

The New England Journal of Medicine

© Copyright, 2000, by the Massachusetts Medical Society

VOLUME 343

JULY 13, 2000

NUMBER 2



ENVIRONMENTAL AND HERITABLE FACTORS IN THE CAUSATION OF CANCER

Analyses of Cohorts of Twins from Sweden, Denmark, and Finland

PAUL LICHTENSTEIN, PH.D., NIELS V. HOLM, M.D., PH.D., PIA K. VERKASALO, M.D., PH.D., ANASTASIA ILIADOU, M.Sc.,
JAAKKO KAPRIO, M.D., PH.D., MARKKU KOSKENVUO, M.D., PH.D., EERO PUKKALA, PH.D., AXEL SKYTTE, M.Sc.,
AND KARI HEMMINKI, M.D., PH.D.

ABSTRACT

Background The contribution of hereditary factors to the causation of sporadic cancer is unclear. Studies of twins make it possible to estimate the overall contribution of inherited genes to the development of malignant diseases.

Methods We combined data on 44,788 pairs of twins listed in the Swedish, Danish, and Finnish twin registries in order to assess the risks of cancer at 28 anatomical sites for the twins of persons with cancer. Statistical modeling was used to estimate the relative importance of heritable and environmental factors in causing cancer at 11 of those sites.

Results At least one cancer occurred in 10,803 persons among 9512 pairs of twins. An increased risk was found among the twins of affected persons for stomach, colorectal, lung, breast, and prostate cancer. Statistically significant effects of heritable factors were observed for prostate cancer (42 percent of the risk may be explained by heritable factors; 95 percent confidence interval, 29 to 50 percent), colorectal cancer (35 percent; 95 percent confidence interval, 10 to 48 percent), and breast cancer (27 percent; 95 percent confidence interval, 4 to 41 percent).

Conclusions Inherited genetic factors make a minor contribution to susceptibility to most types of neoplasms. This finding indicates that the environment has the principal role in causing sporadic cancer. The relatively large effect of heritability in cancer at a few sites (such as prostate and colorectal cancer) suggests major gaps in our knowledge of the genetics of cancer. (N Engl J Med 2000;343:78-85.)

©2000, Massachusetts Medical Society.

EXCEPT for certain types of familial cancer, such as adenomatous polyposis coli, the contribution of hereditary factors to the development of cancer is thought to be relatively minor.¹⁻³ This premise, however, applies mainly to dominant genes, which have been assessed in family studies that cover two or more generations. By contrast, the contributions of recessive traits and combinations of genes to the causation of sporadic cancer are difficult to determine from family studies.⁴ Consequently, the risks associated with single-gene mutations with low penetrance, recessive genes, and oncogenic mechanisms that involve multiple genes are poorly understood.

Family studies of breast, prostate, ovarian, and uterine cancer can estimate risks for siblings and parent-offspring pairs⁵⁻¹² but cannot distinguish between genetic and nongenetic (environmental or infectious) causes of familial aggregations of cancer. By contrast, comparisons of the concordance of cancer between monozygotic and dizygotic pairs of twins provide information on whether the familial pattern is due to hereditary or environmental influences.¹³

If studies of groups of twins show that concordance for cancer is higher among monozygotic twins (who share all genes) than among dizygotic twins (who, on average, share 50 percent of their segregating genes), genetic effects are likely to be important.

From the Department of Medical Epidemiology, Karolinska Institute, Stockholm, Sweden (P.L., A.I.); the Institute of Public Health (Epidemiology) and the Danish Twin Registry, University of Southern Denmark, Odense (N.V.H., A.S.); the Department of Public Health, University of Helsinki, Helsinki, Finland (P.K.V., J.K., M.K.); the Department of Public Health and General Practice, University of Oulu, Oulu, Finland (J.K.); the Department of Public Health, University of Turku, Turku, Finland (M.K.); the Finnish Cancer Registry, Helsinki, Finland (E.P.); and the Department of Biosciences at Novum, Karolinska Institute, Stockholm, Sweden (K.H.). Address reprint requests to Dr. Lichtenstein at the Department of Medical Epidemiology, Karolinska Institute, Box 281, SE-171 77 Stockholm, Sweden, or at paul.lichtenstein@mep.ki.se.

If, however, the concordance is similar for both types of twins, then shared environmental effects are probably important. Furthermore, the use of statistical models to analyze data from large samples of twins makes it possible to estimate the magnitude of the genetic and environmental effects on susceptibility to sporadic cancer.

For these reasons, studies of twins can not only point to hereditary effects, but also estimate heritability, a term denoting the magnitude of the genetic effect. However, the rarity of twins limits this approach, even for the common forms of cancer.¹⁴⁻¹⁹

In the present study, we used data from the Swedish, Danish, and Finnish twin registries to estimate the effects of genetic and environmental factors on the most common cancers. We also assessed how age at the time of diagnosis modified these estimates.

METHODS

Swedish Twins

The Swedish Twin Registry consists of two birth cohorts.²⁰ The first is made up of 10,503 pairs of twins of the same sex who were alive in 1961 and who were born during the period from 1886 through 1925. Information from questionnaires completed by both twins was available for 81 percent of the eligible pairs. A second cohort consists of 12,883 pairs of twins of the same sex born from 1926 through 1958. In this cohort, both twins were living in Sweden in 1972 and had responded to a questionnaire in that year. The rate of response to this questionnaire was 83 percent.

We determined vital status and any diagnoses of cancer from the records of the Swedish Mortality Registry and the Swedish Cancer Registry, using the unique national registration number assigned to each Swedish citizen. According to these records, cancer was diagnosed in 4490 persons in the first cohort from 1961 through 1995 and in 1157 persons in the second cohort from 1973 through 1995.

Danish Twins

The Danish Twin Registry holds data on 8461 pairs of twins of the same sex with known zygosity who were born between 1870 and 1930. This registry, established in 1954, includes all twins born in Denmark from 1870 through 1910,²¹ and it was later expanded to include twins of the same sex born from 1911 through 1930.²¹⁻²³ Included in the registry are all pairs of twins who both survived to the age of six years. A questionnaire was mailed to the twins, or to their closest relatives if one or both twins had emigrated or were dead at the time of identification.

We checked vital status annually through 1979 by obtaining

copies of death certificates from the Central Register of Deaths. After 1979 vital status was regularly updated by linkage to the Civil Registration System, which includes all persons living in Denmark since April 1, 1968.

The Danish Cancer Registry contains information on all malignant diseases diagnosed in Denmark since 1943.²⁴⁻²⁶ All pairs of twins of the same sex who were born from 1870 through 1930 and were both alive on January 1, 1943, have been linked to the Cancer Registry for the period from 1943 through 1993. A total of 3572 persons in this cohort received a diagnosis of cancer (excluding nonmelanoma skin cancer).

Finnish Twins

The Finnish twin cohort includes 12,941 pairs of twins who were born from 1880 through 1958 and who were both living in Finland on December 31, 1975.²⁷ The cohort was compiled from the Central Population Register in 1974. The following year, a questionnaire was mailed to all twins who were 18 years of age or older and for whom an adequate address was available. The overall response rate was 89 percent.

Malignant neoplasms that were diagnosed among the Finnish twins from 1976 through 1996 were identified by linkage of records to national cancer registry data with the use of the personal identification number assigned to every resident of Finland. The Finnish Cancer Registry has information on all cancers diagnosed in Finland since 1953.²⁸ In addition, the study cohort was linked to the Central Population Register to obtain data on death and emigration. Cancer was diagnosed in 1584 persons in the cohort.

Determination of Zygosity

For all three studies, zygosity was determined by a questionnaire that has been shown in validation studies to classify more than 95 percent of pairs of twins correctly.^{22,29,30}

Statistical Analysis

The relative risk of cancer for persons whose twins had a particular type of cancer, as compared with those whose twins did not, was calculated according to sex and zygosity for cancer at each anatomical site. The risk was estimated as an odds ratio. Ninety-five percent confidence intervals were estimated according to the Mantel-Haenszel method.³¹

The absolute risk of cancer for the twin of a person with cancer within the period of the study was calculated as the proportion of all persons with cancer whose twins had cancer at the same site (i.e., the concordance).³² For clinical guidance, we also calculated the risk for twins up to the age of 75 years for the sites for which significant effects of heritable factors were found.

Quantitative genetic analyses were used to estimate the relative importance of hereditary and environmental factors in determining variations in susceptibility to cancer,^{13,33,34} on the usual assumptions of a classic twin study (that there was random mating, no

TABLE 1. EFFECTS ESTIMATED IN THE QUANTITATIVE GENETIC ANALYSES.

| TYPE OF EFFECT | DEFINITION | INDICATION OF EFFECT* | EXAMPLES |
|-------------------------|---|---|---|
| Hereditary | The proportion of phenotypic variance accounted for by inherited genetic differences among persons (heritability) | Similarity greater in MZ twins than in DZ twins | Additive and dominant genetic effects |
| Shared environmental | The proportion of phenotypic variance accounted for by environmental factors shared by both twins, thus contributing to similarity between them | Similarity among both MZ and DZ twins greater than would be expected from genetic effects alone | Environmental factors contributing to similarity in pairs of twins — e.g., passive smoking in childhood family (lung cancer) or similar dietary habits (stomach cancer) |
| Nonshared environmental | The proportion of phenotypic variance accounted for by environmental factors causing differences between twins | Lack of similarity in both MZ and DZ twins | Environments that are not shared by twins — e.g., sporadic mutations, occupational exposure, or viral infections |

*MZ denotes monozygotic, and DZ dizygotic.

interaction between genes and environment, and equivalent environments for monozygotic and dizygotic twins).³⁵ Phenotypic variance was divided into a component due to inherited genetic factors (heritability), a component due to environmental factors common to both members of the pair of twins (the shared environmental component), and a component due to environmental factors unique to each twin (the nonshared environmental component) (Table 1). Structural-equation modeling,³⁶ which incorporates data on all types of twins (male and female monozygotic and dizygotic twins from the three countries) simultaneously, provided estimates of the unobserved variables — that is, additive genetic, shared environmental, and nonshared environmental factors.

The correlations between the genetic and environmental factors for the twins were set to their theoretical values (1.0 and 0.5 for additive genetic effects for monozygotic and dizygotic twins, respectively, and 1.0 for shared environmental effects for both types of twins). These values, 1.0 and 0.5, reflect the facts that monozygotic twins share their entire genomes (theoretical value, 1.0) and dizygotic twins share 50 percent of their segregating genes (theoretical value, 0.5). This method, which uses two-by-two contingency tables of disease status in pairs of twins, also tests for the sex specificity of the genetic and environmental effects. Because of considerations related to statistical power, analyses were performed only for cancers for which there were at least four pairs of twins in which both twins had the cancer. We assumed an underlying normal distribution of susceptibility to the disease.

We defined susceptibility as the sum of the effects of many genetic and environmental factors. When a person receives a diagnosis of cancer, a value (the threshold) in the distribution of susceptibility is assumed to have been exceeded. The threshold value was estimated in the model from the prevalence of the disease.^{13,37} The relative importance of hereditary and environmental effects for individual differences in this underlying susceptibility was then estimated. When sex-specific models did not fit significantly better than a model in which estimates were defined as being the same in men and women, the latter model is presented. Because the birth dates and follow-up periods of the twin cohorts differed among countries, the threshold values were also allowed to differ among countries. However, the estimates for genetic and environmental components were set to be equal in all countries, because no evidence of heterogeneity according to country was found for cancer at any site.

For colorectal, breast, and prostate cancers, there were enough affected persons (i.e., more than 50 pairs of twins who were concordant for the cancer) to enable structural model-fitting analysis to be performed in two age groups. Because of power considerations, the groups were defined so that the younger group included about 35 percent of the affected twins. For colorectal cancer, the younger group was followed up to 63 years of age; for breast cancer, it was followed up to 56 years of age; and for prostate cancer, it was followed up to 70 years of age. Persons in the younger group were followed until they reached the maximal age for that group, until they died, or until the study ended. For the older group, follow-up started when they reached the maximal age for the younger group and ended at death or at the end of the study period. At the end of follow-up in the younger and older groups, each pair of twins was recorded as concordant for cancer (if both twins had had cancer at the same anatomical site) or discordant for cancer (if only one twin had had cancer at a particular site). For colorectal, breast, and prostate cancer, we examined the differences in the ages at which cancer was diagnosed in pairs of twins who were concordant for cancer.

RESULTS

Among the 44,788 pairs of twins included in this analysis, we identified 10,803 persons (among 9512 pairs) in whom at least one cancer had been diagnosed.

Overall, the twin of a person with cancer had an increased risk of having the same cancer. This was especially evident for cancer of the stomach, colorectum, lung, breast, and prostate (Table 2). The twin

TABLE 2. TYPES OF CANCERS INCLUDED IN THE STUDY AND CONCORDANCE ACCORDING TO SEX AND ZYGOSITY IN 44,788 PAIRS OF TWINS FROM SWEDEN, DENMARK, AND FINLAND.*

| SITE OF CANCER AND SUBJECTS | No. OF CONCORDANT AFFECTED PAIRS | No. OF DISCORDANT PAIRS | RELATIVE RISK (95% CI)† | CONCORDANCE‡ | SITE OF CANCER AND SUBJECTS | No. OF CONCORDANT AFFECTED PAIRS | No. OF DISCORDANT PAIRS | RELATIVE RISK (95% CI)† | CONCORDANCE‡ |
|-----------------------------|----------------------------------|-------------------------|-------------------------|--------------|--------------------------------|----------------------------------|-------------------------|-------------------------|--------------|
| Lip | | | | | Larynx | | | | |
| Men | | | | | Men | 2 | 22 | 119.2 (23.6–601.1) | 0.15 |
| MZ | 0 | 27 | — | 0 | MZ | 1 | 36 | 42.4 (5.4–334.6) | 0.05 |
| DZ | 0 | 57 | — | 0 | DZ | 0 | 3 | — | 0 |
| Women | | | | | Women | 0 | 11 | — | 0 |
| MZ | 0 | 4 | — | 0 | MZ | 0 | 3 | — | 0 |
| DZ | 0 | 8 | — | 0 | DZ | 0 | 11 | — | 0 |
| Oral cavity | | | | | Lung | | | | |
| Men | | | | | Men | 15 | 233 | 7.7 (4.4–13.6) | 0.11 |
| MZ | 0 | 10 | — | 0 | MZ | 24 | 436 | 6.7 (4.3–10.5) | 0.10 |
| DZ | 0 | 20 | — | 0 | DZ | 3 | 63 | 25.3 (7.4–87.0) | 0.09 |
| Women | | | | | Women | 1 | 185 | 1.8 (0.2–12.8) | 0.01 |
| MZ | 0 | 6 | — | 0 | MZ | 0 | 2 | — | 0 |
| DZ | 0 | 24 | — | 0 | DZ | 0 | 5 | — | 0 |
| Pharynx§ | | | | | Breast | | | | |
| Men | | | | | Men | 0 | 2 | — | 0 |
| MZ | 0 | 11 | — | 0 | MZ | 0 | 5 | — | 0 |
| DZ | 0 | 19 | — | 0 | DZ | | | | |
| | | | | | Bone | | | | |
| | | | | | Men | | | | |
| | | | | | MZ | | | | |
| | | | | | DZ | | | | |
| | | | | | Brain and other nervous system | | | | |
| | | | | | Men | 1 | 65 | 6.8 (0.9–51.1) | 0.03 |
| | | | | | MZ | 0 | 112 | — | 0 |
| | | | | | DZ | 0 | 66 | — | 0 |
| | | | | | Women | 1 | 142 | 3.0 (0.4–22.0) | 0.01 |
| | | | | | MZ | | | | |
| | | | | | DZ | | | | |
| | | | | | Thyroid | | | | |
| | | | | | Men | 0 | 10 | — | 0 |
| | | | | | MZ | 0 | 19 | — | 0 |
| | | | | | DZ | 0 | 29 | — | 0 |
| | | | | | Women | 0 | 44 | — | 0 |
| | | | | | MZ | | | | |
| | | | | | DZ | | | | |
| | | | | | Men | 0 | 5 | — | 0 |
| | | | | | MZ | | | | |
| | | | | | DZ | | | | |

ENVIRONMENTAL AND HERITABLE FACTORS IN THE CAUSATION OF CANCER

| | | | | | | | | | | | | | |
|-------------------------------|----|-----|------------------|------|----|------|------------------|------|----------|---|-----|------------------|------|
| Women | 0 | 7 | — | 0 | 42 | 505 | 5.2 (3.7-7.4) | 0.14 | DZ | 0 | 8 | — | 0 |
| MZ | 0 | 5 | — | 0 | 52 | 1023 | 2.8 (2.1-3.8) | 0.09 | Women | 0 | 2 | — | 0 |
| DZ | 0 | 16 | — | 0 | 1 | 107 | 2.9 (0.4-21.4) | 0.02 | MZ | 0 | 12 | — | 0 |
| Esophagus | 1 | 50 | 22.0 (2.9-168.3) | 0.04 | 3 | 201 | 4.5 (1.4-14.4) | 0.03 | DZ | 0 | 27 | — | 0 |
| Men | 0 | 17 | — | 0 | 1 | 122 | 2.2 (0.3-16.4) | 0.02 | Women | 0 | 14 | — | 0 |
| MZ | 0 | 31 | — | 0 | 5 | 245 | 5.0 (2.0-12.5) | 0.04 | MZ | 0 | 23 | — | 0 |
| DZ | 6 | 131 | 9.9 (4.1-23.6) | 0.06 | 3 | 125 | 6.4 (2.0-20.9) | 0.05 | DZ | 0 | 36 | — | 0 |
| Women | 8 | 256 | 6.6 (3.2-13.8) | 0.10 | 1 | 230 | 1.1 (0.2-8.2) | 0.01 | Men | 0 | 99 | — | 0 |
| MZ | 5 | 92 | 19.7 (7.5-51.6) | 0.04 | 40 | 299 | 12.3 (8.4-18.1) | 0.21 | MZ | 0 | 45 | — | 0 |
| DZ | 4 | 198 | 6.2 (2.2-17.1) | 0.09 | 20 | 584 | 3.1 (1.9-4.9) | 0.06 | DZ | 0 | 84 | — | 0 |
| Colorectum [¶] | 10 | 202 | 6.9 (3.5-13.6) | 0.08 | 1 | 22 | 59.6 (7.1-501.7) | 0.08 | Women | 0 | 14 | — | 0 |
| Men | 17 | 393 | 5.9 (3.5-9.8) | 0.16 | 0 | 38 | — | 0 | MZ | 0 | 19 | — | 0 |
| MZ | 20 | 214 | 14.3 (8.6-24.0) | 0.06 | 0 | 61 | — | 0 | DZ | 0 | 18 | — | 0 |
| DZ | 15 | 453 | 4.4 (2.5-7.5) | 0.05 | 0 | 108 | — | 0 | Women | 0 | 12 | — | 0 |
| Liver | 0 | 30 | — | 0 | 0 | 48 | — | 0 | MZ | 0 | 36 | — | 0 |
| Men | 1 | 44 | 28.4 (3.7-219.6) | 0.04 | 0 | 99 | — | 0 | DZ | 0 | 57 | — | 0 |
| MZ | 0 | 23 | — | 0 | 5 | 146 | 6.6 (2.6-16.9) | 0.06 | Men | 1 | 36 | 22.2 (2.8-175.3) | 0.05 |
| DZ | 1 | 35 | 50.0 (6.3-396.3) | 0.05 | 2 | 253 | 1.7 (0.4-6.9) | 0.02 | DZ | 0 | 28 | — | 0 |
| Gallbladder and biliary tract | 0 | 16 | — | 0 | 0 | 43 | — | 0 | Women | 0 | 38 | — | 0 |
| Men | 0 | 25 | — | 0 | 0 | 76 | — | 0 | MZ | 0 | 51 | 22.1 (5.0-98.1) | 0.07 |
| MZ | 0 | 51 | 12.9 (1.7-98.8) | 0.04 | 0 | 49 | — | 0 | DZ | 1 | 105 | 5.0 (0.7-36.5) | 0.02 |
| DZ | 1 | 81 | 9.3 (1.2-69.3) | 0.02 | 0 | 78 | 18.0 (4.2-77.2) | 0.05 | Leukemia | 0 | 52 | — | 0 |
| Women | 1 | 81 | — | 0.06 | 2 | 49 | — | 0 | Men | 0 | 93 | 7.1 (1.0-52.2) | 0.02 |
| MZ | 2 | 64 | 14.0 (3.2-60.9) | 0.03 | 0 | 58 | 10.0 (1.3-75.6) | 0.03 | Women | 0 | 262 | 3.8 (3.2-4.5) | 0.29 |
| DZ | 1 | 127 | 12.7 (3.0-54.1) | 0.03 | 0 | 105 | — | 0 | MZ | 0 | 356 | 2.6 (2.3-2.9) | 0.22 |
| Pancreas | 1 | 59 | 9.6 (1.3-73.0) | 0 | 1 | 265 | — | 0 | DZ | 0 | 265 | 3.2 (2.7-3.8) | 0.26 |
| Men | 0 | 127 | — | 0 | 0 | 408 | — | 0 | Women | 0 | 408 | 2.1 (1.9-2.4) | 0.21 |
| MZ | | | | | | | | | MZ | | | | |
| DZ | | | | | | | | | DZ | | | | |

*MZ denotes monozygotic, DZ dizygotic, and CI confidence interval.

†The relative risk was calculated as ad ÷ bc, where a is the number of concordant pairs, b and c are each half the number of discordant pairs, and d is the number of concordant pairs without cancer, or the total number of pairs - a - b - c; the total number of pairs was 7231 for monozygotic men, 15,769 for dizygotic men, 8437 for monozygotic women, and 15,351 for dizygotic women.

‡Concordance was calculated as the proportion of all persons with cancer whose twins had cancer at the same site. §Cancer of the nasopharynx is excluded. ¶Cancer of the anus is excluded.

||Histologically verified benign tumors of the central nervous system and meninges are included.

**Nonmelanoma skin cancer is excluded because of problems of completeness of registration in different periods.

of a male monozygotic twin who had stomach cancer had a risk of stomach cancer that was 9.9 times that of the monozygotic twin of a person without stomach cancer. The concordance for stomach cancer in male monozygotic twins was 0.08, which means that there is an 8 percent probability that the identical twin of a man with stomach cancer will have the same cancer. The concordance was usually less than 0.10, and no concordant pairs were observed for cancers at nine sites (non-Hodgkin's lymphoma, Hodgkin's disease, and cancer of the lip, oral cavity, pharynx, kidney, thyroid, bone, and soft tissue). For cancers at most of the remaining sites, the concordance between monozygotic twins, whether male or female, was greater than the concordance between dizygotic twins.

Table 3 presents the results of model fitting, which we used to obtain estimates of the contributions of heritability and environmental effects. For stomach cancer, for example, heritability was estimated to account for 28 percent of the variation in susceptibility to that neoplasm, shared environmental effects for 10 percent, and nonshared environmental effects for the remaining 62 percent. Stated another way, these estimates indicate that of the various factors that together constitute the total risk of developing stomach cancer, inherited genes contribute 28 percent to the risk, shared environmental effects contribute 10 percent, and nonshared environmental factors make up the remaining 62 percent of the risk. For stomach cancer, therefore, our model predicts the involvement of major environmental factors plus minor genetic components, which may or may not interact with these environmental factors.

The statistical model we used provided an excellent

fit to the observed data ($\chi^2=8.9$, with 38 df; $P=1.0$). Estimated effects of heritability — the proportion of susceptibility to cancer that was accounted for by genetic defects — that were statistically significant (i.e., for which the 95 percent confidence interval did not include zero) were obtained for cancers of the colorectum (35 percent), breast (27 percent), and prostate (42 percent). The estimates for the shared environmental effects ranged from 0 to 20 percent, but none were statistically significant. There were no significant differences between the sexes in the heritability of cancer at any of the sites that we studied.

Table 4 presents the risks of having the same cancer before the age of 75 years among twins of persons with cancer at sites involving statistically significant genetic factors. For colorectal, breast, and prostate cancer, the estimated hereditary components were slightly greater in the younger than in the older groups (data not shown).

The time interval between the diagnoses of prostate cancer was significantly shorter for concordant pairs of monozygotic twins than that for concordant pairs of dizygotic twins (5.7 vs. 8.8 years) (Table 5). There were no significant differences between dizygotic and monozygotic twins in the time between diagnoses of colorectal and breast cancer.

DISCUSSION

Assessments of the contributions of inherited and environmental factors to the causation of cancer in studies of twins have had a relatively small effect on research and clinical practice, because twins are rare, and only a few twin registries go back far enough in time to provide enough cases of cancer for reliable

TABLE 3. EFFECTS OF HERITABLE AND ENVIRONMENTAL FACTORS IN CANCERS AT VARIOUS SITES, ACCORDING TO DATA FROM THE SWEDISH, DANISH, AND FINNISH TWIN REGISTRIES.

| SITE OR TYPE | PROPORTION OF VARIANCE (95% CI)* | | | FIT OF MODEL | |
|----------------|----------------------------------|------------------------------|---------------------------------|---------------|---------|
| | HERITABLE FACTORS | SHARED ENVIRONMENTAL FACTORS | NONSHARED ENVIRONMENTAL FACTORS | χ^2 (df) | P VALUE |
| Stomach | 0.28 (0–0.51) | 0.10 (0–0.34) | 0.62 (0.49–0.76) | 8.9 (38) | 1.0 |
| Colorectum | 0.35 (0.10–0.48) | 0.05 (0–0.23) | 0.60 (0.52–0.70) | 25.8 (38) | 0.93 |
| Pancreas† | 0.36 (0–0.53) | 0 (0–0.35) | 0.64 (0.47–0.86) | 0.5 (3) | 0.92 |
| Lung | 0.26 (0–0.49) | 0.12 (0–0.34) | 0.62 (0.51–0.73) | 28.1 (38) | 0.88 |
| Breast‡ | 0.27 (0.04–0.41) | 0.06 (0–0.22) | 0.67 (0.59–0.76) | 10.1 (18) | 0.93 |
| Cervix uteri†† | 0 (0–0.42) | 0.20 (0–0.35) | 0.80 (0.57–0.97) | 0.3 (3) | 0.96 |
| Corpus uteri‡ | 0 (0–0.35) | 0.17 (0–0.31) | 0.82 (0.64–0.98) | 6.6 (18) | 0.99 |
| Ovary‡ | 0.22 (0–0.41) | 0 (0–0.24) | 0.78 (0.59–0.99) | 6.0 (18) | 1.0 |
| Prostate§ | 0.42 (0.29–0.50) | 0 (0–0.09) | 0.58 (0.50–0.67) | 26.5 (18) | 0.09 |
| Bladder† | 0.31 (0–0.45) | 0 (0–0.28) | 0.69 (0.53–0.86) | 1.7 (3) | 0.64 |
| Leukemia† | 0.21 (0–0.54) | 0.12 (0–0.41) | 0.66 (0.45–0.88) | 0.0 (3) | 0.99 |

*CI denotes confidence interval.

†Data for all countries and both sexes are pooled because of small numbers.

‡Data are for women only.

§Data are for men only.

TABLE 4. ABSOLUTE RISKS OF COLORECTAL, BREAST, AND PROSTATE CANCER (CONCORDANCE RATES) IN TWINS OF AN AFFECTED PERSON UP TO THE AGE OF 75 YEARS.

| SITE OF CANCER | MONOZYGOTIC TWINS | DIZYGOTIC TWINS |
|-------------------|-------------------|-----------------|
| Colorectum | 0.11 | 0.05 |
| Breast (in women) | 0.13 | 0.09 |
| Prostate | 0.18 | 0.03 |

TABLE 5. DIFFERENCE IN AGE AT THE DIAGNOSIS OF CANCER IN CONCORDANT PAIRS OF MONOZYGOTIC AND DIZYGOTIC TWINS.*

| SITE OF CANCER AND SUBJECTS | NO. OF PAIRS | DIFFERENCE IN AGE AT DIAGNOSIS | P VALUE |
|-----------------------------|--------------|--------------------------------|---------|
| | | yr | |
| Colorectum | | | |
| MZ twins | 29 | 9.48±9.49 | 0.86 |
| DZ twins | 32 | 9.94±10.00 | |
| Breast (in women) | | | |
| MZ twins | 41 | 7.90±5.95 | 0.12 |
| DZ twins | 52 | 10.21±8.16 | |
| Prostate | | | |
| MZ twins | 39 | 5.69±3.39 | 0.04 |
| DZ twins | 20 | 8.75±5.66 | |

*Plus-minus values are means ±SD. Data from pairs with a mean difference larger than 3 SD were excluded from these analyses. MZ denotes monozygotic, and DZ dizygotic.

conclusions to be drawn. The fact that we combined data from three Nordic countries with population-based twin registries and well-established cancer registries adds strength to the present study. Moreover, with statistical methods, it was possible to go beyond a simple comparison of concordance rates. With the knowledge that monozygotic twins are genetically identical and dizygotic twins share, on average, 50 percent of their segregating genes, we could estimate the magnitude of the contributions of genetic factors and environmental factors (both shared and non-shared) to the development of cancer.

There is general agreement that environmental factors and somatic events are the predominant contributors to the causation of sporadic cancer,³⁸ and our results support this notion. Even though we found familial (hereditary or nonhereditary) effects for cancer at many sites, the rates of concordance in twins were generally below 0.10. This result indicates that, for nearly all sites, the twin of a person with cancer has only a moderate absolute risk of having cancer at the same site. Using a statistical model, we also es-

timated the contribution of nonshared environmental factors, which include any unique environmental cause of cancer that is not inherited and not shared between twins. For different cancers, this contribution ranged from 58 to 82 percent.

In our model, shared environment, the sum of the common family experiences and habits of the twins, accounted for 0 to 20 percent of causation, but none of these values were statistically significant, partly because studies of twins have limited power to detect shared environmental effects in dichotomous phenotypes (e.g., the presence or absence of cancer).³⁹ Risk factors in the environment shared by a family could include human papillomavirus infection for cervical cancer, smoking (passive or active) for lung cancer, and diet and *Helicobacter pylori* infection for stomach cancer.

The total contribution of hereditary factors to the causation of sporadic cancer is unclear; previous assessments have estimated only the proportion of cancers caused by genetic syndromes. It has been argued that “unmistakable hereditary cancer syndromes” account for about 1 percent of cancers and “upward of 10 to 15 percent of all cancers have a major inherited component, albeit one that may be enigmatic”³³; that “highly penetrant single-gene mutation” accounts for 5 percent⁴¹; and that “primary genetic factors” account for 5 to 10 percent of all cancers.⁴⁰ The results of our study summarize the total effects of heritable factors, but it should be noted that our estimates are population-specific. Thus, if environmental effects are very different in Scandinavia and in other regions, the proportion of susceptibility to cancer that is due to hereditary effects will also differ. And for populations consisting entirely of smokers or of nonsmokers, the contribution of smoking to the variation in risk would be much lower than in a mixed population. Previous studies of cancer in twins have found higher rates of concordance among monozygotic twins than among dizygotic twins for cancer at some sites,^{14-19,41-44} but generally with very wide confidence intervals.

We found statistically significant effects of heritable factors, ranging from 27 percent to 42 percent, for colorectal, breast, and prostate cancer. These results are in agreement with those of most previous studies.^{14-16,43} Our model also revealed suggestive evidence of limited heritability of leukemia and of cancer of the stomach, lung, pancreas, ovary, and bladder, but the estimates did not reach statistical significance. Population-based studies in Utah and Sweden have found a familial effect for cancers at all of these sites.⁴⁵⁻⁴⁷

If we consider that the contribution of inherited genetic factors to the causation of these types of cancer is indeed 27 to 42 percent, and that single-gene mutations in familial cancer syndromes account for 1 to 15 percent of all cancers,^{1,3,40} then there must be major gaps in our understanding of the genetic

basis of colorectal, breast, and prostate cancer. The frequency of mutations in the known high-risk susceptibility genes — *BRCA1* and *BRCA2* in breast cancer, DNA mismatch-repair genes in hereditary nonpolyposis colorectal cancer, and the candidate gene *HPC1* in prostate cancer — is too low to explain more than a fraction of the genetic effects we found. For example, in a recent study of 12 pairs of Swedish monozygotic twins who were concordant for breast cancer (and who were also included in the present study), 2 pairs had a *BRCA2* mutation and none had a *BRCA1* mutation (unpublished data). Our findings suggest that other genes are yet to be identified, but because they are likely to be relatively common and carry only a moderate risk, proving that they are involved in causing cancer will be difficult.²

Although model fitting can be used to estimate the magnitude of the heritable component of susceptibility to cancer, it cannot reveal how this component acts or how it interacts with other factors. For example, a cancer gene could be expressed without any environmental influence or only when activated by environmental factors. For this reason, we cannot exclude a modifying effect of environment on the genetic component found in our analyses of twins. However, without specific environmental measurements, interactions cannot be assessed. For colorectal, breast, and prostate cancer, the estimated hereditary components were slightly higher in the younger than in the older groups; this finding is in accordance with observations that hereditary effects are strongest in early-onset cancers.^{5,7,45,46,48}

The absolute risk of the same cancer before the age of 75 years for the monozygotic twin of a person with colorectal, breast, or prostate cancer was between 11 percent and 18 percent. For dizygotic twins, who have the same degree of genetic similarity as full siblings, the risk of these cancers was 3 to 9 percent. These figures could be valuable in providing clinical guidance not only to the twins of persons with cancer but also to other first-degree relatives.

One limitation of our study is that, despite its size, it did not have enough power to distinguish heritable genetic effects from environmental factors as causes of the familial aggregation of the less common types of cancer. The main reason for this limitation is that birth-cohort and calendar-period restrictions of twin registries set limits for analyses. Because errors in determining zygosity or diagnoses of cancer lead to an overestimation of nonshared environmental effects, the familial effects are, if anything, underestimated. The incidence of cancer among twins in the Finnish study did not differ from that in the general population,¹⁵ and in the Danish study, twins who responded to the questionnaire had the same distribution of zygosity and incidence of cancer as those who did not respond.⁴⁹ Thus, bias due to selective response rates is improbable.

We conclude that the overwhelming contributor to the causation of cancer in the populations of twins that we studied was the environment. For some forms of cancer, in which a shared environment is important, it may be possible to find clues in studies of childhood environment or long-lasting family habits. The relatively large heritability proportions for cancers at some sites, despite the wide confidence intervals, suggest major gaps in our understanding of heritable cancer. Even for cancers for which there is statistically significant evidence of a heritable component, most pairs of twins were discordant for the cancer — indicating that, on the population level, the increase in the risk of cancer even among close relatives of affected persons is generally moderate.

Supported by the Nordic Cancer Union. The Swedish Twin Registry is supported by grants from the John D. and Catherine T. MacArthur Foundation and the Swedish Council for Planning and Coordination of Research. The Danish twin study was supported by research grants from the Danish Cancer Society (36/79), the National Cancer Institute (R35 CA 42581), and the National Institute on Aging (PO1-AG08761). The Finnish part of the study was supported by grants from the Finnish Cancer Organizations and the Academy of Finland.

We are indebted to all the participants and staff of the Nordic twin studies, without whom this study would not have been possible.

REFERENCES

- Li FP. Phenotypes, genotypes, and interventions for hereditary cancers. *Cancer Epidemiol Biomarkers Prev* 1995;4:579-82.
- Easton DF. The inherited component of cancer. *Br Med Bull* 1994;50:527-35.
- Fearon ER. Human cancer syndromes: clues to the origin and nature of cancer. *Science* 1997;278:1043-50.
- Lander ES, Schork NJ. Genetic dissection of complex traits. *Science* 1994;265:2037-48. [Erratum, *Science* 1994;266:353.]
- Thompson WD. Genetic epidemiology of breast cancer. *Cancer* 1994;74:Suppl:279-87.
- Peto J, Easton DF, Matthews FE, Ford D, Swerdlow AJ. Cancer mortality in relatives of women with breast cancer: the OPCS Study. *Int J Cancer* 1996;65:275-83.
- Lesko SM, Rosenberg L, Shapiro S. Family history and prostate cancer risk. *Am J Epidemiol* 1996;144:1041-7.
- Gruber SB, Thompson WD, Cancer and Steroid Hormone Study Group. A population-based study of endometrial cancer and familial risk in younger women. *Cancer Epidemiol Biomarkers Prev* 1996;5:411-7.
- Easton DF, Matthews FE, Ford D, Swerdlow AJ, Peto J. Cancer mortality in relatives of women with ovarian cancer: the OPCS Study. *Int J Cancer* 1996;65:284-94.
- Auranen A, Grénman S, Mäkinen J, Pukkala E, Sankila R, Salmi T. Borderline ovarian tumors in Finland: epidemiology and familial occurrence. *Am J Epidemiol* 1996;144:548-53.
- Aarino M, Sankila R, Pukkala E, et al. Cancer risk in mutation carriers of DNA-mismatch-repair genes. *Int J Cancer* 1999;81:214-8.
- Auranen A, Pukkala E, Mäkinen J, Sankila R, Grénman S, Salmi T. Cancer incidence in the first-degree relatives of ovarian cancer patients. *Br J Cancer* 1996;74:280-4.
- Neale MC, Cardon LR. Methodology for genetic studies of twins and families. Dordrecht, The Netherlands: Kluwer Academic, 1992.
- Ahlbom A, Lichtenstein P, Malmström H, Feychting M, Hemminki K, Pedersen NL. Cancer in twins: genetic and nongenetic familial risk factors. *J Natl Cancer Inst* 1997;89:287-93.
- Verkasalo PK, Kaprio J, Koskenvuo M, Pukkala E. Genetic predisposition, environment and cancer incidence: a nationwide twin study in Finland, 1976-1995. *Int J Cancer* 1999;83:743-9.
- Page WF, Braun MM, Partin AW, Caporaso N, Walsh P. Heredity and prostate cancer: a study of World War II veteran twins. *Prostate* 1997;33:240-5.
- Braun MM, Caporaso NE, Page WF, Hoover RN. Genetic component of lung cancer: cohort study of twins. *Lancet* 1994;344:440-3.

18. Holm NV, Hauge M, Harvald B. Etiologic factors of breast cancer elucidated by a study of unselected twins. *J Natl Cancer Inst* 1980;65:285-98.
19. Holm NV, Hauge M, Jensen OM. Studies of cancer aetiology in a complete twin population: breast cancer, colorectal cancer and leukaemia. *Cancer Surv* 1982;1:17-32.
20. Cederlöf R, Lorich U. The Swedish Twin Registry. In: Nance WE, Allen G, Parisi P, eds. *Twin research. Part B. Biology and epidemiology*. Vol. 24B of *Progress in clinical and biological research*. New York: Alan R. Liss, 1978:189-95.
21. Hauge M, Harvald B, Fischer M, et al. The Danish Twin Register. *Acta Genet Med Gemellol (Roma)* 1968;17:315-32.
22. Hauge M. The Danish Twin Register. In: Mednick SA, Baert AE, eds. *Prospective longitudinal research: an empirical basis for the primary prevention of psychosocial disorders*. Oxford, England: Oxford University Press, 1981:217-21.
23. Kyvik KO, Christensen K, Skytthe A, Harvald B, Holm NV. The Danish Twin Register. *Dan Med Bull* 1996;43:467-70.
24. Clemmesen J. *Statistical studies in the aetiology of malignant neoplasms. Vol. 1. Review and results*. Copenhagen, Denmark: Munksgaard, 1965.
25. Storm HH, Michelsen EV, Clemmensen IH, Pihl J. The Danish Cancer Registry: history, content, quality and use. *Dan Med Bull* 1997;44:535-9.
26. Storm HH, Pihl J, Michelsen E, Nielsen AL. *Cancer incidence in Denmark 1993*. Copenhagen, Denmark: Danish Cancer Society, 1996.
27. Kaprio J, Sarna S, Koskenvuo M, Rantasalo I. The Finnish Twin Registry: formation and compilation, questionnaire study, zygosity determination procedures, and research program. *Prog Clin Biol Res* 1978;24:179-84.
28. Teppo L, Pukkala E, Lehtonen M. Data quality and quality control of a population-based cancer registry: experience in Finland. *Acta Oncol* 1994;33:365-9.
29. Cederlöf R, Friberg L, Jonsson E, Kaij L. Studies on similarity diagnosis in twins with the aid of mailed questionnaires. *Acta Genet (Basel)* 1961;11:338-62.
30. Sarna S, Kaprio J, Sistonen P, Koskenvuo M. Diagnosis of twin zygosity by mailed questionnaire. *Hum Hered* 1978;28:241-54.
31. Kuritz SJ, Landis JR, Koch GG. A general overview of Mantel-Haenszel methods: applications and recent developments. *Annu Rev Public Health* 1988;9:123-60.
32. McGue M. When assessing twin concordance, use the probandwise not the pairwise rate. *Schizophr Bull* 1992;18:171-6.
33. Heath AC, Neale MC, Hewitt JK, Eaves LJ, Fulker DW. Testing structural equation models for twin data using LISREL. *Behav Genet* 1989;19:9-35.
34. Sham P. *Statistics in human genetics*. London: Arnold, 1997.
35. Martin N, Boomsma D, Machin G. A twin-pronged attack on complex traits. *Nat Genet* 1997;17:387-92.
36. Neale MC. *Mx: statistical modeling*. 5th ed. Richmond: Virginia Commonwealth University, 1999.
37. Falconer DS. The inheritance of liability to certain diseases, estimated from the incidence among relatives. *Ann Hum Genet* 1965;29:51-76.
38. *Cancer: causes, occurrence and control*. Lyon, France: International Agency for Research on Cancer, 1990. (IARC scientific publications no. 100.)
39. Neale MC, Eaves LJ, Kendler KS. The power of the classical twin study to resolve variation in threshold traits. *Behav Genet* 1994;24:239-58.
40. Lynch HT, Smyrk T. Hereditary nonpolyposis colorectal cancer (Lynch syndrome): an updated review. *Cancer* 1996;78:1149-67.
41. Swerdlow AJ, De Stavola BL, Swanwick MA, Maconochie NES. Risks of breast and testicular cancers in young adult twins in England and Wales: evidence on prenatal and genetic aetiology. *Lancet* 1997;350:1723-8.
42. Mack TM, Cozen W, Shibata DK, et al. Concordance for Hodgkin's disease in identical twins suggesting genetic susceptibility to the young-adult form of the disease. *N Engl J Med* 1995;332:413-8.
43. Grönberg H, Damber L, Damber J-E. Studies of genetic factors in prostate cancer in a twin population. *J Urol* 1994;152:1484-9.
44. Braun MM, Caporaso NE, Page WF, Hoover RN. A cohort study of twins and cancer. *Cancer Epidemiol Biomarkers Prev* 1995;4:469-73.
45. Goldgar DE, Easton DF, Cannon-Albright LA, Skolnick MH. Systematic population-based assessment of cancer risk in first-degree relatives of cancer probands. *J Natl Cancer Inst* 1994;86:1600-8.
46. Hemminki K, Vaittinen P, Kyrrönen P. Age-specific familial risks in common cancers of the offspring. *Int J Cancer* 1998;78:172-5.
47. Hemminki K, Vaittinen P. Familial cancers in a nationwide family cancer database: age distribution and prevalence. *Eur J Cancer* 1999;35:1109-17.
48. Grönberg H, Smith J, Emanuelsson M, et al. In Swedish families with hereditary prostate cancer, linkage to the HPC1 locus on chromosome 1q24-25 is restricted to families with early-onset prostate cancer. *Am J Hum Genet* 1999;65:134-40.
49. Holm NV. Application of twin studies in the elucidation of causal relation in diseases of complex etiology with cancer as an example. (Ph.D. thesis. Odense, Denmark: Odense University, 1983.) (In Danish.)