USES AND ABSUSES OF THE "GAIL" MODEL FOR PROJECTING ABSOLUTE BREAST CANCER RISK

Mitchell H. Gail

Biostatistics Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD 20892, USA

Abstract

We review the concept of absolute risk and its applications, using breast cancer as the leading example. The strengths and weakness of the Gail model for projecting absolute breast cancer are discussed in relation to applications of the model. Among the more promising applications are designing breast cancer prevention trials, whose power depends on average absolute breast cancer risk, counseling women regarding their risk, and offering guidance on weighing the risks and benefits of interventions such as tamoxifen administration to prevent breast cancer. Because of the limited ability of the model to discriminate women who will develop breast cancer from those who will not, the model should not be used, for example, to screen the population for women who would receive a more definitive breast cancer evaluation; such a strategy would prevent many women who will eventually develop breast cancer from receiving the more definitive evaluation. Scope for future research includes including more powerful risk factors in such a model to improve its discriminatory power, validating and if necessary modifying the model for use in minority populations, and developing individualized absolute risk models for other diseases to assist in weighing risks and benefits of interventions.

KEY WORDS: Absolute risk; crude risk; competing risk; risks and benefits; breast cancer.

1. INTRODUCTION

Gail et al. (1989) used data from the Breast Cancer Detection Demonstration Project (BCDDP) to develop a model that estimates the absolute risk of breast cancer over a defined age interval for a woman with certain risk factors. A modified version of this model has been used to design and define the eligibility criteria for a prevention trial of tamoxifen (Fisher et al.,1998), to counsel women on breast cancer risk, and to provide guidance concerning the clinical decision of whether or not a woman should take tamoxifen to prevent breast cancer. In this article, we define the concept of absolute risk and some of its potential applications, and we

discuss validation data and the strengths and limitations of the "Gail" model, together with some possible ways to improve the model. We will use the example and terminology of breast cancer throughout, although the concept of absolute risk can be applied to any disease.

2. DEFINITON OF ABSOLUTE RISK

Absolute risk is the probability that a breast cancer will be diagnosed in a defined age interval for a woman with a particular set of risk factors. Absolute risk is sometimes called "crude" risk, because it is affected by competing causes of death. If we let $h_1(t)$ be the hazard of breast cancer incidence at age t for a woman with no risk factors, $h_1(t) \exp\{\beta^T X(t)\}$ be the hazard of breast cancer for a woman with risk factors X(t), and $h_2(t)$ be the hazard of dying of non-breast cancer causes, then the absolute risk of developing breast cancer in the age interval [a, a + τ) is

$$\int_a^{a+\tau} h_1(t)\exp\{\beta^T X(t)\}\exp[-\int_a^t [h_1(u)\exp\{\beta^T X(u)\} + h_2(u)]du]dt$$

The absolute risk increases with the width of the age interval, τ, and with age a, because $h_1(t)$ increases with age. A large hazard of competing risks $h_2(t)$ results in reduced absolute risk of breast cancer. Because one cannot predict future values of risk factors X(t), one usually regards the factors as fixed at their initial value at age a, X(a). Nonetheless, equation (1) does not require a proportional hazards assumption nor that covariates be fixed, because time-dependent covariates are permitted. The Gail model includes an interaction between age and number of biopsies; proportional hazards is assumed only for those under age 50 and separately for those aged 50 and above.

As described in the Appendix to Costantino et al (1999), the Gail model was modified by Stewart J. Anderson and Carol K. Redmond to reflect breast cancer rates in the National Cancer Institute's Surveillance and End Results (SEER) study population in the United States, rather than in the BCDDP. This modified model can be obtained through http://brca.nci.nih.gov/brc/. As an example, a 50 year old white woman wants to know her projected risk to ages 55 and 90. She began menstruating at age 11, had a first live birth at age 19, had one previous breast biopsy with unknown atypical hyperplasia status, and had a sister with breast cancer. From the program, the projected absolute risks are 2.5% and 21.6% respectively. Approximate 95% confidence intervals are, respectively, (1.9%, 3.3%) and (16%, 28%), as estimated from Figure 3 in Benichou et al. (1996). The 21.6% risk is high

in this case because the age interval is long and because there is one affected mother or sisters.

As is indicated from this example, the risk factors used in the Gail model are age at menarche, number of breast biopsies (with different relative risks for those below and at or above age 50), whether or not atypical hyperplasia was found on a biopsy, age at first live birth, and number of mother or sisters with a breast cancer diagnosis. The Claus model (Claus et al, 1994) for projecting breast cancer risk, includes more detailed information on family history, but no information on reproductive and medical history.

3. VALIDATION AND EVALUATION OF THE GAIL MODEL

Constantino et al (1999) recently evaluated the performance of the modified Gail model on independent data from the Cancer and Hormone Study (CASH) (Wingo et al, 1988), the Breast Cancer Prevention Trial (BCPT) (Fisher et al., 1998), and the Nurses Health Study (NHS) (Rockhill et al, 2001). They found substantial agreement in relative risks, represented by $\exp(\beta)$ for the parameters in equation (1), when the model was refitted to the various independent data sets. The effects of family history, however, were somewhat stronger in the BCDDP data (Gail model) and in CASH data, which were derived from case-control studies, than in NHS, a cohort study.

In addition to checking relative risk features of the model, Costantino et al (1999) and Rockhill et al (2001) compared the absolute risk estimate from the model ("expected" counts, E) with the numbers of observed breast cancer cases (O) in subgroups of women in the BCPT and in the NHS. For all women in the BCPT, the E/O ratio was 1.03, with 95% confidence interval (CI) of 0.88-1.21. Corresponding E/O ratios (with 95% CI) were 0.93 (0.72-1.22) for women under age 50, 1.13 (0.83-1.55) for women aged 50 to 59, and 1.05 (0.80-1.41) for women aged 60 and over. Data from Costantino et al. show that the modified Gail model predicted risk just as well in high risk women, such as those with a strong family history, as in low risk women. Rockhill et al (2001) likewise found good agreement between observed and expected counts in the NHS. The overall E/O ratio was 0.94 with 95% CI 0.89-0.99, and agreement was also good in higher risk women.

Although the modified Gail model predicts risk well groups of women in these studies, it is important to recall that the model was based on white women in the BCDDP, and the validation data is preponderantly among white women. Therefore, there is a need or further validation studies in minority populations and perhaps for modifications of the model for such populations. Moreover, the model

was not designed for women who had had a previous breast cancer or ductal or lobular carcinoma *in situ*. The model would also be misleading for women from certain parts of the world, such as rural China where breast cancer rates are low. A woman who is known to carry a disease-producing mutation in BRCA1 or BRCA2 or who has rare hereditary condition such as the Li-Fraumeni syndrome that places her at increased risk of breast cancer should seek genetic counseling to take advantage of this special information, rather than rely solely on the Gail model to project risk. These exceptions indicate that it is advisable that a health care provider who is aware of the limitations of the Gail model be available to assist in the interpretation of projections from the model.

Even though Rockhill et al. (2001) found that the modified Gail model predicted observed counts well in subgroups of women, they noted that the model had limited ability to discriminate which particular women in NHS would develop breast cancer and which women would not develop breast cancer. If one plots histograms of the projected risks in women who did and did not develop breast cancer, there is considerable overlap, indicating limited discriminatory power. Rockhill et al. estimated the concordance statistic, which is the area under the receiver operating curve based on varying cut-points of projected risk, as 0.58 (95% CI 0.56-0.60). This analysis takes only the discriminatory power of the risk factors apart from age into account. Our unreported calculations based on data from the 2000 National Health Interview Survey (NHIS) indicate a concordance near 0.75, if one also includes age as a risk factor in this analysis. It seems more appropriate to assess the discriminatory power conditional on age, however, in which case the estimated concordance from NHIS data is approximately 0.62, in line with the estimate of Rockhill et al. The limited discriminatory power of the model was also indicated by the fact that the ratio of average risk of the women in NHS in the highest decile of risk compared to that of women in the lowest decile of risk was only 2.83 (95% CI 2.19-3.65).

Thus the Gail model predicts the numbers of observed breast cancers that will develop in subgroups of women well but has limited ability to identify precisely who will develop breast cancer. These features dictate for which applications the model is best suited, as well as some inappropriate applications.

4. SOME APPROPRIATE AND INAPPROPRIATE USES OF THE GAIL MODEL

The strength of the Gail model is its ability to accurately predict the probability of developing breast cancer. One successful application of the model

is therefore to aid in designing breast cancer prevention trials. The power and required sample size of such trials depends on the numbers of breast cancers that will eventually develop during the study. The expected numbers of breast cancers that develop is the sample size times the average absolute breast cancer risk of the study participants. This idea was used to calculate the sample size needed for the BCPT.

Having estimates of absolute breast cancer risks with 5, 10 and 30 year time horizons can be useful in the counseling women who have concerns about breast cancer. Many women seriously overestimate their breast cancer risk and are anxious about it.

Absolute risk estimates are useful in counseling women about the risks and benefits of a particular intervention to prevent breast cancer, because the commensurate scale for weighing risks of various health outcomes is absolute risk. For example, tamoxifen reduces the risk of invasive breast cancers, *in situ* breast cancers, and hip fractures, but increases the risks of endometrial cancer, stroke, pulmonary emboli and deep vein thrombosis. Gail et al. (1999) assigned weights of 1.0 to life threatening events (invasive breast cancer, hip fracture, endometrial cancer, stroke, pulmonary embolism) and 0.5 to severe events (*in situ* breast cancer, deep vein thrombosis) and combined these weights with estimates of the absolute risks of various outcomes in the presence and absence of tamoxifen to determine a net benefit/risk index for women categorized by race/ethicity, age, and risk factors in the Gail model. Young women with high breast cancer risk had positive benefit /risk indices. Older women often had negative benefit/risk indices because the increased risks from endometrial cancer and cardiovascular events outweighed the benefits of reduced breast cancer and hip fracture risk. The methodology requires accurate prediction of the absolute risks of various outcomes in the absence of tamoxifen. The Gail model is well validated, but absolute risk estimates for outcomes such as stroke and hip fracture were based only on age and ethnicity, because the literature still does not have validated individualized absolute risk models for many endpoints. Using the benefit/risk indices developed by Gail et al (1999) and U.S. national survey data, Freedman et al. (2003) estimated that whereas 18.7% of white women aged 35-79 in the U.S. were eligible to take tamoxifen to prevent breast cancer according to Food and Drug Administration labeling criteria, only 4.9% had a positive benefit/risk ratio.

By combining estimates of the prevalence of risk factor combinations in a population with absolute risk estimates, one can gauge the burden of disease in a population. If some of the risk factors in the absolute risk model are modifiable, one can also use such models to estimate the health impact on the population from changing the prevalence of risk factors in the population.

The previous applications of absolute risk models depend on the ability of the model to predict absolute risk accurately, as reflected in good agreement between expected and observed disease incidence in groups. Some potential applications of absolute risk models would be inappropriate for models with limited discriminatory power, such as the Gail model. For example, it would be inappropriate to use the Gail model to screen the entire population and then to offer further, more definitive screening for breast cancer, such as mammography, only to those women who had an estimated absolute risk above a particular level. This strategy would prevent many women who will eventually develop breast cancer from receiving the more definitive evaluation.

5. DISCUSSION

Absolute risk is a very useful quantity for designing studies and as an aid to clinical decision-making and counseling. Although the Gail model has been shown to predict risk accurately in independent validation data, further work is needed to validate and, if necessary, modify the model for minority populations. Incorporating other strong risk factors, such as mammographic density (Byrne et al., 1995) may improve the ability of the model to discriminate between women who will and will not develop breast cancer. To aid in weighing risks and benefits, it would be useful to develop and validate individualized absolute risk models for other health outcomes, such as stroke and endometrial cancer. From a methodologic perspective, it would be interesting to explore the possibility of developing joint models for the evolution of risk factor values as well as disease outcome over time, rather than having to make simplifying assumptions, such as the assumption that X(t) will remain constant at its initial value, X(a), in equation (1).

REFERENCES

BENICHOU J., GAIL M.H., MULVIHILL J.J., 1996: Graphs to estimate an individualized risk of breast cancer. *Journal of Clinical Oncology*, 14,103-110.

BYRNE C., SCHAIRER C., WOLFE J., PAREKH N., SALANE M., BRINTON L.A., HOOVER R., HAILE R., 1995: Mammographic features and breast cancer risk: effects with time, age, and menopause status. *Journal of the National Cancer Institute*, 87,1622-1629.

CLAUS E.B., RISCH N., THOMPSON W.D., 1994: Autosomal dominant inheritance of early-onset breast cancer. Implications for risk prediction. *Cancer* ,73, 643-651.

COSTANTINO J.P., GAIL M.H., PEE D., ANDERSON S., REDMOND C.K., BENICHOU J., WIEAND H.S., 1999: Validation studies for models projecting the risk of invasive and total breast cancer incidence. *Journal of the National Cancer Institute*, 91,1541-1548.

FISHER B., COSTANTINO J.P., WICKERHAM D.L., REDMOND C.K., KAVANAH M., CRONIN W.M., VOGEL V., ROBIDOUX A., DIMITROV N., ATKINS J., DALY M., WIEAND S., TAN-CHIU E., FORD L., WOLMARK N., 1998: Tamoxifen for prevention of breast cancer: report of the National Surgical Adjuvant Breast and Bowel Project P-1 Study. *Journal of the National Cancer Institute* 90, 1371-1388.

FREEDMAN A.N., GRAUBARD BI., RAO S.R., MCCASKILL-STEVENS W., BALLARD-BARBASH R, GAIL M.H., 2003: Estimates of the number of US women who could benefit from tamoxifen for breast cancer chemoprevention. *Journal of the National Cancer Institute,* 95, 526-532.

GAIL M.H., BRINTON L.A., BYAR D.P., CORLE D.K., GREEN S.B., SCHAIRER C., MULVIHILL J.J., 1989: Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. *Journal of the National Cancer Institute,* 81, 1879-1886.

GAIL M.H., COSTANTINO J.P., BRYANT J., CROYLE R., FREEDMAN L., HELZLSOUER K., VOGEL V., 1999: Weighing the risks and benefits of tamoxifen treatment for preventing breast cancer. *Journal of the National Cancer Institute,* 91, 1829-1846.

ROCKHILL B., SPIEGELMAN D., BYRNE C., HUNTER D.J., COLDITZ G.A., 2001: Validation of the Gail et al. model of breast cancer risk prediction and implications for chemoprevention. *Journal of the National Cancer Institute,* 93, 358-366.

WINGO P.A., ORY H.W., LAYDE P.M., LEE N.C.,1988: The evaluation of the data collection process for a multicenter, population-based, case-control design. *American Journal of Epidemiology*, 128, 206-217.

USO E ABUSO DEL MODELLO DI GAIL PER LA PROIEZIONE DEL RISCHIO ASSOLUTO DEL TUMORE

Riassunto

In questo lavoro viene rivisto il concetto di rischio assoluto e delle sue applicazioni, utilizzando come riferimento il caso del tumore alla mammella. Saranno discussi vantaggi e limiti dell'uso del Modello di Gail nella proiezione del rischio assoluto per tumore alla mammella in diversi ambiti applicativi. Tra le applicazioni più promettenti del modello sono considerate: il disegno di studi di prevenzione per tumore alla mammella, la cui potenza dipende dal valore medio del rischio assoluto per questo tumore nella popolazione in studio e l'attività di counseling per donne che tenendo conto del previsto rischio individuale, permetta di definire procedure di sorveglianza che pesino opportunamente i rischi e i benefici di politiche d'intervento quali ad esempio la somministrazione di tamoxifen allo scopo di prevenire l'insorgenza di tumore alla mammella. Come conseguenza della limitata capacità del modello di discriminare donne che svilupperanno tumore alla mammella da quelle che tale tumore non svilupperanno, il modello non dovrebbe essere utilizzato, per esempio, per identificare nella popolazione donne da sottoporre ad una più attenta valutazione per tumore alla mammella. Infatti tale strategia escluderebbe da una più attenta politica di sorveglianza molte donne che successivamente svilupperanno il

tumore alla mammella. Obiettivi per ricerche future comprendono l'inclusione nel modello di fattori di rischio più rilevanti al fine di aumentarne il potere discriminante, la validazione ed eventuale modifica del modello per sottogruppi di popolazione e lo sviluppo di modelli per la stima del rischio assoluto per altre malattie al fine di contribuire alla definizione di strategie di prevenzione in grado di meglio pesare i vantaggi e i rischi di politiche di intervento.