HERITABILITY OF MAMMOGRAPHIC DENSITY, A RISK FACTOR FOR BREAST CANCER


ABSTRACT

Background Women with extensive dense breast tissue visible on a mammogram have a risk of breast cancer that is 1.8 to 6.0 times that of women of the same age with little or no density. Menopausal status, weight, and parity account for 20 to 30 percent of the age-adjusted variation in the percentage of dense tissue.

Methods We undertook two studies of twins to determine the proportion of the residual variation in the percentage of density measured by mammography that can be explained by unmeasured additive genetic factors (heritability). A total of 353 pairs of monozygotic twins and 246 pairs of dizygotic twins were recruited from the Australian Twin Registry, and 218 pairs of monozygotic twins and 134 pairs of dizygotic twins were recruited in Canada and the United States. Information on putative determinants of breast density was obtained by questionnaire. Mammograms were digitized, randomly ordered, and read by a blinded investigator.

Results After adjustment for age and measured covariates, the correlation coefficient for the percentage of dense tissue was 0.61 for monozygotic pairs in Australia, 0.67 for monozygotic pairs in North America, 0.25 for dizygotic pairs in Australia, and 0.27 for dizygotic pairs in North America. According to the classic twin model, heritability (the proportion of variants attributable to additive genetic factors) accounted for 60 percent of the variation in density (95 percent confidence interval, 54 to 66) in Australian twins, 67 percent (95 percent confidence interval, 59 to 75) in North American twins, and 63 percent (95 percent confidence interval, 59 to 67) in all twins studied.

Conclusions These results show that the population variation in the percentage of dense tissue on mammography at a given age has high heritability. Because mammographic density is associated with an increased risk of breast cancer, finding the genes responsible for this phenotype could be important for understanding the causes of the disease. (N Engl J Med 2002;347:886-94.)

The New England Journal of Medicine

Copyright © 2002 Massachusetts Medical Society.

From the Division of Epidemiology and Statistics, Ontario Cancer Institute (N.F.B., J.S., A.G., D.T.); Cancer Care Ontario (A.C.); and Imaging Research, Sunnybrook and Women’s College Health Sciences Centre (M.J.Y.) — all in Toronto; the Centre for Genetic Epidemiology, University of Melbourne (G.S.D., J.L.H.); and the Centre for Cancer Epidemiology, Cancer Council Victoria (D.R.E., G.G.G.) — both in Melbourne, Australia; and the Department of Preventive and Social Medicine, University of Otago, Dunedin, New Zealand (M.R.E.M.). Address reprint requests to Dr. Boyd at the Division of Epidemiology and Statistics, Ontario Cancer Institute, 610 University Ave., Toronto, ON M5G 2M9, Canada, or at boyd@uhnres.utoronto.ca.
METHODS

Study Subjects

We assembled two samples of pairs of female monozygotic and dizygotic twins — one sample from Australia and one from Canada and the United States. We collected data on risk factors for breast cancer as putative determinants and obtained mammograms in which mammographic density was measured.

In both Australia and North America, pairs of female monozygotic and dizygotic twins were eligible if they were between 40 and 70 years of age at the time of the interview; if both had undergone or were willing to undergo mammography; if they understood written and spoken English; and if they provided written informed consent. Mammograms in each pair of twins had to have been obtained within 36 months of each other and within 36 months of the time data were collected. Pairs were excluded if one or both twins had had breast cancer, breast augmentation, or breast-reduction surgery before the mammogram was obtained.

Twins were recruited in Melbourne, Sydney, and Perth, Australia, between January 1995 and July 1999 through the Australian Twin Registry. A letter to the twins from the principal investigator at each of these three sites explained the aims of the study, invited participation, and included consent forms for participation and release of mammograms, as well as a postage-paid reply envelope. Twins who agreed to participate were contacted by a research assistant and, if they had not undergone mammography within the previous two years, were given information on how to make an appointment with a state-run mammographic-screening program.

In North America, twins were recruited between May 1997 and February 2001 through print and electronic media; the annual Twins Days Festival held in Twinsburg, Ohio; mammography units of the Ontario Breast Screening Program; and the Twins Foundation (a Rhode Island–based nonprofit organization with a resource center containing information on twins).

Data Collection

Twins provided written informed consent for participation, including permission to release their most recent mammogram. In Australia, films were digitized at a center in Melbourne and sent on compact disk to Toronto. In North America, all films were digitized in Toronto. A questionnaire (described below) was completed by each participating twin; questionnaires were administered by telephone interview in Australia and were self-administered in North America, with telephone interviews used only to clarify incomplete or ambiguous responses.

Questions addressed demographic information, weight, height,
physical activity, smoking history, alcohol consumption, reproductive history, cessation of menstrual periods, use of oral contraceptives and hormone-replacement therapy, breast examination, and family history of cancer.

One cranio-caudal view of one breast was used for each woman (the side was randomly selected for women in North America, and the right side was used for women in Australia). Only original films were used, and all were digitized at a pixel size of 260 µm and a precision of 12 bits. All mammograms were evaluated by a single observer using a technique called “interactive thresholding” that has been described previously,[25] in which the total area of the breast appearing on the mammogram and the area of dense tissue appearing on the mammogram were measured. The percentage of dense tissue was then calculated as the dense area = the total area × 100. Mammograms were read in sets of approximately 120 by an observer who was unaware of the zygosity of each woman or the identity of her twin. Each set contained mammograms from both monozygotic and dizygotic twins, and mammograms from both members of a given pair were included in the same set; the mammograms were randomly ordered within the set. The eight or nine sets per study were read over the course of one to two weeks, each set requiring approximately 45 minutes. The average reliability for the measurement of the percentage of dense tissue within and between sets was approximately 45 minutes. The average reliability for the measurement of the percentage of dense tissue within and between sets was approximately 45 minutes. The average reliability for the measure-

Statistical Analysis

We fitted a fixed-effects model and a random-effects model to the data for the percentage of dense tissue. The fixed effects included age as a polynomial centered on 40 years and linear functions of other measured covariates. The variance and covariance structure was modeled in several ways. Descriptive models involved the residual variance (σ²) and, for pairs of monozygotic and dizygotic twins, either separate covariances (σMZ² and σDZ²) or separate correlations (ρMZ and ρDZ). We also fitted classic twin models, which assume that the residual variance can be partitioned into three components of variance: σ², representing the effects of additive genetic factors; σ², representing the effects of environmental factors that are common to twins within the same pair; and σ², representing person-specific environmental factors, including measurement error.[26] The key assumption of this model is that the degree to which the effects of common environmental factors are shared by twins is the same for monozygotic pairs as it is for dizygotic pairs. According to this model, the total residual variance is σ² = σ² + σ² + σ², the covariance for monozygotic pairs is σ² + σ², and the covariance for dizygotic pairs is σ² + σ², given that monozygotic twins in the same pair share all genetic variants, whereas dizygotic twins share, on average, half their genetic variants.[27] According to this model, the heritability or proportion of residual variance attributed to additive genetic factors is σ²/σ².

The Fisher statistical package was used to fit all models according to maximum likelihood[28] and to test the assumptions of the models.[29] Statistical inference and the choice of parsimonious models were based on standard asymptotic likelihood theory and Akaike’s information criterion.[30] This bivariate normal model was also used to test for differences between defined groups in the means of continuous measured characteristics; for binary variables, an estimate of the asymptotic variance of prevalence that takes into account the measurement error.[31,32] The key assumption of this model is that the degree to which the effects of common environmental factors are shared by twins is the same for monozygotic pairs as it is for dizygotic pairs. According to this model, the total residual variance is σ² = σ² + σ² + σ², the covariance for monozygotic pairs is σ² + σ², and the covariance for dizygotic pairs is σ² + σ², given that monozygotic twins in the same pair share all genetic variants, whereas dizygotic twins share, on average, half their genetic variants.[27] According to this model, the heritability or proportion of residual variance attributed to additive genetic factors is σ²/σ².

The Fisher statistical package was used to fit all models according to maximum likelihood[28] and to test the assumptions of the models.[29] Statistical inference and the choice of parsimonious models were based on standard asymptotic likelihood theory and Akaike’s information criterion.[30] This bivariate normal model was also used to test for differences between defined groups in the means of continuous measured characteristics; for binary variables, an estimate of the asymptotic variance of prevalence that takes into account the concordance of twin pairs was used to assess differences in proportions.[30] All quoted P values are nominal and two-sided.

RESULTS

Characteristics of the Women

Within both Australia and North America, the monozygotic twins were similar to the dizygotic twins in terms of all the characteristics we considered (P>
0.05 for all comparisons) (Table 1). Twins in North America were, on average, three to four years older than twins in Australia at the time of the interview and at the time of mammography (P<0.001 for both comparisons); had a body-mass index (the weight in kilograms divided by the square of the height in meters) that was about half a unit higher than that of the twins in Australia (P=0.08); and were about six months younger at menarche than the twins in Australia (P<0.001). More of the twins in Australia were parous (88 percent vs. 81 percent, P<0.001), but among parous women, the average age at first live birth was similar. A higher proportion of the twins in North America had ceased menstruating (64 percent vs. 54 percent, P<0.001). The average absolute time between mammography and the interview was six months shorter among the twins in Australia (P=0.04) but was independent of zygosity within each population. In all, 11 percent of the twins reported having at least one first-degree relative with breast cancer. In the combined populations, the 112 pairs of twins who reported having a first-degree relative with breast cancer had, after adjustment for age and other covariates, a mean (±SE) percentage of dense tissue of 40.0±1.5 percent, as compared with 37.0±0.6 percent among the 812 pairs who reported that they had no such first-degree relative (P=0.08).

In Australia, more pairs of monozygotic twins than pairs of dizygotic twins underwent mammography in the same clinic as one another (P=0.04), but in North America the proportions were similar among the two types of twins (P=1.00). In both studies, correlations in the percentage of dense tissue for monozygotic pairs and for dizygotic pairs were similar whether the twins attended the same or different clinics (data not shown).

**Percentage of Dense Tissue According to Age**

In each age group, the distribution of the percentage of dense tissue among the twins in Australia was similar to that among the twins in North America (Fig. 2). The variances were large in all age groups, and the means appeared to decrease with increasing age after 50 years of age.

**Adjustment of Mean Percentage of Dense Tissue for Age and Measured Covariates**

To identify variables associated with the mean percentage of dense tissue, we fitted models that included age and other covariates. The best-fitting model for age alone was a quadratic model (data not shown). Body-mass index, age at menarche, cessation or continuation of menstruation, parity, and among parous women, the number of live births and the age at first delivery were associated with the percentage of dense tissue and had similar effects in both studies (data not shown). These variables accounted for about one quarter of the age-adjusted variance in the percentage of dense tissue, and most of the variance attributable to them was due to differences in body-mass index (data not shown). There was no evidence in either study of an independent association between the percentage of dense tissue and the present or previous use of oral contraceptives or hormone-replacement therapy, and a history of smoking or alcohol consumption (data not shown).

**Correlation and Covariance of the Percentage of Dense Tissue within Pairs of Twins**

The residual variance (σ²) was similar in the two studies after adjustment for age and the other covariates (P=1.0) (Table 2). In both studies, the correlations and covariances in the percentage of dense tissue were greater, by a factor of about two, in monozygotic pairs than in dizygotic pairs, whether the analysis was adjusted for age alone (data not shown) or for age and other covariates (P<0.001 for all comparisons). With adjustment for age alone, the correlations in the percentage of dense tissue did not differ significantly between Australia and North America for dizygotic pairs (correlation coefficient, 0.31 and 0.42, respectively; P=0.23) but did differ slightly for monozygotic pairs (correlation coefficient, 0.65 and 0.74, respectively; P=0.03). After adjustment for age alone, the correlation coefficient for the percentage of dense tissue was 0.61 for monozygotic pairs in Australia, 0.67 for monozygotic pairs in North America, 0.25 for dizygotic pairs in Australia, and 0.27 for dizygotic pairs in North America. The correlation coefficients in the combined studies were 0.63 for monozygotic pairs and 0.27 for dizygotic pairs. Scatter plots of the residuals for the percentage of dense tissue in pairs of monozygotic and dizygotic twins, adjusted for age and the other covariates, are shown in Figure 3, which illustrates the stronger correlation in the percentage of dense tissue in monozygotic pairs than in dizygotic pairs in both studies. We found no evidence that the variances, covariances, or correlations varied according to age (data not shown).

**Analysis of the Components of Variance**

In both studies, the best-fitting model of components of variance included only components for additive genetic factors and for person-specific environmental factors, whether it was adjusted for age alone or for age and other covariates. The estimated contribution of the common environmental factors was not significant in any model containing all three components, and in all models containing the additive-genetic and person-specific environmental components resulted in a better fit than the model containing the common environmental and person-specific environmental components. The estimate of
Figure 2. Distribution of the Percentage of Dense Tissue According to Age at Mammography in Twins in Australia and North America.
The central dot in the box plot represents the median, the boxed areas represent the interquartile ranges, and the whiskers extend to 1.5 times the interquartile range.
heritability based on the model containing additive genetic and individual-specific environmental components was 65 percent in the Australian study and 74 percent in the North American study with adjustment for age alone; the estimate was 60 percent in the Australian study and 67 percent in the North American study with adjustment for age and other covariates.

**DISCUSSION**

The results of our two twin studies replicate each other in providing compelling evidence that the wide variation in the percentage of dense tissue on mammography in women 40 to 70 years of age is strongly influenced by genetic factors. The correlation between monozygotic twins was approximately twice as strong as that between dizygotic twins, a finding that is consistent with an additive genetic cause. According to the classic twin model, genetic factors explained the majority of variation in breast density, with estimates of heritability ranging from 60 to 75 percent. When known major determinants of mammographic density were taken into account, the estimates of heritability were reduced by just 10 percent. Measurement was performed by one observer who was blinded to the pairing and zygosity of twins, and the method of measurement was reliable.

The two studies were carried out in geographically separate populations, both predominately European in origin and living in “Western” environments, so our findings do not rule out the possibility of a greater influence of environmental factors at levels of exposure that lie outside the range usually seen in Western societies. Furthermore, although the estimates of the effect of environmental factors common to twins were negative and not significant, the study did not have adequate statistical power to rule out small effects.

Twins are an important natural experimental model. Monozygotic twins within the same pair are genetic copies of each other, so that any differences between them must be the result of environmental factors and measurement error. Dizygotic twins share, on average, half their genes, so that differences within same-sex pairs may be due to environmental and genetic factors, as well as to measurement error. Both monozygotic and dizygotic twins usually share a common environment, at least during early life. The classic twin model makes the critical assumption that the correlation between twins in the strength of the effects of common environmental factors on the trait of interest is the same for monozygotic pairs as it is for dizygotic pairs of the same sex. Under this assumption, the demonstration for a given trait of a difference in covariances (and if variance is independent of zygosity, in correlations) between monozygotic pairs and dizygotic pairs is consistent with the existence of one or more genetic factors that determine variation in that trait.

Our finding that the degree of mammographic density for age is highly heritable is in keeping with the limited existing evidence. Wolfe et al. found more agreement in parenchymal patterns of the breast in mother–daughter pairs and in pairs of sisters than in age-matched control pairs, and Pankow et al. found an age-adjusted correlation in mammographic density between sisters of 0.22. The only other published study in twins showed a greater correlation of parenchymal patterns of the breast in 7 pairs of monozygotic twins than in 23 pairs of dizygotic twins. In premenopausal women, variations in mammographic density are associated with blood and tissue levels of insulin-like growth factor I, and in postmenopausal women with blood levels of prolactin. Inherited variations in the production or metabolism of these and other mitogens that act on the breast might underlie the heritability of mammographic density.

A high percentage of dense tissue is associated with

### Table 2. Estimates of Residual Variance, Correlations Between Monozygotic Pairs of Twins and Dizygotic Pairs of Twins, and Heritability of the Percentage of Dense Tissue, Adjusted for Age and Other Covariates.*

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>AUSTRALIA</th>
<th>NORTH AMERICA</th>
<th>COMBINED STUDIES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residual variance ($\sigma^2$)</td>
<td>306.40±13.97</td>
<td>306.00±18.61</td>
<td>310.00±11.33</td>
</tr>
<tr>
<td>Correlation coefficient for monozygotic pairs</td>
<td>0.61±0.03</td>
<td>0.67±0.03</td>
<td>0.63±0.02</td>
</tr>
<tr>
<td>Correlation coefficient for dizygotic pairs</td>
<td>0.25±0.06</td>
<td>0.27±0.08</td>
<td>0.27±0.05</td>
</tr>
<tr>
<td>Heritability</td>
<td>0.60±0.03</td>
<td>0.67±0.04</td>
<td>0.63±0.02</td>
</tr>
</tbody>
</table>

*Plus–minus values are estimates ±SE. The mean percentage of dense tissue was adjusted for body-mass index, age at menarche, cessation or continuation of menstruation, parity, and in parous women, number of live births and age at first birth. The symbol $\sigma^2$ denotes the residual variance, in units of the square of the percentage of dense tissue. A more detailed analysis appears in Supplementary Appendix 1 (available with the full text of this article at http://www.nejm.org).
large relative and attributable risks of breast cancer, and the evidence that it has high heritability has implications for our understanding of breast cancer in general, as well as of familial aggregation of the disease. Familial aggregation of breast cancer on a population basis is important because the apparently moderate doubling of risk (on average) that is associated with having an affected first-degree relative can only occur if there are strong underlying familial risk factors. Current evidence suggests that the principal known susceptibility genes (BRCA1 and BRCA2) explain 20 percent or less of familial aggregation of breast cancer on a population basis (although they explain a greater proportion in rare families with multiple cases). Perhaps 10 percent of this familial aggregation may be explained by the currently known lifestyle-related risk factors, so it is plausible that there are other breast-cancer susceptibility genes that have moderate or strong effects on the level of risk.

The average relative risk of breast cancer for women in the highest category of the percentage of dense tissue as compared with those in the lowest category is about 4.0, and the correlation between dizygotic twin sisters in our studies was 0.27. On the basis of these data, we estimate that familial associations in the percentage of dense tissue alone could increase

---

**Figure 3.** Correlations in the Percentage of Dense Tissue for Pairs of Twins in Australia and North America. Plus–minus values are estimates ±SE.
risk in the first-degree relatives of affected women by a factor of 1.05 to 1.08 and explain an additional 5 to 8 percent of familial aggregation on a population basis. The number of genes that influence mammographic density remains to be determined, as does any role they may have in causing breast cancer. However, extensive mammographic density is common, is associated with a markedly increased risk of breast cancer, and may account for a large proportion of cases of the disease. Thus, the genes responsible for familial correlation in mammographic density could influence susceptibility to breast cancer in a large fraction of the population and contribute to some of the familial aggregation of the disease. Our two twin studies suggest that a potentially fruitful approach to the identification of new susceptibility genes for breast cancer may be to study pairs of sisters to analyze variants in candidate genes in the potential causal pathways for high breast density.

Supported by grants from the National Breast Cancer Foundation (Australia), the National Health and Medical Research Council (Australia), the Merck, Sharp & Dohme Research Foundation (Australia), and the Canadian Breast Cancer Research Initiative.

We are indebted to Jennifer Cawson, Penny McIver, Helen O'Connor, Maggie Angelakos, the twins who participated in the study, and the Australian Twin Registry.

APPENDIX

Questions used to determine zygosity:
1. Were you and your twin “as alike as two peas in a pod”? As alike as two peas in a pod
   - Usual sibling similarity
   - Quite different
2. Were you and your twin mixed up as children?
   - Yes, very often
   - Now and then
   - Never
3. In that case, by whom were you mixed up?
   - Parents
   - Teachers
   - Others
   - Nobody

REFERENCES
