

Letters to the Editor

Am. J. Hum. Genet. 52:627–629, 1993

Is There Heterogeneity of Age at Onset for Breast Cancer?

To the Editor:

Recently, Margaritte et al. (1992) challenged portions of a linkage analysis of breast cancer families by Hall et al. (1990). Specifically, they argued that the inverse relationship between age at onset and lod scores observed by Hall et al. (1990) is an artifact of the erroneous use of age-dependent penetrance functions: for affected individuals with known age at disease onset, the density function, rather than the distribution function, of onset age should be used, which was what Hall et al. (1990) had failed to do. Margaritte et al. (1992) show by means of a clever argument that, as a consequence of this omission, sporadic cases with high age at onset have a tendency to be misclassified as inherited cases. These false-positive genetic cases introduce false recombinants, which make a family with older cases falsely look unlinked. In a previous study we also reanalyzed the data of Hall et al. (1990) but arrived at a conclusion opposite that of Margaritte et al. (1992)—namely, that genetic cases tend to be restricted to families with early age at onset, with a significance that was even stronger than that obtained by Hall et al. (1990). Here we wish to explore the reasons for this discrepancy.

The main point of Margaritte et al. (1992) is that, in a linkage analysis for a disease with age-dependent penetrance, it is the density function of age at onset that should be used for affected individuals (at age at disease onset), and the cumulative density (the distribution function) should be used for unaffected individuals (at current age). This distinction has been well known for a number of years (Elston 1973), but linkage analysts of-

ten do not seem to be aware of it (Ott 1991) and do not provide for separate liability classes for affected and unaffected individuals. For ease of use, age-at-onset distributions were several years ago incorporated into the LIPED program (Ott 1991, p. 160), such that appropriate curves are automatically used for affected and unaffected individuals. Thus, the LIPED program does *not* work only with cumulative incidence, as claimed by Hall et al. (1992, p. 1236).

We agree, of course, with Margaritte et al. (1992) on their main point, except that their formulas are correct only for complete penetrance, whereas, in breast cancer, a lifetime penetrance (i.e., penetrance at high age) of less than 100% is generally assumed (for formulas with incomplete lifetime penetrance, see Elston 1973, 1981; Ott 1991). To elucidate the reasons for the discrepant conclusions of Margaritte et al. (1992) and Mérette et al. (1992), we duplicated the construction of rates of sporadic cases (random affected individuals) as carried out in the former paper, but we used our own parameterization for age at onset.

In our analysis, in families of the linked type, we used a normal density of age at onset, with means $\mu_{Dd} = 43.0$ and $\mu_{dd} = 55.5$ for Dd and dd genotypes, respectively, where D is the disease allele; the common SD was estimated as $\sigma = 13$. Lifetime penetrances were taken to be $p = .82$ for Dd genotypes and $t = .081$ for dd genotypes (Newman et al. 1988). With this parameterization, the relative probability that a random affected individual is a sporadic case versus a genetic case (Margaritte et al. 1992) is given by $R(x) = K(p/t) \phi(x; \mu_{dd}, \sigma^2) / \phi(x; \mu_{Dd}, \sigma^2)$, where ϕ is the normal density, x is the given individual's age at onset, K stands for $(1 - q)^2 / [q(2 - q)]$, and q is the disease gene frequency. As table

Table 1**Relative Probability, $R(x)$, of Sporadic versus Inherited Breast Cancer, for Affected Women**

AGE (years)	$R(x)$ FOR					
	Linked Families			Unlinked Families		
	Hall et al. (1990) ^a	Mérette et al. (1992) ^b	Margaritte et al. (1992) ^c	Hall et al. (1990) ^a	Mérette et al. (1992) ^d	Margaritte et al. (1992) ^c
2753	.93	.53	.53	4.86	.53
47	2.08	4.12	4.06	2.08	4.86	4.06
67	4.84	18.08	16.23	4.84	4.86	16.23

^a $R(x=67)/R(x=27) = 9.1$.^b $R(x=67)/R(x=27) = 19.4$.^c $R(x=67)/R(x=27) = 30.6$.^d $R(x=67)/R(x=27) = 1$.

1 shows, for linked families, our values of $R(x)$ for the three age classes in Margaritte et al. (1992) are similar to those obtained by Margaritte et al. (1992), even though we allow for incomplete lifetime penetrance.

When the disease in a family is unlinked to the marker under study, we assume the same elevated mean age at onset, irrespective of the genotype, at some disease-causing locus other than the one linked to the marker under study. Therefore, our relative probability, $R(x) = Kp/t$, is constant over genotypes, whereas, in Margaritte et al. (1992), $R(x)$ depends on the genotypes, as it does for linked families.

On the basis of the facts outlined above, we see essentially two important differences between the three analyses discussed here. The first difference relates to a comparison between age classes within linked families (see table 1). As shown by Margaritte et al. (1992), the relative probability $R(x)$ of being a sporadic case versus a genetic case is too low in older age classes of Hall et al. (1990), which may create false genetic cases, and, thus, false recombinants in families with older cases, thus leading to false evidence for heterogeneity. However, as seen by the ratio of $R(x = 67 \text{ years})$ versus $R(x = 27 \text{ years})$ in table 1, Margaritte et al. (1992) seem to have overcorrected the increase in the relative probability of being a sporadic case, such that they were unable to detect the heterogeneity in the data. This overcorrection may be due to their failure to allow for incomplete penetrance in genetic and nongenetic cases.

The second important difference between the three analyses relates to a comparison between linked and unlinked families. In contrast to the other investigators, we allow for an elevated mean age at onset for genetic

cases in unlinked families, such that all genotypes in unlinked families have the same (high) mean age at onset. If some families have a reduced recombination rate *and* decreased mean age at onset, these two effects jointly lead to stronger heterogeneity than does reduced recombination frequency (linkage) alone. However, a difference in mean age at onset was not even considered in the analysis by Margaritte et al. (1992). Therefore, the question of an age-at-onset difference between linked and unlinked families is not strictly addressed in their analysis. We agree with their explanation for the difference between their results and Hall et al.'s (1990) results, but that explanation cannot serve as the basis for a claim of no age-at-onset difference between linked and unlinked families.

The graph (fig. 1 in Mérette et al. 1992) of mean age at onset per family versus estimated recombination fraction in Hall et al.'s (1990) data clearly suggests a relationship between the two quantities. Hall et al. (1990) essentially came to the correct conclusion, even though (a) their analysis did not allow for reduced age at onset in linked families and (b) they did not use age at disease onset (but only age at diagnosis) for affected individuals. Hall et al.'s (1990) analysis was not, in principle, wrong but, rather, was incomplete—they neglected to use some of the information in the data. Their analysis is correct for the situation in which age at disease onset is not known for each case considered.

JURG OTT AND CHANTAL MÉRETTE

Columbia University
New York

References

- Elston RC (1973) Ascertainment and age of onset in pedigree analysis. *Hum Hered* 23:105–112
- (1981) Segregation analysis. In: Harris H, Hirschhorn K (eds) *Advances in human genetics*, vol 11. Plenum, New York, pp 63–120
- Hall JM, Friedman L, Guenther C, Lee MK, Weber JL, Black DM, King M-C (1992) Closing in on a breast cancer gene on chromosome 17q. *Am J Hum Genet* 50:1235–1242
- Hall JM, Lee MK, Newman B, Morrow JE, Anderson LA, Huey B, King M-C (1990) Linkage of early-onset familial breast cancer to chromosome 17q21. *Science* 250:1684–1689
- Margaritte P, Bonaiti-Pellie C, King M-C, Clerget-Darpoux F (1992) Linkage of familial breast cancer to chromosome 17q21 may not be restricted to early-onset disease. *Am J Hum Genet* 50:1231–1234
- Mérette C, King M-C, Ott J (1992) Heterogeneity analysis of breast cancer families by using age at onset as a covariate. *Am J Hum Genet* 50:515–519
- Newman B, Austin MA, Lee M, King M-C (1988) Inheritance of human breast cancer: evidence for autosomal dominant transmission in high risk families. *Proc Natl Acad Sci USA* 85:3044–3048
- Ott J (1991) *Analysis of human genetic linkage*, 2d ed. Johns Hopkins University Press, Baltimore

© 1993 by The American Society of Human Genetics. All rights reserved.
0002-9297/93/5203-0020\$02.00

Am. J. Hum. Genet. 52:629–630, 1993

Reply to Ott and Mérette

To the Editor:

In their letter, Ott and Mérette propose an explanation for the discrepancies between their results (Mérette et al. 1992) and ours (Margaritte et al. 1992), with regard to the analysis of cancer family data reported by Hall et al. (1990). The reason invoked is our study's "failure of allowing for incomplete penetrance in genetic and non-genetic cases." It is certainly not the correct reason, since, in fact, in our analysis, the lifetime penetrances were set to .82 for the gene carriers and .081 for the noncarriers. These values are the same as those used by Mérette et al. (1992) and correspond to the estimations by Newman et al. (1988). Besides, our formulas $R(x)$ and $R'(x)$ do not require any assumption about the lifetime penetrance values, since they were established by

using directly the incidence functions obtained from Newman et al. (1988). On the contrary, Mérette et al. modeled the incidence through a normal density function, which consequently has to be multiplied by the lifetime penetrance.

In our view, the discrepancy is probably due to the two studies' different assumptions about the age-at-onset distributions. Mérette et al. not only assumed normal distributions for inherited and sporadic cases but also fixed a mean age at onset $\mu_{dd} = 55.5$ years in sporadic cases, whereas we just used the step functions provided by Newman et al. (1988). The conclusion of a linkage homogeneity test is valid only on the condition that the assumptions are valid. In particular, the conclusion that there are two age-at-onset distributions among inherited cases in Mérette et al. (1992), as well as the results of table 1 in the letter by Ott and Mérette, depends on the correctness of the μ_{dd} value. Note that Mérette et al. (1992) estimated this value on the basis of data (Mettlin et al. 1990) other than those that Newman et al. (1988) used for estimating the lifetime penetrances. This value (55.5 years) is surprisingly low, compared with values published in the literature (68.99 in Claus et al. 1991). Furthermore, under their assumption of normality, this value implies that half the sporadic cases would have an age at onset that is more than 55.5 years. This is not compatible with the step functions of Newman et al. (1988), which predict that two-thirds of sporadic cases would have an age at onset that is more than 55 years.

The second point raised by Ott and Mérette is that "a difference in age at onset between linked and unlinked families is not strictly addressed" in our analysis. Is it not obvious that such a question did not have to be addressed, since our homogeneity tests did not indicate that families with late onset were unlinked but that they could be explained by the presence of sporadic cases? Of course, we do not exclude the existence of two (or more) age-at-onset distributions among inherited cases. However, at the present time there is no convincing argument for this. In particular, the data presented at the last meeting of The American Society of Human Genetics (Skolnick et al. 1992) do not favor the existence of such a heterogeneity.

FRANÇOISE CLERGET-DARPOUX,
PATRICIA MARGARITTE,
AND CATHERINE BONAÏTI-PELLIÉ

*INSERM Unité 155
Château de Longchamp, Bois de Boulogne
Paris*