Agreement Between Breast Cancer Risk Estimation Methods

Kimberly A. McGuigan, Patricia A. Ganz, Claudine Breant

Accurate identification of women at high risk of developing breast cancer has important uses in both research and clinical settings. Identification of high-risk groups to study chemopreventive agents requires a reliable and valid risk projection model (1). Targeting chemopreventive agents to those at highest risk decreases the required sample size for clinical trials and improves the risk/benefit ratio for participants (2). Clinical counseling is also facilitated by more specific risk information, because many women overestimate their actual risk of breast cancer (3,4).

Two model-based approaches to individualized risk prediction have been offered for use in research and clinical applications. The first model was developed by Gail et al. (5) (Gail model) using case and control subjects from the Breast Cancer Detection Demonstration Project (BCDDP). The second model, by Claus et al. (6) (Claus model), was developed from subjects in the Cancer and Steroid Hormone (CASH) Study.

Risk estimation from the Gail model is based on indicators of endogenous estrogen exposure, family history, and clinical risk (5,7,8). The Claus model uses genetic modeling to tabulate risk on the basis of the number and age at diagnosis of first- and second-degree relatives previously diagnosed with breast cancer. These predictions are based on the assumption of a single autosomal dominant gene for breast cancer.

Validation studies of the Gail model indicate some discrepancies between numbers of projected and actual cases, overpredicting risk for premenopausal women and for women with "extensive" family history (9) and for women who did not adhere to screening guidelines (10). No validation work has been reported for the Claus model.

Reliability, or agreement between these two methods, has not yet been assessed. In contrast to validation, where predicted risk is compared with observed clinical diagnosis of disease (9,10), reliability indicates how closely two measurements agree. For both scientific and clinical counseling applications, it is important to establish that two methods which purport to measure the same underlying construct—here, probability of developing breast cancer—give similar estimates for the same subjects.

In this study, we compare the estimated risk for both the Gail and Claus models using previously collected data from women attending the high-risk breast cancer clinic of the University of California at Los Angeles from 1992 through 1994.

One hundred thirty-nine women had computerized records reporting family history, risk factors, and clinical information. Of these, 128 had sufficient information to calculate risk based on the Gail model, and 116 had a first- or second-degree relative with breast cancer for risk estimation by the Claus model. A total of 111 women had risk estimates for both methods.

We calculated remaining lifetime risk (to age 80 years for the Gail method and to age 79 years for the Claus method) for all subjects. This risk is expressed as the probability of developing breast cancer by age 80 (the Gail method) or by age 79 (the Claus method), conditional on no previous breast cancer diagnosis at the woman's present age. We obtained point estimates and 95% confidence intervals (CLs) (11) for the Gail method and point estimates only for the Claus method, because interval estimation for the Claus method was not available.

To assess concordance, we used two approaches. First, we calculated chance-corrected agreement between methods using both parametric and nonparametric statistics. Intraclass correlation (ICC) was used to determine concordance between the point estimates (12). Similarly, ICC for ranks was calculated to assess how well the relative orderings of the two methods agree. The ICC for point estimates was 0.43 (P<.001), whereas the corresponding measure for ranks was slightly higher, 0.55 (P<.001). Second, an indicator of concordance was constructed by identifying whether the Claus estimate fell within the 95% CI about the Gail point estimate (Fig. 1). Only 22% of the subjects showed agreement by use of this criterion, whereas 19% fell above the 95% CI and 59% fell below the 95% CI. Using this indicator, where 1 indicates that the Claus estimate is contained within the Gail 95% confidence region and 0 indicates otherwise, we regressed this outcome on the set of risk factors listed in Table 1 to identify determinants of agreement. Logistic regression showed the greatest discrepancies between risk estimates for women with no live births (odds ratio = 0.21; 95% CI = 0.07-0.66) and for those who had had a previous breast biopsy (odds ratio = 0.41; 95% CI = 0.21-0.93). Greater concordance was seen for women who had more than one first-degree relative diagnosed with breast cancer (odds ratio = 7.24; 95% CI = 2.03-25.89). Overall, these results reflect differences in risk factors considered by each of the two models. These findings are consistent with those of validation studies, where the Gail model overpredicted risk for some subgroups of subjects (9,10). Here, when
The implications of these results pertain to both clinical and research applications of these instruments. Clinicians who use either method for counseling purposes should be aware that the two approaches may result in substantially different estimates of breast cancer risk for some patients. For researchers, these comparisons highlight the need to distinguish and report sources of uncertainty in risk estimation. These sources include model uncertainty, as indicated by variation around point estimates, as well as consistency across methods that measure the same underlying construct.

**Table 1. Descriptive statistics for risk factors for women at high risk of developing breast cancer**

<table>
<thead>
<tr>
<th>Risk factor</th>
<th>Mean (standard error)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, y</td>
<td>43.9 (1.1)</td>
</tr>
<tr>
<td>Age at menarche, y</td>
<td>12.6 (0.1)</td>
</tr>
<tr>
<td>Age at first birth, y</td>
<td>25.6 (0.3)</td>
</tr>
<tr>
<td>% who have had no live births</td>
<td>49.5 (4.8)</td>
</tr>
<tr>
<td>% with first-degree relative with breast cancer</td>
<td>95.5 (2.0)</td>
</tr>
<tr>
<td>Age of first-degree relative with breast cancer (n = 106)</td>
<td>49.2 (1.1)</td>
</tr>
<tr>
<td>% with second-degree relative with breast cancer</td>
<td>20.7 (3.9)</td>
</tr>
<tr>
<td>Age of second-degree relative with breast cancer (n = 19)</td>
<td>60.5 (3.6)</td>
</tr>
<tr>
<td>% with one or more biopsies</td>
<td>41.4 (4.7)</td>
</tr>
<tr>
<td>% with atypical hyperplasia</td>
<td>2.7 (1.5)</td>
</tr>
</tbody>
</table>

*A total of 111 women had risk estimates for both methods.
†Four subjects did not know, or did not report, age of second-degree relative with breast cancer.

**References**

Notes

Supported by the Revlon/UCLA Women's Health Research Program. K. A. McGuigan was supported by Public Health Service National Research Service Award 1T32HS00046-01 from the Agency for Health Care Policy and Research, Office of the Assistant Secretary for Health, Department of Health and Human Services, and by Public Health Service grant CA49565 from the National Cancer Institute, National Institutes of Health, Department of Health and Human Services. Database development was supported by a grant from the Drown Foundation.

We thank the professional staff of the High Risk Clinic at the Revlon/UCLA Breast Center—S. Goldman, Dr. D. Wellisch, Dr. K. Duvall, Dr. A. Hoffman, and the center director Dr. S. M. Love. Manuscript received March 14, 1996; revised May 22, 1996; accepted June 17, 1996.

Hereditary nonpolyposis colorectal cancer (HNPCC) is one of the most common autosomal dominantly inherited diseases, affecting as many as one in every 200-400 individuals in the Western world. HNPCC is responsible for up to 15% of all colorectal cancers (1). It has recently been shown that the mismatch repair genes, hMSH2, hMLH1, hPMS1, and hPMS2, are mutated in the germline of affected members in HNPCC families (2-6), as well as for some apparently nonhereditary cancer patients (7-9). A combination of linkage and mutational analysis has indicated that hMSH2 and hMLH1 are likely to be the prevalent genes responsible for HNPCC, accounting for 80%-90% of HNPCC cases (10-13), while hPMS1 and hPMS2 are thought to account for only a minor fraction of cases. The identification of the causative mutations in HNPCC families is desirable, since it allows the carrier status of unaffected relatives at risk to be determined.

The Korean Hereditary Colorectal Cancer Registry was established in 1991, and we have registered HNPCC families on the basis of the minimal criteria proposed by the International Collaborative Group on HNPCC (ICG-HNPPC) (14). We have also registered families who do not fulfill the ICG-HNPPC criteria but where a genetic basis of colon cancer is strongly suspected (categorized as suspected HNPCC families) because of the following features: 1) vertical transmission of colorectal cancer or at least two siblings affected with colorectal cancer in a family; and 2) development of multiple colorectal tumors or at least one colorectal cancer case diagnosed before the age of 50 years. Data concerning the early-onset patients who had developed colorectal cancer before the age of 40 years without any family history of disease were also collected through the Department of Surgery, Seoul National University Hospital.

To investigate the genetic status of hMSH2 and hMLH1 genes in 25 Korean HNPCC kindreds, 17 suspected HNPCC patients, and 22 early-onset colorectal cancer patients, single-strand conformation polymorphism (SSCP) was used to screen for the mutations, followed by sequencing of the DNA fragments displaying abnormal SSCP pattern. Genomic DNAs were prepared from white blood cells as described by Blin and Stafford (15). The polymerase chain reaction primers and methods for amplification of each exon in hMSH2 and hMLH1 and sequence analysis have been previously described (10,13).

The germline mutations detected in this study are summarized in Table 1. A total of 13 germline mutations were detected by SSCP analysis.

The frequency of mutations in hMSH2 and hMLH1 for the patients is given in Table 2. Among 25 classic HNPCC families, we found eight mutations in the hMLH1 gene (32%) but none in the hMSH2 gene. In particular, codon 586 of exon 16 in the hMLH1 gene was the frequently mutated site, accounting for four of 12 mutations.