup> (<40 yrs)	13 (71%)	3 (17%)	1 (6%)	1 (6%)	18
Seattle <sup>b</sup> (<35 yrs)	4 (34%)	6 (50%)	1 (8%)	1 (8%)	12
Singapore <sup>c</sup> (<40 yrs)	3 (50%)	1 (17%)		2 (33%)	6
U.K. <sup>d</sup> (<45 yrs)	22 (73%)	3 (10%)	3 (10%)	2 (7%)	30

Population-based study (age at onset)	No. of 1st degree female relatives with breast cancer			Total
	0	1	>1	
Ashkenazi women, Washington, D.C. <sup>e</sup> (<40 yrs)	6 (67%)	3 (33%)		9
(≥40 yrs)	9 (50%)	6 (33%)	3 (17%)	18

<sup>a</sup> Hopper *et al.* (7).<sup>b</sup> *BRCA1* only. Malone *et al.* (5).<sup>c</sup> *BRCA1* only. Sng *et al.* (6).<sup>d</sup> Peto *et al.* (8).<sup>e</sup> Ashkenazi founder mutations. Hartge *et al.* (10).Table 2 Proportion of breast cancer cases who carry a mutation in *BRCA1* or *BRCA2* by family history

Population-based study (age at onset)	No. of 1st or 2nd degree female relatives with breast cancer			
	0	1	2	>2
Australia <sup>a</sup> (<40 yrs)	1 in 20	1 in 16	1 in 8	1 in 2
Seattle <sup>b</sup> (<35 yrs)	1 in 29	1 in 10	1 in 11	1 in 4
Singapore <sup>c</sup> (<40 yrs)	1 in 18	1 in 11		1 in 1
United Kingdom <sup>d</sup> (<45 yrs)	1 in 31	1 in 24	1 in 7	1 in 3

Population-based study (age at onset)	No. of 1st degree female relatives with breast cancer		
	0	1	>1
Ashkenazi women, Washington, D.C. <sup>e</sup> (<40 yrs)	1 in 5	1 in 3	
(≥40 yrs)	1 in 20	1 in 11	1 in 4

<sup>a</sup> Hopper *et al.* (7).<sup>b</sup> *BRCA1* only. Malone *et al.* (5).<sup>c</sup> *BRCA1* only. Sng *et al.* (6).<sup>d</sup> Peto *et al.* (8).<sup>e</sup> Ashkenazi founder mutations. Hartge *et al.* (10).

mutations in *BRCA1* and *BRCA2* in Ashkenazi and non-Ashkenazi populations. The size and structure of families were based on the population-based ABCFS. Breast cancer incidence rates in Australia are as high as in other Western countries.

For simplicity, we have ignored sources of familial aggregation other than dominant inheritance of high-risk mutations in autosomal loci. We have also restricted attention to breast cancer in females only. Although an increased risk of ovarian cancer in women who carry a mutation in *BRCA1* or *BRCA2* has been observed in the BCLC families (14, 15), population-based studies typically have found few cases of ovarian cancer in the families ascertained through breast cancer in a mutation carrier. Similarly, an increased risk of male breast cancer in *BRCA2* mutation carriers is evident from the BCLC families (16), but male breast cancer is not a feature of population-based families with a *BRCA2* mutation. In the population-based studies listed in Table 1, there were only 4 cases of ovarian cancer reported in the first- and second-degree relatives of the 66 mutation-carrying case probands (*i.e.*, 6% had a family history of ovarian cancer) and no cases of male breast cancer.

## Methods

**Pedigree Structure and Sampling.** Simulated pedigrees consisted of three generations based on four grandparents in the first generation. The gender of individuals in the subsequent generations was determined at random so that there was an equal chance of being male or female (see Fig. 1). For each sisterhood in the second and third generation, the number of females was assumed to be distributed as an independent Poisson variable with mean 1.25, based on census data from the Australian Bureau of Statistics (18). Each pedigree had to have at least one affected female member in the third generation, called the case proband, as indicated by the solid symbol in Fig. 1. If two or more females in the third generation were affected, one was randomly chosen to be the case proband.

**Age Distribution.** The calendar age ( $t_b$ ) of each individual, calculated from birth, was assumed to follow an independent normal distribution. For these simulation studies to be relevant to families selected through early-onset cases, the means were chosen to be 80, 60, and 35 years for the first, second, and third

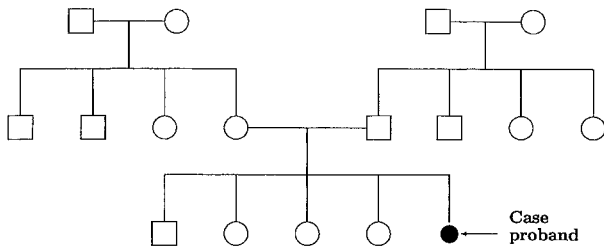


Fig. 1. The pedigree structure and sampling scheme, in which the number of female siblings is distributed as a Poisson variable.

generations, with SDs 7, 8, and 6 years, respectively, based on the ages of relatives of affected probands under the age of 40 years at the time of diagnosis observed in the ABCFS.

Following Li and Thompson (19), age at death ( $t_d$ ) was simulated according to the density function

$$f(t_d) = \frac{\beta[\ln(t_d)]^{\beta-1}}{t_d[\ln(100)]^\beta}$$

where  $1 \leq t_d \leq 100$  and  $\beta = 15$ , so as to give a median age of death of 81 years, consistent with Australian female population data (20). The minimum of calendar age and age at death was taken as the censored age  $t_c = \min(t_b, t_d)$ , where *min* stands for minimum.

**Genetic Model.** Although there are two known autosomal loci, *BRCA1* and *BRCA2*, associated with breast cancer, the probability that an individual inherits a mutation in either gene is small. For simplicity, we have ignored the very rare possibility that more than one mutation is segregating in a family. In effect, we assumed a single autosomal locus model, with two alleles *A* and *a*, where *a* represents a mutation in either *BRCA1* or *BRCA2* and has allele frequency *p*.

Mating is considered to have been at random with respect to these loci, and Hardy-Weinberg equilibrium is presumed to have existed. Therefore, the genotype of each grandparent can be represented as *AA*, *Aa*, or *aa* with probabilities  $(1 - p)^2$ ,  $2p(1 - p)$ , and  $p^2$ , respectively, independent of that of other grandparents in the first generation. For individuals in the second and third generations, their genotypes were generated by randomly and independently sampling one allele from the mother and one from the father.

This genotype-assigning process was first carried out for the grandparents. We assumed that a marriage has occurred in the second generation. If there was more than one daughter of the maternal grandparents and more than one son of the paternal grandparents, one daughter and one son were chosen at random to mate. The process was then repeated, assigning genotypes to members in the second and third generations.

The values of *p* chosen for simulations were 0.001, 0.005, and 0.01. The former two cover the range that is thought to be applicable for *BRCA1* and *BRCA2* mutations combined in a general Western population (21–25), and the latter for the combined three founder mutations in people of Ashkenazi Jewish descent (9).

For females with genotype *AA*, the hazard function (*i.e.*, the conditional probability of disease in the next age interval, given being alive at the current age *t*, measured in years) was assumed to be:

$$h_0(t) = \alpha t^{\alpha-1}$$

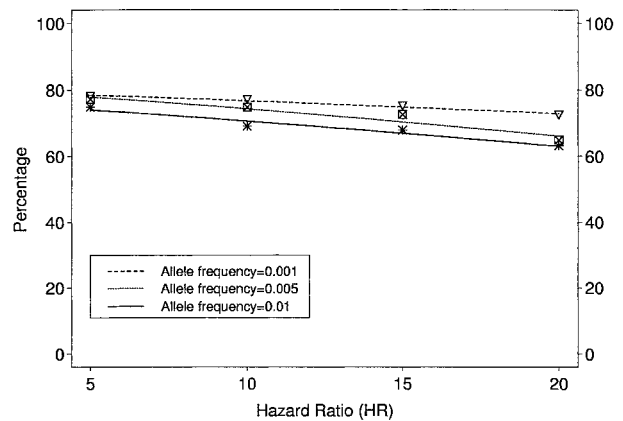


Fig. 2. Percentage of case probands with no family history.

where  $\alpha = 4.21$  and  $\lambda = 9.95 \times 10^{-10}$ . That is, the age at onset was assumed to follow a Weibull distribution, with parameters consistent with breast cancer incidence rates derived from Australian cancer registries (26). The cumulative risk was 6% to age 70.

Because we are considering hereditary cases to be those with a germ line mutation in *BRCA1* or *BRCA2*, we assumed dominant inheritance. For females with genotypes *Aa* or *aa*, the hazard function was

$$h_1(t) = h_0(t) \times \text{HR}$$

The HR, was allowed to take the values 5, 10, 15, and 20, equivalent to cumulative risks to age 70 of 25, 44, 58, and 70%, respectively. HR = 10 corresponds to the population-based estimate of cumulative risk derived from the ABCFS, and HR = 20 approximates the cumulative risk found from the analysis of the BCLC families. As in the ABCFS, the United Kingdom population-based studies also found about one-half the number of affected relatives as would be predicted by the BCLC-estimated penetrance; therefore, the value of HR = 10 would be generally appropriate for that population too. To simplify our illustration of how the distribution of family history changes with HR, we let HR be independent of age *t*, although segregation analyses and penetrance analyses of the BCLC families, but not of the relatively small number of ABCFS mutation-carrying families, suggest that HR may decrease with age.

The disease status of each female relative of a case proband was derived from the population hazard, her genotype-specific HR, and her censored age ( $t_c$ ), according to the formula:

$$\text{Pr}(\text{disease at or before age } t_c) = 1 - \exp(-\lambda_g t_c^\alpha)$$

given the presumed dominant inheritance,  $\lambda_g = \lambda$  for genotype *AA* and  $\lambda \text{HR}$  for genotype *Aa* or *aa*, where Pr stands for probability. All of the male relatives were presumed to be unaffected because this simulation study was focused on female breast cancer only.

**Results**

A total of 4000 families with at least one affected female in the third generation were generated for each of the 12 combinations of allele frequency (0.001, 0.005, and 0.01) and HR (5, 10, 15, and 20). The average numbers of sisters and of maternal and

Table 3 Percentage of case probands by the number of affected relatives, for each value of allele frequency and HR

Allele frequency	No. of affected female relatives	HR			
		5	10	15	20
0.001	0	77.8	76.7	75.0	71.9
	1	19.4	20.4	21.9	24.2
	2	2.5	2.5	2.7	3.3
	>2	0.3	0.4	0.4	0.6
0.005	0	76.7	<b>72.9<sup>a</sup></b>	71.4	64.9
	1	20.3	<b>24.0</b>	22.5	28.4
	2	2.8	<b>2.7</b>	5.5	5.6
	>2	0.3	<b>0.4</b>	0.6	1.1
0.01	0	74.9	69.1	67.9	62.8
	1	22.3	26.4	25.8	28.7
	2	2.5	4.1	5.6	6.9
	>2	0.3	0.4	0.7	1.6

<sup>a</sup> The percentages printed in bold for  $p = 0.005$  and  $HR = 10$  are used to draw Fig. 4. See text for details.

paternal aunts of case probands were 1.2, 1.3, and 1.3, respectively, and close to those observed in the ABCFS [see Hopper *et al.* (27)].

**Proportion of Probands by Family History.** Fig. 2 shows that, in all situations, the majority of case probands (60–80%) did not have a family history (*i.e.*, had no affected first- or second-degree relatives). The proportion of case probands without a family history decreased as the HR, or the allele frequency, increased. In absolute terms, however, the effect of these genetic parameters on the probability of not having a family history was not large. The continuous curves in Fig. 2 were fitted using logistic regression.

As shown in Table 3, the percentage of case probands with one affected relative was fairly stable at ~20–30%. The proportion with two affected relatives ranged between 2 and 7%, and the proportion with more than two affected relatives was between 0.3 and 1.6%. These percentages increased as the HR, or the allele frequency, increased. For  $p = 0.001$ , all of the proportions were weakly dependent on HR, in terms of both absolute and relative changes. For example, the proportion with no affected relatives varied only between 72 and 78%, whereas the proportion with two or more affected relatives varied between 0.3 and 0.6%. For  $p = 0.01$ , the proportion with no affected relatives varied from 62 to 75%, and the proportion with two or more affected relatives from 0.3 to 1.6%. That is, the effect of HR on the probability of  $n$  affected relatives increased as  $p$  increased and, in absolute and proportional terms, increased as  $n$  increased.

#### Prevalence of Mutation Carriers among Case Probands.

The prevalence of mutation carriers among case probands was small, and increased as the allele frequency ( $p$ ) or the HR increased. Moreover, the effect of HR increased as  $p$  increased. For example, for  $p = 0.001$ , the prevalence went from 0.01 to 0.02 to 0.04 as HR went from 5 to 10 to 20, whereas for  $p = 0.01$ , the corresponding prevalences were 0.1, 0.16, and 0.28. The prevalence did not exceed 5% when  $p = 0.001$ , was less than 10% when  $p = 0.005$  and  $HR < 10$ , and was at a maximum of only 30% when  $p = 0.01$  and  $HR = 20$ .

**Prevalence of Mutation Carriers among Case Probands by Family History.** Fig. 3 shows the prevalence of mutation carriers among case probands classified by their family history, where the continuous curve represents a fitted logistic regression model that included linear and quadratic terms for HR. The

prevalence increased as the allele frequency, extent of family history, or HR increased. For  $p = 0.001$ , the prevalence of mutation carriers among case probands without a family history was between 1 and 2%. For  $p = 0.01$ , the prevalence was between 8 and 14%. The prevalence of mutation carriers among probands with more than two affected relatives ranged from 6 to 50% for  $p = 0.001$ , and from 40 to 90% for  $p = 0.01$ .

For  $p = 0.001$ , the prevalence of mutation carriers among case probands with one affected relative was below 10%. Even when HR was as large as 20, the prevalence in case probands who had two or more affected relatives was less than 50%.

When  $p = 0.005$  and HR was between 5 and 10, the prevalence of mutation carriers among case probands was below 5% for cases with no family history. As the number of affected relatives increased from one to two and to more than two, the prevalence increased from between 8 and 15%, to between 15 and 30% and to between 20 and 50%, respectively. That is, each affected female relative increased the risk of being a mutation carrier by about 2- to 3-fold (11, 28).

#### Percentage of Mutation Carriers among Case Probands by the Number of Affected Relatives.

Shown in Table 4 are the percentage of mutation carriers among case probands by the number of affected relatives, for each fixed allele frequency and HR. We see that the percentage of mutation-carrying probands without a family history decreased as the HR increased. This percentage declined, from 62% when  $HR = 5$  and  $p = 0.001$ , to 30% when  $HR = 20$  and  $p = 0.01$ . When  $HR \leq 10$ , more than 50% of mutation carriers among case probands had no family history for  $0.001 \leq p \leq 0.01$ .

**Tower of Breast Cancer Genetics.** Figure 1 of Hopper *et al.* (27) summarized findings from mutation testing in 400 cases of early-onset breast cancer by a “tower” with a broad base representing sporadic cases (*i.e.*, cases with no family history). Familial cases (those who had at least one affected first- or second-degree female relative) were represented above the base, with the different categories of family history labeled on the left-hand side of the figure. The tower became progressively narrower as the number of affected relatives increased. The sizes of the sporadic and familial boxes were proportional to the breakdown of cases by family history in the population-based sample. Hereditary cases (*i.e.*, cases with a *BRCA1* or *BRCA2* mutation) were represented by a vertical core through the middle of the rectangles, with a fixed width of one unit. Therefore, as the width of the column of the tower decreased with increasing numbers of affected relatives, the proportion of hereditary cases for each category—written in terms of “1 in  $x$ ” in the right-hand half of each box—also increased with increasing family history. The percentage on the right-hand side of the figure represents the proportion of all of the mutation carriers among the case probands, broken down for each category of family history.

Fig. 4 shows the simulated tower for  $HR = 10$  and  $p = 0.005$ , the genetic parameters compatible with the Australian population-based data on which Figure 1 of Hopper *et al.* (27) was based. The relative sizes of the boxes were based on the percentages of case probands by the number of affected relatives given in bold for  $HR = 10$  and  $p = 0.005$  in Table 3. The percentages of mutation carriers among case probands for each category of family history listed on the right-hand side are derived from the numbers given in bold font in Table 4. The proportions of carriers among case probands for each category of family history, written as “1 in  $x$ ,” are derived from Fig. 3,  $p = 0.005$  and  $HR = 10$ . Using the same approach, towers for other scenarios can be drawn (not shown).



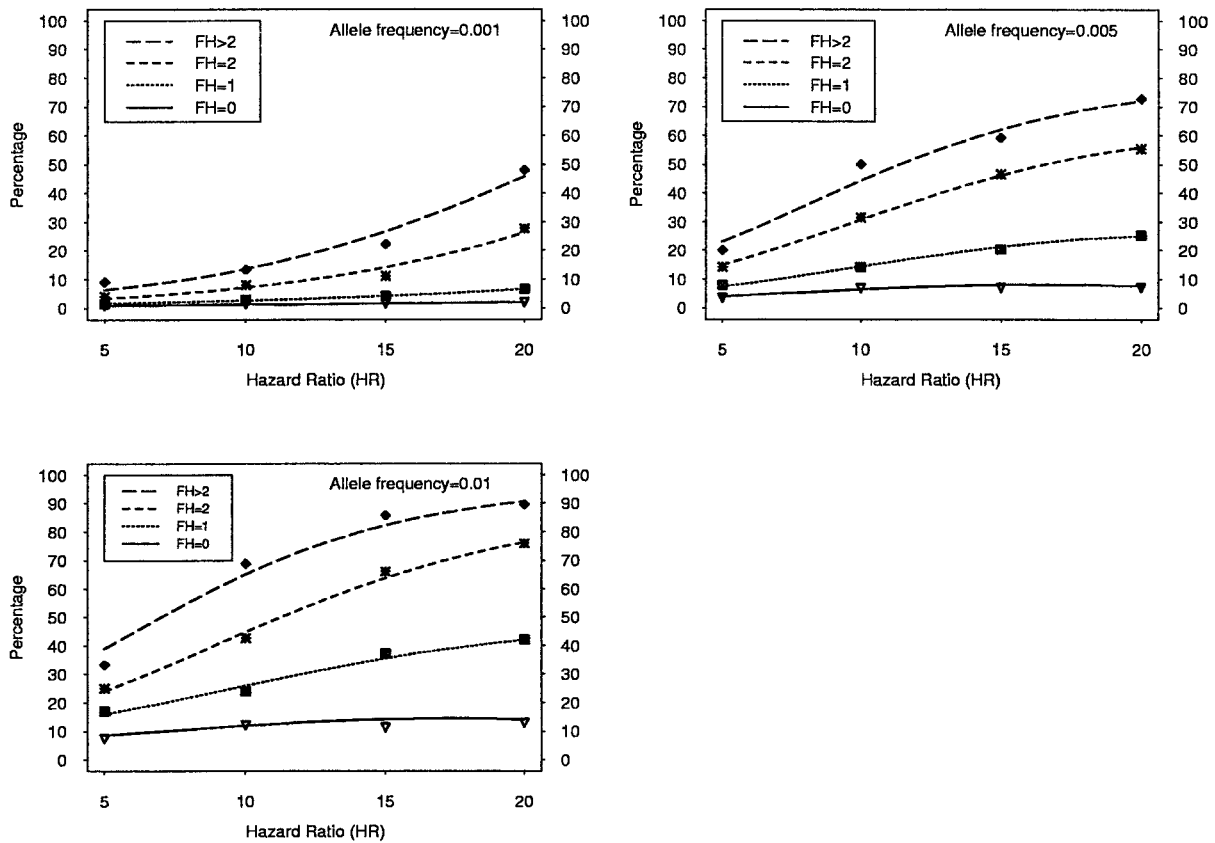


Fig. 3. Prevalence of mutation carriers among case probands by the number of affected first- or second-degree relatives.

Table 4 Percentage of mutation carriers among case probands by the number of affected relatives, for each value of allele frequency and HR

Allele frequency	No. of affected female relatives	HR			
		5	10	15	20
0.001	0	62.2	54.1	46.4	33.3
	1	26.7	32.4	37.1	38.1
	2	8.9	10.8	12.4	21.4
	>2	2.2	2.7	4.1	7.2
0.005	0	56.8	<b>52.7<sup>a</sup></b>	39.8	29.5
	1	33.7	<b>36.2</b>	36.5	45.7
	2	8.4	<b>9.0</b>	20.5	19.7
	>2	1.1	<b>2.1</b>	3.2	5.1
0.01	0	55.3	50.0	35.7	30.1
	1	37.5	38.0	44.4	45.0
	2	6.2	10.5	17.1	19.3
	>2	1.0	1.5	2.8	5.6

<sup>a</sup> The percentages printed in bold for  $p = 0.005$  and  $HR = 10$  are used to draw Fig. 4. See text (p. 808) for details.

Fig. 4 shows, therefore, that the majority of hereditary cases (*i.e.*, 53%) had no family history, and that only 2% of hereditary case probands had more than two affected relatives, and only 11% had two or more affected. Comparison with Table 1 shows that the observed proportion with no affected relatives (typically 50% or more and on average 64%) is similar to that predicted by the model with  $HR = 10$ .

Where the model appears not to fit the observed data well is in the smallest category of two or more affected relatives, into which 6 (9%) of the 66 mutation-carrying case families were classified, compared with the model's prediction of just 1.3 (2%). That is, there appear to be too many mutation-carrying families with more than two affected female first- or second-degree relatives than would be expected if the HR were 10 and independent of age.

In addition, Fig. 4 shows that the probability that a case carries a mutation approximately doubles or triples for each additional affected relative. This is similar to the effect being observed empirically, both for population-based studies (see Table 2) and for studies of women of Ashkenazi Jewish descent (10, 11, 28).

In contrast, using the values  $HR = 20$  and  $p = 0.005$  appropriate for the multiple-case families of the BCLC, Table 4 shows that only 29% did not have a family history, inconsistent with the observations listed in Table 1. The model also predicts that at the other end of the family history distribution, 5% would have at least two affected relatives, and that 25% would have two or more affected relatives, more comparable with the observed 11 (17%) of the population-based families in Table 1.

When we conducted simulations restricting to only first-degree relatives and used the parameters appropriate for the combined Ashkenazi founder mutations ( $HR = 15$ ,  $p = 0.01$ ), we found that about 65% of case carriers did not have an affected first-degree relative, and 32% had one affected rela-

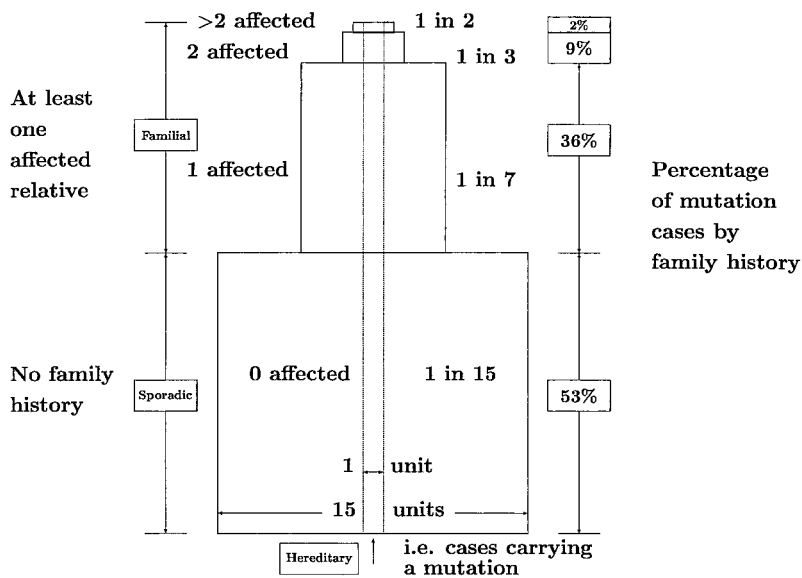


Fig. 4. Tower of breast cancer genetics when allele frequency  $p = 0.005$  and  $HR = 10$ . Women with breast cancer categorized according to the number of affected first- or second-degree female relatives (0, 1, 2, or >2). The relative areas of the boxes indicate the proportion of all those affected in each category of family history. Familial cases have one or more affected first- or second-degree relatives. Sporadic cases have no family history. Hereditary cases have a germ line mutation in either *BRCA1* or *BRCA2* and can occur in any category. The width of the hereditary column is 1 unit. The proportion of hereditary cases in each category of family history, in terms of 1 in  $x$ , where  $x$  is the width of the tower for that category, is shown in the right-hand half of each box. The proportion of all of the hereditary cases for each category of family history, in terms of a percentage, is shown on the right-hand side of the figure.

tive, close to that observed from the community-based Ashkenazi study (see Table 1).

## Discussion

In general, using  $HR = 10$  and  $p = 0.005$ , we found that the simulated breakdowns by family history of case probands and of mutation carriers among case probands were similar to those observed in the population-based studies, from which those values of the genetic parameters were estimated, or deemed appropriate from other sources of information. The shape of the tower given in Fig. 4 is similar to that derived from the ABCFS (27), although it perhaps varies in the extreme category from that suggested by the pooling of the population-based studies in Table 2.

Whereas the population-based estimate of  $HR = 10$  best describes the mutation status of the majority of cases, for whom either no relative or one relative is affected (*base of the tower*), it appears to underestimate the proportion with two or more affected relatives (*top of the tower*). In contrast, the BCLC-based estimate of  $HR = 20$  appears to describe the mutation status of cases with multiple affected relatives better but gives poor predictions for the majority of cases. That is, for the vast majority of mutation carriers, the observed data may be best described by a model with  $HR = 10$  or even less [when one considers the proportion of mutation carriers among cases with no family history (see Table 2)], but there may also be a small subset for whom  $HR = 20$  is appropriate. It is tempting, therefore, to suggest that there is heterogeneity of risk. This could be attributable to there being classes of mutations associated with different risks, or that there are other familial risk “triggers” (environmental or genetic) that modify the risk in mutation carriers. Clearly, larger numbers of population-sampled case carriers, and further simulations on models that allow for heterogeneity and age-dependence of risk, are needed to clarify this issue.

These simulation studies have confirmed that in the population setting, even for  $HR = 20$ , family history of breast cancer is not a strong predictor of mutation status, which is consistent with the observations of population-based studies (7). Each affected relative increased the risk of being a mutation

carrier by  $\sim 2$ - to 3-fold (11, 28). The probability of being a mutation carrier was generally low, except in families with extreme histories of breast cancer, such as those deliberately ascertained for gene-hunting by the BCLC.

**Why Are the Majority of Hereditary Cases of Early-onset Breast Cancer Sporadic?** First, there may not be many female mutation carriers in the family. Their probability of developing breast cancer is not necessarily high, especially when their age is taken into consideration. It now appears that the average increased risk of breast cancer attributable to mutations that cause early-onset breast cancer in the population is about 10-fold, and, hence, the lifetime risk is less than 50%. Consequently, more than one-half of female mutation carriers in a family are unlikely to be affected, especially if they are not old. Furthermore, on average, only one in every two female first-degree relatives of a mutation carrier is a carrier herself, and mutations can be passed down through the paternal line(s) so there may be no female carriers in the family.

The second issue is the potential for under-reporting of affected relatives in mutation carrying cases found by the population-based studies. The population-based studies listed in Table 1, however, have made major efforts to record the family histories of the few case carriers that they have identified, and they have even systematically sought an interview with first- and second-degree relatives as in the ABCFS. Nevertheless, there are some families for whom there is little knowledge about the disease status of relatives, especially for second-degree relatives.

The third issue is *de novo* mutations, which would increase the number of cases without a family history. However there is little evidence that this is a common occurrence for *BRCA1* and *BRCA2* (as it is for the *APC* gene). One instance of a *de novo* mutation in *BRCA1* has been identified by the ABCFS (29), and we are aware of one other anecdotal report.

The fourth factor that may affect the interpretation of our results is the sensitivity of testing for *BRCA1* and *BRCA2* mutations which, in practice, has yet to reach 100%. Some mutation carriers may not have been detected, and, although this would result in a lower prevalence of mutation carriers among case probands, the breakdown by family history would

be altered only if the mutations detected were associated with a different risk of disease than those not detected.

The fifth factor is the risk of ovarian cancer in mutation-carrying relatives. Because of the poor prognosis in those who develop ovarian cancer, the probability that such carriers would also develop breast cancer would be low. Contrary to the experience of the BCLC, which had specifically over-sampled “breast-ovary” families, ovarian cancers are not a common feature of *BRCA1* or *BRCA2* mutation-carrying families as ascertained through a population-sampled case. Only 6% of such families listed in Table 1 have been observed to have a family history of ovarian cancer. That is, whereas ovarian cancer may be a predictor of the presence of a *BRCA1* or even a *BRCA2* mutation, especially when it occurs in a family with multiple-cases of breast cancer, it is not a typical feature of mutation-carrying families in the population.

Family size is another potentially important variable. In this study, the mean number of female siblings in the second and third generation was assumed to be 1.25. However, when we increased this to 1.8, the percentage of mutation carriers who did not have a family history only went from 53 to 59%, when  $p = 0.005$  and  $HR = 10$ .

Although the findings presented in this report match recent population-based studies reasonably well, caution should be exercised when interpreting our results. We have assumed that HR at all ages is a constant, which represents an averaged genetic relative risk. To study the effects that a heterogeneity of risk between the young and the old would have on the family history distribution, further simulations should allow for HR to depend on a woman’s age.

We assessed the reliability of the percentages in Table 3 and Table 4 by calculating approximate SEs and confidence intervals based on the analytical variance estimate from the binomial distribution (30). For example, in Table 3, 72.9% of case probands did not have a family history, and 0.4% had more than two affected female relatives, when  $p = 0.005$  and  $HR = 10$ . Because these estimates came from 4000 simulated families, their SEs were approximately 0.7% and 0.1%, respectively. Accordingly, the approximate 95% confidence intervals were from 71.5 to 74.3% and from 0.2 to 0.6%, respectively. Similarly, in Table 4, 52.7 and 2.1% of case probands who were mutation carriers had no family history and had more than two affected female relatives, respectively. Because these estimates came from 2916 ( $4000 \times 72.9\%$ ) and 16 ( $4000 \times 2.1\%$ ) families, the 95% confidence intervals were 50.9–54.5% and 0–9.2%, respectively.

These simulations have shown that the findings of the population-based studies are not inconsistent with a genetic model for *BRCA1* and *BRCA2* mutations. They are reassuring, because early conference presentations and submitted manuscripts that were based on these findings (*i.e.*, Refs. 31 and 7, respectively) were met by considerable skepticism. Both the observed and predicted towers have a large and high base, and the great majority of case carriers have no affected relative or only one. Therefore, the multiple-case families who present to cancer genetics clinics, represented by the box at the top of the tower, constitute only a small proportion of all mutation carriers. It is, therefore, likely that only a small percentage of all mutation carriers will ever be detected by current mutation detection, which is almost exclusively limited to the so-called high-risk families. Although a small percent of women with breast cancer may carry a mutation in *BRCA1* or *BRCA2*, the prospect of detecting more than a small percentage of them is not realistic given the current cost and acceptability of mutation testing. The population-based perspective of breast cancer ge-

netics is quite different from that provided by the earlier studies, which focused on families with multiple cases of breast cancer.

## References

- Hall, J. M., Lee, M. K., Newman, B., Morrow, J. E., Anderson, L. A., Huey, B., and King, M. C. Linkage of early-onset familial breast cancer to chromosome 17q21. *Science* (Washington DC), 250: 1684–1689, 1990.
- Wooster, R., Neuhausen, S. L., Mangion, J., Quirk, Y., Ford, D., Collins, N., Nguyen, K., Seal, S., Tran, T., Averill, D., Fields, P., Marshall, G., Narod, S., Lenoir, G. M., Lynch, H., Feunteun, J., Devilee, P., Cornelisse, C. J., Menko, F. H., Daly, P. A., Ormiston, W., McManus, R., Pye, C., Lewis, C. M., Cannon-Albright, L. A., Peto, J., Ponder, B. A. J., Skolnick, M. H., Easton, D. F., Goldgar, D. E., and Stratton, M. R. Localization of a breast cancer susceptibility gene, *BRCA2*, to chromosome 13q12-13. *Science* (Washington DC), 265: 2088–2090, 1994.
- Miki, Y., Swensen, J., Shattuck-Eidens, D., Futreal, P. A., Harshman, K., Tavtigian, S., Liu, Q., Cochran, C., Bennett, L. M., Ding, W., Bell, R., Rosenthal, J., Hussey, C., Tran, T., McClure, M., Frye, C., Hattier, T., Phelps, R., Haugen-Strano, A., Katcher, H., Yakumo, K., Gholami, Z., Shaffer, D., Stone, S., Bayer, S., Wray, C., Bogden, R., Dayananth, P., Ward, J., Tonin, P., Narod, S., Bristow, P. K., Norris, F. H., Helvering, L., Morrison, P., Rostock, P., Lai, M., Barrett, J. C., Lewis, C., Neuhausen, S., Cannon-Albright, L., Goldgar, D., Wiseman, R., Kamb, A., and Skolnick, M. H. A strong candidate for the breast and ovarian cancer susceptibility gene *BRCA1*. *Science* (Washington DC) 266: 66–71, 1994.
- Wooster, R., Bignell, G., Lancaster, J., Swift, S., Seal, S., Mangion, J., Collins, N., Gregory, S., Gumbs, C., Micklem, G., Barfoot, R., Hamoudi, R., Patel, S., Rice, C., Biggs, P., Hashim, Y., Smith, A., Connor, F., Arason, A., Gudmundsson, J., Ficenc, D., Kelsell, D., Ford, D., Tonin, P., Bishop, D. T., Spurr, N. K., Ponder, B. A. J., Eeles, R., Peto, J., Devilee, P., Cornelisse, C., Lynch, H., Narod, S., Lenoir, G., Egilsson, V., Barkadottir, R. B., Easton, D. F., Bentley, D. R., Futreal, P. A., Ashworth, A., and Stratton, M. R. Identification of the breast-cancer susceptibility gene *BRCA2*. *Nature* (Lond.), 378: 789–792, 1995.
- Malone, K. E., Daling, J. R., Thompson, J. D., O’Brien, C. A., Francisco, L. V., and Ostrander, E. A. *BRCA1* mutations and breast cancer in the general population—analyses in women before age 35 years and in women before age 45 years with first-degree family history. *J. Am. Med. Assoc.*, 279: 922–929, 1998.
- Sng, J-H., Chang, J., Feroze, F., Rahman, N., Tan, W., Lim, S., Lehnert, M., van der Pool, S., and Wong, J. The prevalence of *BRCA1* mutations in Chinese patients with early onset breast cancer and affected relatives. *Br. J. Cancer*, 82: 538–542, 2000.
- Hopper, J. L., Southey, M. C., Dite, G. S., Jolley, D. J., Giles, G. G., McCredie, M. R. E., Easton, D. F., Venter, D. J., and the Australian Breast Cancer Family Study. Population-based estimate of the average age-specific cumulative risk of breast cancer for a defined set of protein-truncating mutations in *BRCA1* and *BRCA2*. *Cancer Epidemiol. Biomark. Prev.*, 8: 741–747, 1999.
- Peto, J., Collins, N., Barfoot, R., Seal, S., Warren, W., Rahman, N., Easton, D. F., Evans, C., Deacon, J., and Stratton, M. R. Prevalence of *BRCA1* and *BRCA2* gene mutations in patients with early-onset breast cancer. *J. Natl. Cancer Inst.*, 91: 943–949, 1999.
- Struwing, J. P., Hartge, P., Wacholder, S., Baker, S. M., Berlin, M., McAdams, M., Timmerman, M. M., Brody, L. C., and Tucker, M. The risk of cancer associated with specific mutation of *BRCA1* and *BRCA2* among Ashkenazi Jews. *N. Engl. J. Med.*, 336: 1401–1408, 1997.
- Hartge, P., Struwing, J. P., Wacholder, S., Brody, L. C., and Tucker, M. A. The prevalence of common *BRCA1* and *BRCA2* mutations among Ashkenazi Jews. *Am. J. Hum. Genet.*, 64: 963–970, 1999.
- Hopper, J. L., and Jenkins, M. A. Modeling the probability that Ashkenazi Jewish mutation carry a founder mutation in *BRCA1* or *BRCA2*. *Am. J. Hum. Genet.*, 65: 1775–1776, 1999.
- Lynch, H. T., and Hirayama, T. *Genetic Epidemiology of Cancer*, p. 310. Boca Raton, FL: CRC Press, Inc., 1989.
- Easton, D. F., Ford, D., Bishop, D. T., Crockford, G. P., and the Breast Cancer Linkage Consortium. Genetic linkage analysis in familial breast and ovarian cancer: results from 214 families. *Am. J. Hum. Genet.*, 52: 678–701, 1993.
- Narod, S., Ford, D., Devilee, P., Barkadottir, R. B., Eyfjord, J., Lenoir, G., Serova, O., Easton, D., Goldgar, D., and the Breast Cancer Linkage Consortium. Genetic heterogeneity of breast-ovarian cancer revisited. *Am. J. Hum. Genet.*, 57: 957–958, 1995.
- Ford, D., Easton, D. F., Stratton, M., Narod, S., Goldgar, D., Devilee, P., Bishop, D. T., Weber, B., Lenoir, G., Chang-Claude, J., Sobol, H., Teare, M. D., Struwing, J., Arason, A., Scherneck, S., Peto, J., Rebbeck, T. R., Tonin, P., Neuhausen, S., Barkadottir, R., Eyfjord, J., Lynch, H., Ponder, A. J., Gayther,

- S. A., Birch, J. M., Lindblom, A., Stoppa-Lyonnet, D., Bignon, Y., Borg, A., Hamann, U., Haites, N., Scott, R. J., Maugard, C. M., Vasen, H., Seitz, S., Cannon-Albright, L. A., Schofield, A., Zelada-Hedman, M., and the Breast Cancer Linkage Consortium. Genetic heterogeneity and penetrance analysis of the *BRCA1* and *BRCA2* genes in breast cancer families. *Am. J. Hum. Genet.*, *62*: 676–689, 1998.
16. The Breast Cancer Linkage Consortium. Cancer risks in *BRCA2* mutation carriers. *J. Natl. Cancer Inst.*, *91*: 1310–1316, 1999.
17. Fodor, F. H., Weston, A., Bleiweiss, I. J., McCurdy, L. D., Walsh, M. M., Tartter, P. I., Brower, S. T., and Eng, C. M. Frequency and carrier risk associated with common *BRCA1* and *BRCA2* mutations in Ashkenazi Jewish breast cancer patients. *Am. J. Hum. Genet.*, *63*: 45–51, 1998.
18. Australian Bureau of Statistics. Births, Australia, 1998. Canberra: Australian Bureau of Statistics, 1998.
19. Li, H., and Thompson, E. A. Semiparametric estimation of major gene and family-specific random effects for age of onset. *Biometrics*, *53*: 282–293, 1997.
20. Australian Bureau of Statistics. Deaths, Australia, 1997. Canberra: Australian Bureau of Statistics, 1997.
21. Claus, E. B., Risch, N. J., and Thompson, W. D. Genetic analysis of breast cancer in the Cancer and Steroid Hormone Study. *Am. J. Hum. Genet.*, *48*: 232–242, 1991.
22. Whittemore, A. S., Gong, G., and Itnyre, J. Prevalence and contribution of *BRCA1* mutations in breast cancer and ovarian cancer: results from three U. S. population-based case-control studies of ovarian cancer. *Am. J. Hum. Genet.*, *60*: 496–504, 1997.
23. Easton, D. F., Matthew, F. E., Ford, D., Swerdlow, A. J., and Peto, J. Cancer mortality in relatives of women with ovarian cancer: the OPCS study. *Int. J. Cancer*, *65*: 284–294, 1996.
24. Peto, J., Easton, D. F., Matthews, F. E., Ford, D., and Swerdlow, A. J. Cancer mortality in relatives or women with breast cancer: the OPCS study. *Int. J. Cancer*, *65*: 275–283, 1996.
25. Ford, D., Easton, D. F., and Peto, J. Estimates of the gene frequency of *BRCA1* and its contribution to breast and ovarian cancer incidence. *Am. J. Hum. Genet.*, *57*: 1457–1462, 1995.
26. McCredie, M. R. E., Hopper, J. L., and Cawson, J. N. Risk factors and preventive strategies for breast cancer. *Med. J. Aust.*, *163*: 435–440, 1995.
27. Hopper, J. L., Chenevix-Trench, G., Jolley, D. J., Dite, G. S., Jenkins, M. A., Venter, D. J., and McCredie, M. R. E. Design and analysis issues in a population-based, case-control-family study of the genetic epidemiology of breast cancer and the Co-operative Family Registry for Breast Cancer Studies (CFRBCS). *J. Natl. Cancer Inst. Monogr.*, *26*: 95–100, 1999.
28. Foulkes, W. D., Brunet, J.-S., Warner, E., Goodwin, P. J., Meschino, W., Narod, S. A., Goss, P. E., and Glendon, G. The importance of a family history of breast cancer in predicting the presence of a *BRCA* mutation. *Am. J. Hum. Genet.*, *65*: 1779–1782, 1999.
29. Tesoriero, A., Andersen, C., Southey, M., Somers, G., McKay, M., Armes, J., McCredie, M., Giles, G., Hopper, J. L., and Venter, D. J. *de novo BRCA1* mutation in a patient with breast cancer and an inherited *BRCA2* mutation. *Am. J. Hum. Genet.*, *65*: 567–569, 1999.
30. Hogg, R. V., and Craig, A. T. Introduction to mathematical statistics. Eaglewood Cliffs, NJ: Prentice-Hall, Inc., 1995.
31. Hopper, J. L., McCredie, M. R. E., Giles, G. G., Dite, G. S., Southey, M. C., and Venter, D. J. The Australian breast cancer family study: putting breast cancer genes into perspective. *Am. J. Hum. Genet.*, *61* (Suppl.): A15, 1997.