

*Hypothesis/Commentary***Epidemiology and Prevention of Breast Cancer****Graham A. Colditz**

Channing Laboratory, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, The Harvard Center for Cancer Prevention, Boston Massachusetts and School of Population Health, University of Queensland, Brisbane, Australia

For over 140 years, medical scientists have searched for ways to identify women at high risk of breast cancer from those not at elevated risk (1). Recently, we have seen this search come to a logical end point with a proof of principle that risk reduction is in fact possible through selective use of chemopreventive agents (2, 3). Here I review some of the history of the search, summarize what we know, and what we might do to improve our identification of women at risk. Finally, I summarize some of the strategies that we could already be pursuing.

The evidence that breast cancer could be preventable comes from studies of international variation of breast cancer (4) and studies of breast cancer in migrants (5). Together, these two sources of evidence powerfully support the contention that we might prevent breast cancer. Alas, it is very common to see articles conclude with statements such as "much of breast cancer epidemiology is not explained by known risk factors", implying that we must continue the search for new risk factors rather than take what we know and act on it.

The international variation in breast cancer incidence during the 1980s is summarized in Fig. 1 to highlight the magnitude of the difference between San Francisco, typical of U.S. incidence rates, Japan, and China. More recently, rates in Japan (6), Singapore (7), and China (8) have been rising. Using a log-scale to plot breast cancer incidence, we see a sharp inflection at around age 50. We have a 4-fold increased rate of breast cancer in San Francisco, which is typical of western society compared with traditional agrarian societies. This cross section of incidence rates from tumor registries summarizes data in 1980s; thus, the women who are in their 70s and 80s in Japan and China were born around 1900. In fact, that generation of Japanese women, who have this low rate of breast cancer in their 70s and 80s, were paid a bounty by the Japanese government when they reached their 10th child (9), as the Japanese government tried to build its army to fight the Chinese. Meanwhile, the Chinese were having six to eight children as a typical reproductive pattern.

Also of note is the rate of increase from late teen years through to 30 or 35, which is in fact quite comparable between a country like the United States, where we end up with a very high incidence rate, and the Asian countries, with significantly lower rates of breast cancer (10). This accumulation of risk will be considered in the context of models of breast cancer incidence, paying particular attention to the timing of exposures. These have been developed and applied to data

from the Nurses' Health Study (NHS) in collaboration with my colleague, Bernard Rosner.

The Nurse's Health Study began in 1976, following pilot studies conducted by Dr Speizer. The major hypothesis at the initiation of the prospective cohort study was the relation of oral contraceptives to risk of breast cancer. Additional hypotheses included smoking, postmenopausal hormone use, and hair dye use in relation to cancer in women. The National Cancer Institute has funded the study since it began.

The 121,700 participants are followed with a mailed questionnaire every 2 years, of whom ~90% respond. Updating of exposure and outcome information occurs with this questionnaire cycle, providing a unique resource, even compared with most of the other cohort studies, because of this repeatedly updated exposure information.

Drawing on these data, we have evaluated the relation between reproductive and other exposures to risk of breast cancer (11-13). The important feature of this work is that we focus on the temporal relation between the exposure and breast cancer incidence. If we can do this and improve our modeling and understanding of these relations, our hope is that we will actually create a better summary of the pathways to breast cancer and improve prediction of risk both at the population level (that is the total number of cases expected), and at the individual level, so that we might identify high risk women more accurately. By focusing on the timing of exposure, we might also identify areas that could maximize reduction in risk and hopefully lower the burden of breast cancer.

The concepts underlying modeling were set forth by Armitage and Doll (14), describing a multistage theory of carcinogenesis. Moolgavcar applied this approach to breast cancer (10) and noted the increase in incidence in premenopausal women, which is consistent around the world. Our more recent work has applied these concepts to breast cancer (11). In 1983, Pike set the stage for much of our work (15), describing a model of breast cancer incidence. It was based on the concept of breast tissue aging as the root cause of breast cancer. We are using the same approach, but think of it as an accumulation of cell division and molecular damage on the pathway to breast cancer. Tissue damage or aging starts at a constant, rate C, at menarche continues to T1 the time of the first pregnancy at which point there is a decrease in the rate of tissue aging. That continues to menopause when there is an additional decrease and after menopause, up to current age, the rate continues at a far lower level. Pike included in his model a term for an increase in risk with the first pregnancy. Pregnancy thus has dual effects: an increase in risk that is a one-time adverse effect and a decrease in the subsequent rate of tissue aging. The area under the incidence curve represents the cumulative tissue aging.

In extending the Pike breast cancer incidence model, we first focused on the effect of the first birth, which varies according to the duration from menarche to first birth (11). Petrakis encouraged us to explore this under the assumption that a later first pregnancy would allow for more time for

Cancer Epidemiol Biomarkers Prev 2005;14(4):768-72

Received 2/27/04; revised 12/2/04; accepted 12/14/04.

Grant support: CA87969 (Nurses' Health Study).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Note: G. Colditz is an American Cancer Society Clinical Research professor.

Based on the 2003 DeWitt Goodman Lecture at the AACR Meeting, July 12, 2003, Washington, DC.

Requests for reprints: Graham A. Colditz, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, 181 Longwood Avenue, Boston, MA 02115. Phone: 617-525-2258; Fax: 617-525-2008. E-mail: graham.colditz@channing.harvard.edu

Copyright © 2005 American Association for Cancer Research.

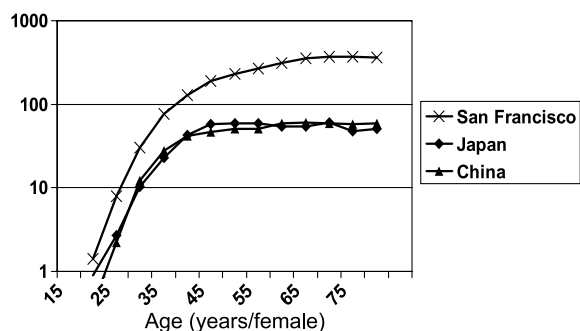


Figure 1. Breast cancer average annual incidence per 100,000 by age, 1982.

DNA damage. The proliferation due to pregnancy hormones would therefore be acting on a more damaged set of DNA and carry a greater adverse effect. In fact, this is confirmed in the incidence model (13).

Initial modeling followed the assumption that incidence is related to a power k , the number of hits required to develop cancer. Peer review feedback suggested an alternative analytic approach producing terms that are interpretable on the relative risk (RR) scale and hence more accessible to the average reader (13). The original approach gave some mathematical terms that we had trouble talking about in the context of risk and RR. Our ongoing work is now applying this model in terms of estrogen receptor (ER)-positive (ER⁺) and ER-negative (ER⁻) breast cancer (16).

Our multiple birth model, as we apply it, shows an adverse effect of the first pregnancy, a decrease in the rate of tissue aging with each subsequent pregnancy. We in fact can test for an adverse effect with the second or third pregnancy and see no significant adverse effect (11). As we shorten the interval between births, we observe a lower rate of tissue aging and incidence of breast cancer. The cumulative effect of exposure to cycling endogenous hormones of the premenopausal women is thus reduced.

One feature of our work is to focus on the magnitude of risk as it varies by age and calculate the cumulative risk to age 70 (12). We believe this is a more useful summary of risk because the effects of some risk factors vary by age; for example, body mass index reduces risk in premenopausal women but increases in risk for postmenopausal women leading to challenges in risk communication.

The spacing of births is significantly related to reduced risk; the closer the births are, the lower the risk. Plotting the age incidence curve for a typical woman with age at menarche 13 and age at natural menopause 50, we have a nulliparous woman with cumulative incidence (area under the curve) of 6.92%; having one pregnancy at age 35 has an adverse effect that persists as increased risk throughout life compared with the nulliparous women [95% confidence interval (95% CI), 8.03%]. A woman with four births at ages 20, 23, 26, and 29 has a significantly lower risk than the nulliparous woman (5.07%).

With regard to cumulative risk to age 70, we see a 16% increase in risk for the woman with one birth at age 35 and a 27% reduction in risk by having four children starting at age 20. Thus, even in the typical U.S. population, pregnancy history can contribute substantially to variation in risk and may explain much of the Black/White crossover in breast cancer incidence. A 4-year difference in age at menarche from 11 to 15 would give a 30% reduction in risk of breast cancer. Whereas we do not have any effective interventions to promote changes in age at menarche, evidence suggest that physical activity and diet may reasonably affect age at menarche. Our current young age at menarche is not a

longstanding tradition of western society. Scandinavian countries back in the 1860s, at the turn of the industrial revolution, had a mean age at menarche at 16 to 17 (ref. 17; see Fig. 1). More recent data from China in 1980 show an average age at menarche of 17 among rural Chinese women (18).

In contrast, age at menopause does not vary over time. We can see that a 10-year difference in age at natural menopause, comparing a woman with menopause at 45 to a woman at 55, lowers risk by some 44% (95% CI, 26-64). This points to the importance of the age at menopause and the effect of the reduced rate of increase in breast cancer risk after menopause. Early menopause reduces risk, and we have growing evidence that high circulating hormone levels after menopause increase risk (19). Furthermore, antiestrogen selective ER modulators may effectively reduce risk of breast cancer (2, 3), but we are left with the challenges to whom we might target such preventive therapies, and how do we balance risks and benefits (20).

The evidence for hormonal exposure after menopause starts with obesity and its negative effect on survival. We have a long history of tamoxifen as an agent to reduce mortality (21) after breast cancer and now we have evidence for tamoxifen and raloxifene as agents to prevent breast cancer (2, 3).

The other intriguing piece of evidence on the role of hormones and the potential to identify women who may benefit most by use of selective ER modulators comes from the MORE study, where the baseline estradiol level was used to stratify out the women in the raloxifene trial (22). In the placebo group, the breast cancer risk increased with baseline estradiol levels and the greatest reduction in risk among women allocated to raloxifene was observed among women with the highest baseline estradiol levels. Thus, we can start to see a model that if we can identify women either with high baseline estradiol levels or at high risk in this hormonal milieu and if we have the right agent available, we might achieve substantial reduction in risk.

If we evaluate use of exogenous postmenopausal hormones, with menarche at 13 and natural menopause at age 50, 10 years of use of estrogen plus progestin or estrogen alone lead to substantial and significant increases in risk of breast cancer, consistent with a growing body of epidemiologic data (23). In fact, the Women's Health Initiative shows the same effect for the combination estrogen plus progestin as the observational epidemiologic studies do (24, 25). We also see that women who gained <5 kg (11 lbs) because leaving high school have the most dramatic adverse effect of adding exogenous hormones so that 10 years of use of E plus P gives about a fourfold risk compared with never users (26). With greater weight gain, we observe lower RR to the point that for women who have gained over 20 kg after leaving high school, adding hormones does not do much, presumably because their circulating estrogen levels are already so high. This in itself is consistent with the raloxifene data, emphasizing the role of the underlying hormonal exposure.

Do these risk factors apply equally to ER⁺ cases, where we have reasonable treatment strategies, and ER⁻ cases? Furthermore, given the different age incidence patterns for ER⁺ and ER⁻ tumors (27), could application of this model give new insights? Several studies (28-30) have shown that nulliparity is inversely associated with risk of developing ER⁻/progesterone receptor-negative (PR⁻) tumors among postmenopausal women, whereas other studies (29, 31-33) have shown that parity was inversely associated with ER⁺ tumors but not with ER⁻, suggesting that there are differing influences of parity on incidence according to ER status of tumors. We fit our model with more explicit modeling of the timing of births, classified tumors into four categories (ER⁺/PR⁺, ER⁺/PR⁻, ER⁻/PR⁺, and ER⁻/PR⁻), compare coefficients, and observed that the adverse effect of the first

pregnancy is only present for the ER⁻/PR⁻ tumors (16). At the same time, the protective effect of parity is strong for the ER⁺ tumors. The use of postmenopausal hormones drives up risk of both tumor types equally. Similar to Potter et al. (30), we have found that epidemiologic risk factors vary by the hormone receptor expression of the breast cancer, indicating that these receptor expression categories represent distinct stable phenotypes in human breast cancer (34) rather than a single disease with a single biological pathway.

As a society, we have maximized the interval from menarche to first birth, in part reflecting greater education and more effective contraception. This translates into a greater breast cancer burden, which will include a higher proportion of ER⁺ tumors. Thus, we must ask ourselves, in terms of prevention, what can we do given this social context? Obviously, we cannot undo the increasing duration from menarche to first pregnancy. There is no current hormonal manipulation that might mimic first pregnancy and generate the ensuing benefit.

Accordingly, we have focused our attention on the interval from menarche to first birth, to understand how exposures during this interval may offer a route for prevention. We have examined high school diet among participants in the NHS II, a cohort of younger nurses. We have evaluated risk of proliferative benign breast disease, a marker of breast cancer risk, and also breast cancer incidence. Some 29,000 women free from benign breast disease in NHS II completed a questionnaire in 1998, which had some 131 food items, asking them to recall their typical dietary intake during high school. Incident cases ($n = 470$) of proliferative benign breast disease were confirmed through centralized pathology review. We observed a positive relation between animal fat and incident proliferative disease and an inverse relation between vegetable fat and benign breast disease (35). Vitamin E and fiber are also inversely related to benign breast disease. However, contrary to our prior hypothesis (36), vitamin A, carotenoids, and folate intakes, show no relation. Components of diet might, therefore, offer useful strategies to lower risk of benign breast disease as supported by the underlying basic biology and the potential for their manipulation in the food supply.

We have also evaluated adolescent diet in relation to incidence using 361 breast cancer cases. We see that vegetable fat is significantly inversely related to risk of breast cancer (RR top quintile versus the bottom quintile, 0.58; 95% CI, 0.38-0.86; $P_{\text{trend}} = 0.005$; ref. 37). Vitamin E also showed a significant reduction in risk for the top versus bottom quintiles (Q₅ versus Q₁ multivariate RR, 0.61; 95% CI, 0.42-0.89; $P_{\text{trend}} = 0.003$). Again, vitamins A, C, D, and folate were not related to risk of breast cancer. Both of these findings give us some hope, although they need replication before we move forward with preventive action. Certainly, these findings suggest that exposure between menarche and first pregnancy may offer a real avenue for future prevention. Alternatively, a long lag between dietary exposure and breast cancer incidence may be required; a feature that is rarely evaluated in studies of diet and breast cancer.

Breast cancer risk accumulates through the life course. Selective ER modulators may offer one potential strategy for risk reduction. But, how do we identify the subgroup of the population that will get the most benefit with the fewest side effects, and what can we recommend for women to reduce breast cancer risk while we try to refine the identification of risk groups? One option is to take the modeling that Dr. Rosner and I have completed to evaluate if it can do a better job than the Gail model to identify high-risk women who would potentially have the most benefit from the selective ER modulator-type strategy. We have taken the model, developed it, and now validated, it with goodness of fit and its ability to predict risk and discriminate between

those who will get breast cancer and those who will not. Overall, the model predicts well (38). We can apply it to a separate 4 years of incident cases (1994-1998) and get a good fit to the data. We also observe a RR of ≥ 5 , comparing the top to bottom decile. This suggests we are starting to separate out high-risk from low-risk women; the statistic for discriminatory accuracy is 0.62. If we are tossing a coin to separate out our women of the same age and whether or not they will get breast cancer, we would say we have a discriminatory accuracy of 0.5. Cardiovascular prediction rules run around 0.8 for the best estimate of discriminatory accuracy (39). For breast cancer, there is room for improvement in discriminatory accuracy.

We can also compare performance to the application of the Gail model within the NHS. Rockhill has shown that the model from Gail clearly applies well in the NHS (40). It gets the number of cases correct but has a discriminatory accuracy statistic of 0.58. Identifying which women will get cancer and who will remain free from cancer is really not all that successful. We observe a RR comparing the top versus bottom decile of risk by the Gail model of only 2.8 compared with 5 in the Rosner model, which includes more risk factors (38). Of note, 44% of women who actually developed breast cancer in 5 years started with a risk of ≥ 1.67 ; the cut point that is typically used to identify women for tamoxifen. The majority of breast cancer cases in the NHS cohort occur in women who are not separated out by the recommended cut point of 1.67. The women who remain free from breast cancer and the women with the estimated 5-year risk overlap. Ideally, we would like to separate the two and be able to identify risk.

Neither the Gail nor the Rosner model includes endogenous hormone levels. Adding hormone levels should increase discrimination. Preliminary data suggest a value of about 0.68 for the discriminatory test, an increase from about 0.62 to 0.68, but our CIs are very wide because we have few cases from this nested case-control data set. Addition of breast density may further increase our predictive ability. The Harvard Specialized Programs of Research Excellence in breast cancer is facilitating such work.

The importance of hormones in breast cancer etiology and prevention is clear. We have data from the tamoxifen prevention trial as a proof of principle for chemoprevention as a feasible strategy to reduce risk, particularly for ER⁺ breast cancer (2). We must determine which agent is going to give us the risk benefit ratio that is acceptable to women, remembering that the cholesterol search took 30 years to get us to statins from cholestyramine and other early cholesterol lowering agents.

In the mean time, modifiable risk factors are amenable to change. Alcohol intake is consistently related to increased risk of breast cancer; limiting alcohol is one option. Exciting evidence from the cohort (41) shows that among women with high plasma folate levels, there is no adverse effect of alcohol. Thus, increasing plasma folate levels through the use of multivitamins or fortification of the food supply may prevent breast cancer, particularly for women who are consuming alcohol.

If we think about how we present risk, there are a number of questions we might ask that point to whether we are doing a service or a disservice in the way we present the results of epidemiologic studies. The disservice includes undue anxiety about breast cancer risk. The study of breast cancer risk by Black et al. at Dartmouth College (42), observing that the most common response to what "1 in 10 risk of breast cancer" means is that on average one woman in 10 will die in the next 10 years. One in five was their mean estimate of how many women will be diagnosed with breast cancer in 10 years; again, grossly misestimating risk! Public health professionals and the media have combined to give misleading information

to women. We need to move beyond talking about lifetime risk, because lifetime for any one person is hard to estimate. Accordingly, in our own work, we have shifted to using age 70, which women may at least be able to relate to more easily than their unknown lifetime risk.

Returning to preventability, I pointed to international variation as strong evidence supporting this. I have shown some of the variation that reproductive characteristics can contribute as well as growth, obesity, and postmenopausal hormones. We can add to this data on alcohol and folate. Our retrospective findings for vitamin E and vegetable fat intake during adolescence are provocative, suggesting up to 50% reduction in risk (35, 37, 43). Replication of this work, together with prospective confirmation, must be a high priority before we can frame prevention strategies.

We have modeled breast cancer incidence, focusing on risk accumulation. This points most strongly to the interval from menarche to first birth, an interval where we have less epidemiologic data than in the rest of life in relation to breast cancer risk. Globally, industrialization will drive breast cancer risk as the interval lengthens with development. We also have a range of social pressures to complete education and establish careers before childbirth. These forces exacerbate the population risk of breast cancer.

Prevention should be considered in the context of levels of intervention that we need in place if we are to succeed at any prevention strategy, be it tobacco, lowering the risk of breast cancer, or any other preventive endeavor. We must identify what we can do through health care providers, through regulatory approaches, and at the community level (44). In terms of the risk of breast cancer, it is clear that health care providers can already be counseling on diet, physical activity, and avoidance of weight gain, to lower risk of breast cancer and other chronic diseases. Ideally, we will have strategies to identify high-risk women and focus chemopreventive interventions on them. Facilitating lactation (45), physical activity (46), and fortification of the food supply to raise folate levels are examples of how regulatory approaches can lower breast cancer risk. At the community level, we need a range of changes that would foster and support physical activity and ultimately access to medical care.

We have much work to do to translate our findings from epidemiologic and clinical studies to successful prevention of breast cancer. If we can apply what we already know, there is real potential to lower the burden of breast cancer in our society today.

Acknowledgments

I thank the contributions of the participants in the NHS and the efforts of many faculty and long-term staff of the NHS research group and Dr. Rosner for providing collaborative input to many of the analyses summarized in this article.

References

- Paget J. Address in surgery. *Lancet* 1862;80:158–64.
- Fisher B, Costantino J, Wickerham D, et al., for the National Surgical Adjuvant Breast and Bowel Project Investigators. Tamoxifen for prevention of breast cancer: report of the National Surgical Adjuvant Breast and Bowel Project P-1 study. *J Natl Cancer Inst* 1998;90:1371–88.
- Cummings SR, Eckert S, Krueger KA, et al. The effect of raloxifene on risk of breast cancer in postmenopausal women. *JAMA* 1999;281:2189–97.
- Parkin D, Muir C, Whelan S, Gao Y-T, Ferlay J, Powell J. Cancer incidence in five continents. Vol. 5. IARC scientific publication no. 120. Lyon: IARC; 1992.
- Kolonel L. Cancer patterns of four ethnic groups in Hawaii. *J Natl Cancer Inst* 1980;65:1127–39.
- Nagata C, Kawakami N, Shimizu H. Trends in the incidence rate and risk factors for breast cancer in Japan. *Breast Cancer Res Treat* 1997;44:75–82.
- Seow A, Duffy S, McGee M, Lee J, Lee H. Breast cancer in Singapore: trends in incidence 1968–1992. *Int J Epidemiol* 1996;25:40–5.
- Jin F, Shu X-O, Devesa S, Zheng W, Blot W, Gao Y-T. Incidence trends for cancers of the breast, ovary, and corpus uteri in urban Shanghai, 1972–1989. *Cancer Causes Control* 1993;4:355–60.
- Miyaji N, Lock M. Monitoring motherhood: sociocultural and historical aspects of maternal and child health in Japan. *Daedalus* 1994;134:87–112.
- Moolgavkar SH, Day NE, Stevens RG. Two-stage model for carcinogenesis: epidemiology of breast cancer in females. *J Natl Cancer Inst* 1980;65:559–69.
- Rosner B, Colditz GA, Willett WC. Reproductive risk factors in a prospective study of breast cancer: the Nurses' Health Study. *Am J Epidemiol* 1994;139:819–35.
- Colditz G, Rosner B. Cumulative risk of breast cancer to age 70 years according to risk factor status: data from the Nurses' Health Study. *Am J Epidemiol* 2000;152:950–64.
- Rosner B, Colditz G. Log-incidence mathematical model of breast cancer incidence. *J Natl Cancer Inst* 1996;88:359–64.
- Armitage P, Doll R. The age distribution of cancer and a multistage theory of carcinogenesis. *Br J Cancer* 1954;8:1–12.
- Pike MC, Krailo MD, Henderson BE, Casagrande JT, Hoel DG. "Hormonal" risk factors, "breast tissue age" and the age-incidence of breast cancer. *Nature* 1983;303:767–70.
- Colditz G, Rosner B, Chen WY, Holmes M, Hankinson SE. Risk factors for breast cancer: according to estrogen and progesterone receptor status. *J Natl Cancer Inst* 2004;96:218–28.
- van Weiringen J. Secular growth changes. In: Falkner F, Tanner JM, editors. *Human growth: a comprehensive treatise*. New York: Plenum Press; 1986.
- Chen J, Campbell TC, Junyao L, Peto R. Diet, life-style, and mortality in China: a study of the characteristics of 65 Chinese counties. Oxford (England): Oxford University Press; 1990.
- The Endogenous Hormones and Breast Cancer Collaborative Group. Endogenous sex hormones and breast cancer in postmenopausal women: reanalysis of nine prospective studies. *J Natl Cancer Inst* 2002;94:606–16.
- Gail MH, Costantino JP, Bryant J, et al. Weighing the risks and benefits of tamoxifen treatment for preventing breast cancer. *J Natl Cancer Inst* 1999;91:1829–46.
- Early breast cancer trialists' collaborative, g. Effects of adjuvant tamoxifen and of cytotoxic therapy on mortality in early breast cancer. An overview of 61 randomized trials among 28,896 women. *N Engl J Med* 1988;319:1681–92.
- Cummings SR, Duong T, Kenyon E, Cauley JA, Whitehead M, Krueger KA. Serum estradiol level and risk of breast cancer during treatment with raloxifene. *JAMA* 2002;287:216–20.
- Beral V. Breast cancer and hormone-replacement therapy in the Million Women Study. *Lancet* 2003;362:419–27.
- Rossouw JE, Anderson GL, Prentice RL, et al. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results from the Women's Health Initiative randomized controlled trial. *JAMA* 2002;288:321–33.
- Fletcher SW, Colditz GA. Failure of estrogen plus progestin therapy for prevention. *JAMA* 2002;288:366–8.
- Huang Z, Hankinson SE, Colditz GA, et al. Dual effects of weight and weight gain on breast cancer risk. *J Am Med Assoc* 1997;278:1407–11.
- Yasui Y, Potter J. The shape of the age-incidence curves of female breast cancer by hormone-receptor status. *Cancer Causes Control* 1999;10:431–7.
- Hildreth NG, Kelsey JL, Eisenfeld AJ, LiVolsi VA, Holford TR, Fischer DB. Differences in breast cancer risk factors according to the estrogen receptor level of the tumor. *J Natl Cancer Inst* 1983;70:1027–31.
- Hislop TG, Coldman AJ, Elwood JM, Skippen DH, Kan L. Relationship between risk factors for breast cancer and hormonal status. *Int J Epidemiol* 1986;15:469–76.
- Potter JD, Cerhan JR, Sellers TA, et al. Progesterone and estrogen receptors and mammary neoplasia in the Iowa Women's Health Study: how many kinds of breast cancer are there? *Cancer Epidemiol Biomarkers Prev* 1995;4:319–26.
- Cooper JA, Rohan TE, Cant EL, Horsfall DJ, Tilley WD. Risk factors for breast cancer by oestrogen receptor status: a population-based case-control study. *Br J Cancer* 1989;59:119–25.
- Stanford JL, Szklo M, Boring CC, et al. A case-control study of breast cancer stratified by estrogen receptor status. *Am J Epidemiol* 1987;125:184–94.
- Kreiger N, King WD, Rosenberg L, Clarke EA, Palmer JR, Shapiro S. Steroid receptor status and the epidemiology of breast cancer. *Ann Epidemiol* 1991;1:513–23.
- Robertson J. Oestrogen receptor: a stable phenotype in breast cancer. *Br J Cancer* 1996;73:5–12.
- Baer HJ, Schnitt SJ, Connolly JL, et al. Adolescent diet and incidence of proliferative benign breast disease. *Cancer Epidemiol Biomarkers Prev* 2003;12:1159–67.
- Colditz G, Frazier A. Models of breast cancer show that risk is set by events of early life: prevention efforts must shift focus. *Cancer Epidemiol Biomarkers Prev* 1995;4:567–71.
- Frazier AL, Li L, Cho E, Willett WC, GA C. Adolescent Diet and Risk of Breast Cancer (United States). *Cancer Causes Control* 2004;15:73–82.
- Rockhill B, Byrne C, Rosner B, Louie MM, Colditz G. Breast cancer risk prediction with a log-incidence model: evaluation of accuracy. *J Clin Epidemiol* 2003;56:856–61.
- D'Agostino RB Sr, Grundy S, Sullivan LM, Wilson P. Validation of the

- Framingham coronary heart disease prediction scores: results of a multiple ethnic groups investigation. *JAMA* 2001;286:180–7.
40. Rockhill B, Spiegelman D, Byrne C, Hunter DJ, Colditz GA. Validation of the Gail et al. model of breast cancer risk prediction and implications for chemoprevention. *J Natl Cancer Inst* 2001;93:358–66.
 41. Zhang SM, Willett WC, Selhub J, et al. Plasma folate, vitamin B6, vitamin B12, homocysteine, and risk of breast cancer. *J Natl Cancer Inst* 2003;95:373–80.
 42. Black WC, Nease RF Jr, Tosteson AN. Perceptions of breast cancer risk and screening effectiveness in women younger than 50 years of age. *J Natl Cancer Inst* 1995;87:720–31.
 43. Frazier AL, Ryan CT, Rockett H, Willett WC, Colditz GA. Adolescent diet and risk of breast cancer. *Breast Cancer Res* 2003;5:R59–64.
 44. Atwood K, Colditz G, Kawachi I. Implementing prevention policies: relevance of the Richmond model to health policy judgments. *Am J Public Health* 1997;87:1603–6.
 45. Beral V. Breast cancer and breastfeeding: collaborative reanalysis of individual data from 47 epidemiological studies in 30 countries, including 50302 women with breast cancer and 96973 women without the disease. *The Lancet* 2002;360:187–95.
 46. IARC. Weight control and physical activity. IARC handbook on cancer prevention. Lyon: IARC; 2002.