VOLUME 22 · NUMBER 1 · JANUARY 1 2004

JOURNAL OF CLINICAL ONCOLOGY

ORIGINAL REPORT

Menopausal Status Dependence of Early Mortality Reduction Due to Diagnosis of Smaller Breast Cancers (T1 *v* T2-T3): Relevance to Screening

Romano Demicheli, Gianni Bonadonna, William J.M. Hrushesky, Michael W. Retsky, and Pinuccia Valagussa

From the Istituto Nazionale Tumori, Milano, Italy; Department of Surgery, Children's Hospital and Harvard Medical School, Boston, MA; and Dorn VA Medical Center, Columbia, SC.

Submitted December 26, 2002; accepted October 30, 2003.

Authors' disclosures of potential conflicts of interest are found at the end of this article.

Address reprint requests to Romano Demicheli MD, PhD, Istituto Nazionale Tumori, Department of Medical Oncology, Via Venezian 1, 20137 Milano, Italy; e-mail: demicheli@istitutotumori.mi.it.

0732-183X/04/2201-102/$20.00

DOI: 10.1200/JCO.2004.12.139

A B S T R A C T

Purpose
To provide data relevant to the paradoxical mortality excess for women age 40 to 49 years observed during the first 6 to 8 years in the invited group in all mammography screening studies.

Patients and Methods
In 1,173 patients undergoing mastectomy alone as primary treatment, allocated to subsets according to menopausal status and tumor size, hazard rates for death were calculated. The ratios between the hazard rate for T2-T3 patients and the corresponding value for T1 patients were assessed over time.

Results
For postmenopausal patients, the ratio appeared to be time-dependent, dropping from the maximum value of approximately 5 at the first year after surgery to a near constant value of approximately 2 after 5 to 6 years. Premenopausal patients, on the contrary, showed a nearly constant ratio of approximately 3. Therefore, although in each T-category the 10-year survival of premenopausal and postmenopausal patients was similar, its time distribution was menopause-dependent. In particular, the difference between cumulative survival of premenopausal and postmenopausal T2-T3 patients attained statistical significance after 3 years.

Conclusion
The mortality reduction due to the diagnosis of smaller tumors is significantly higher for postmenopausal women than for premenopausal women during early postsurgery time. According to the hypothesis that primary tumor surgical removal, occurring sooner in the invited group than in the control arm of screening trials, results in some acceleration of metastasis development, a greater number of unfavorable events (recurrence and death) should occur in the invited group. We suggest that for younger women, the early balance between benefit from tumor downsizing and harm from surgery-induced metastasis acceleration results in harm. This disadvantage does not occur in postmenopausal women.

J Clin Oncol 22:102-107. © 2004 by American Society of Clinical Oncology

INTRODUCTION

Mammography screening was first studied in a large randomized controlled trial (RCT) in New York (the Health Insurance Plan of Greater New York trial) in the 1960s [1] and was further assessed in other RCTs (Malmo, Two-County, Stockholm, and Goteborg) in Sweden in the 1970s and 1980s. The Swedish trials have been recently reviewed by an overview committee that confirmed fundamentally the results previously reported by the individual research groups [2]. Even the results of a trial in the United Kingdom (the Edinburgh RCT) were quite similar [3].

Mammography screening apparently reduces breast cancer mortality of the invited women by approximately 20% to 30% in comparison with controls. This benefit, however, occurs primarily in women older than 50 to 55 years. For instance, the Swedish overview, which summarizes a significant fraction of performed RCTs, found a 15% mortality reduction in the age group 40 to 54 years and a 25% reduction for the age

group 55 to 74 years. It is noteworthy but seldom clearly stated that all studies demonstrated a breast-cancer mortality excess in the screened group of younger women during the first 6 to 8 years, whereas advantage, if any occurred, developed later.

The Health Insurance Plan, Swedish, and Edinburgh RCTs had been designed primarily for women age 40 to 65 or 75 years, and findings for women ages 40 to 49 resulted from subset analysis of the studied population. The only RCT specifically designed to evaluate the benefit of mammography (plus clinical breast exam) screening for women ages 40 to 49 years was carried out in Canada. Despite the specificity of the 25,000 subjects per arm, the expectation that it would show screening benefit in comparison with usual community care even for this age group was frustrated. In particular, the early excess of deaths among screened women remained present at more than 10 years of follow-up [4], thus replicating findings of other RCTs.

Following the National Institutes of Health Consensus Development Conference Breast Cancer Screening for Women Ages 40 to 49, two different and contradicting reports were published [5]. A consensus panel voted that data do not support a universal recommendation of screening for all women age 40 to 49 years, whereas a minority report came to the opposite conclusion. The resultant controversy became even more complicated when a recent paper raised doubts about the value of mammography screening for women of all ages [6] and earned harsh criticism (and some partial support) [7,8,9]. It is surprising that, even during this heated controversy, no attention was paid to the paradoxical breast-cancer mortality surge for younger women invited to undergo screening.

We report observations, based on our study of the survival pattern of 1,173 patients undergoing mastectomy alone as primary treatment for operable breast cancer, that may clarify some age-related differences in screening effectiveness. This analysis demonstrates that the mortality increase that is associated with the diagnosis of larger tumors (larger than 2 cm *v* smaller than 2 cm) depends on the menopausal state at discovery and the time following surgery. Since screening is believed to result in the improvement of breast cancer outcome by detecting significantly smaller tumors in the intervention arm, the screening-induced "improvement" of prognosis may be age dependent (via the menopausal status) and change as follow-up progresses. We suggest a biologic mechanism underlies the differences in breast cancer behavior between younger and older women.

PATIENTS AND METHODS

All patients entered onto three different clinical trials between 1964 through 1980 at the Milan Cancer Institute, with mastectomy alone as primary treatment for operable breast cancer, were retrospectively evaluated for this analysis. Before surgery, all patients underwent standard staging: complete physical examination, x-ray study of chest, skull, spine, and pelvis, bilateral mammography, electrocardiogram, complete hemogram, and routine biochemical tests. Primary tumor was treated by radical or modified radical mastectomy. No patient received postoperative radiotherapy or chemotherapy. Patients experiencing relapse were given systemic treatment according to the standards used at the time of recurrence. More detailed characteristics of the studied series have been reported [10].

Menopausal status was defined as "postmenopausal" if 1 year was elapsed since the last menstrual period. The date of surgery was assumed as the start point for survival determination, and the death date was obtained by the death certificate. Because the cause of death was not accurately identified for all patients, both deaths from all causes and deaths following documented recurrence from breast cancer were examined. The former may be considered a reasonable estimate of breast cancer related deaths. Indeed, in two more recent clinical trials carried out at the Milan Cancer Institute, involving more than 900 patients and with careful assessment of the cause of death, 98.8% of deaths following recurrence were attributed to breast cancer.

The death hazard rate was calculated as the conditional probability of dying in a time interval, given that the patient was alive at the beginning of the interval. The yearly discrete hazard rates for death and the survival curves were calculated by means of the life-table method over the course of 10 years. Because the hazard rates can fluctuate due to random variation, a Kernel-like smoothing procedure [11] was adopted, and a window width of 3 years was empirically chosen. Cumulative survival curves were compared by the two-sided log-rank test.

RESULTS

A total of 1,173 patients were examined. Median age was 52 years (range, 23 to 82 years), 516 patients were premenopause (T1, 222; T2-T3, 294), and 657 patients were postmenopause (T1, 237; T2-T3, 420). A total of 588 patients experienced disease relapse within 15 years from surgery, and 625 patients died, whereas 554 and 361 patients were surviving at risk at 10 and 15 years, respectively.

Patients were allocated to four subsets according to menopausal status (pre- *v* post-) and tumor size (T1 *v* T2-T3), and the corresponding hazard rates were calculated for both deaths from all causes and deaths after breast cancer recurrence (Fig 1). The analysis was limited to 10 years following surgery in order to have at least 150 patients at risk in each group.

Findings from both death analyses were very similar. Premenopausal patients yielded hazard rate curves displaying a near parallel pattern for T1 and T2-T3 subsets, with an initial mortality wave, peaking at four to five years after surgery, followed by a second mortality peak at the eighth year. Somewhat differently, postmenopausal patients showed an early major mortality surge peaking at the third year for T2-T3 tumors, dropping afterward to a very modest increase at the eighth year, whereas the

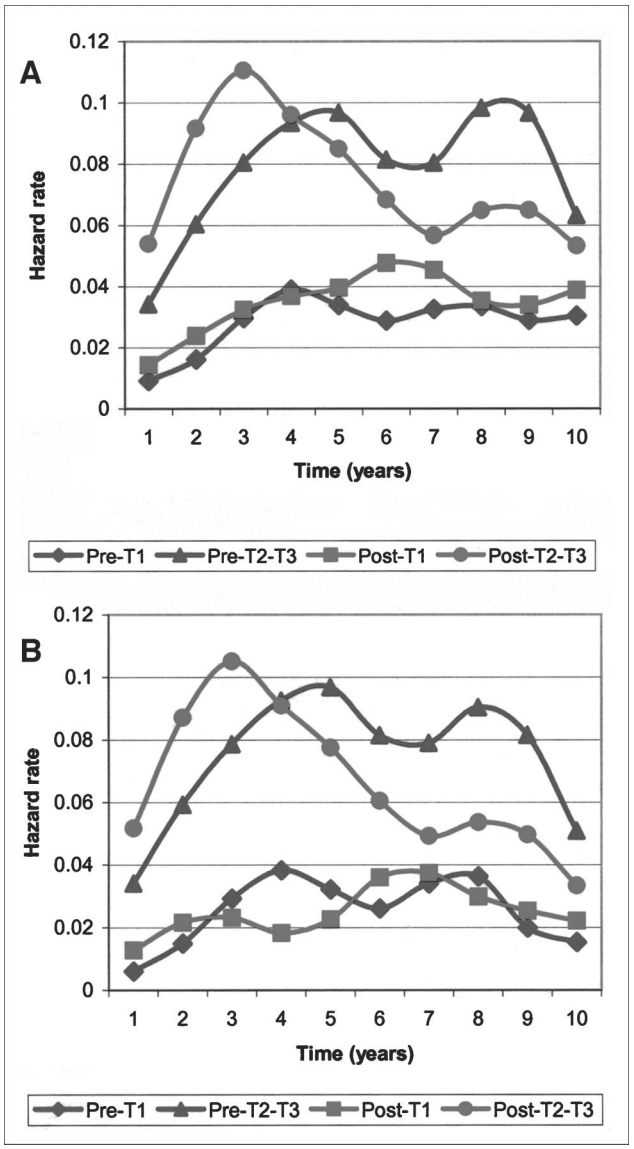

Fig 1. Yearly death risk: time distribution of the hazard rates for deaths from all causes (A) and for deaths following recurrence (B) in 1,173 patients undergoing mastectomy without any adjuvant treatment. Pre, premenopausal; Post, postmenopausal.

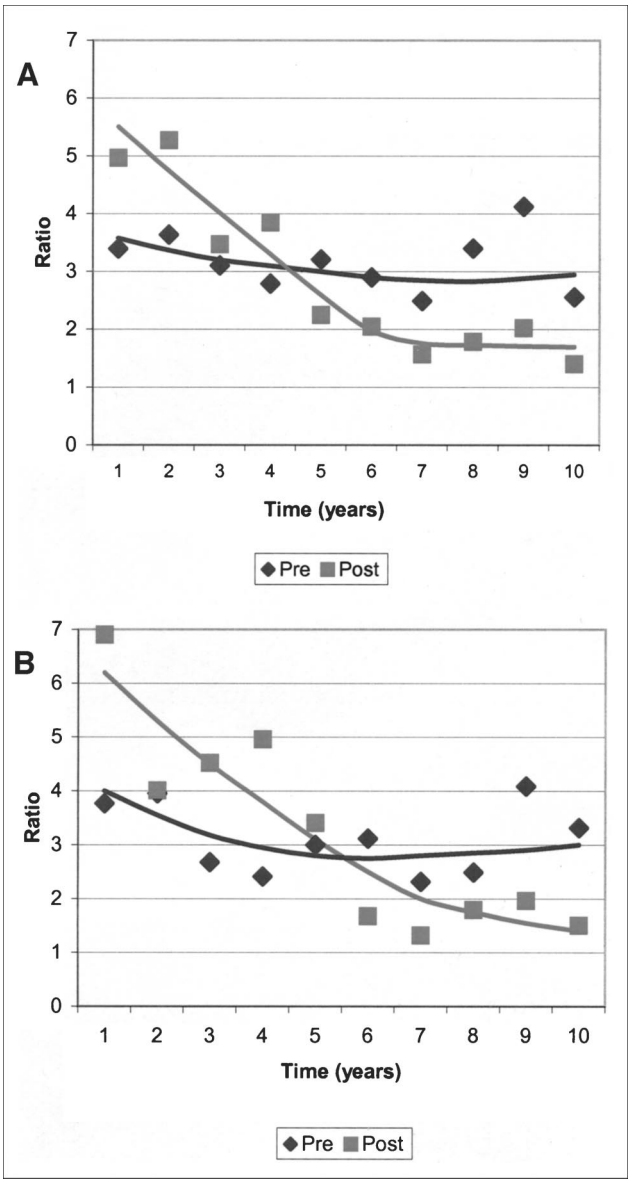

Fig 2. Death risk ratio between the values of the hazard rate for deaths from all causes (A) and for deaths following recurrence (B) for T2-T3 patients and the corresponding values for T1 patients. Pre, premenopausal; Post, postmenopausal.

hazard curve of patients with T1 tumors did not parallel this pattern.

For both menopausal subsets, the ratios between the hazard rate for T2-T3 patients and the corresponding value for T1 patients were calculated (Fig 2). Results provided by the analysis of deaths from all causes and deaths following breast cancer recurrence were very similar. In the subset of postmenopausal patients, the ratio appeared to be definitely time-dependent, and moving from T1 to T2-T3 category resulted in considerable worsening of early mortality (approximately five-fold during the first 2 years), whereas the detrimental effect lowered afterward approaching a constant value of approximately two-fold after approximately 5 to 6 years. Premenopausal patients, quite differently, showed a nearly constant ratio (approximately 3) during the whole investigated time, with the possible exception of the first 1 to 2 years. It is noteworthy that the curves crossed each other at 4 to 5 years following surgery, when postmenopausal patients began to display a less unfavorable ratio in comparison with premenopausal patients.

The above reported findings were confirmed by a careful reading of the classical cumulative survival curves (Fig 3), in which one can observe that, even if the 10-year survival of premenopausal and postmenopausal patients was similar both in T1 and T2-T3 categories, in the latter group,

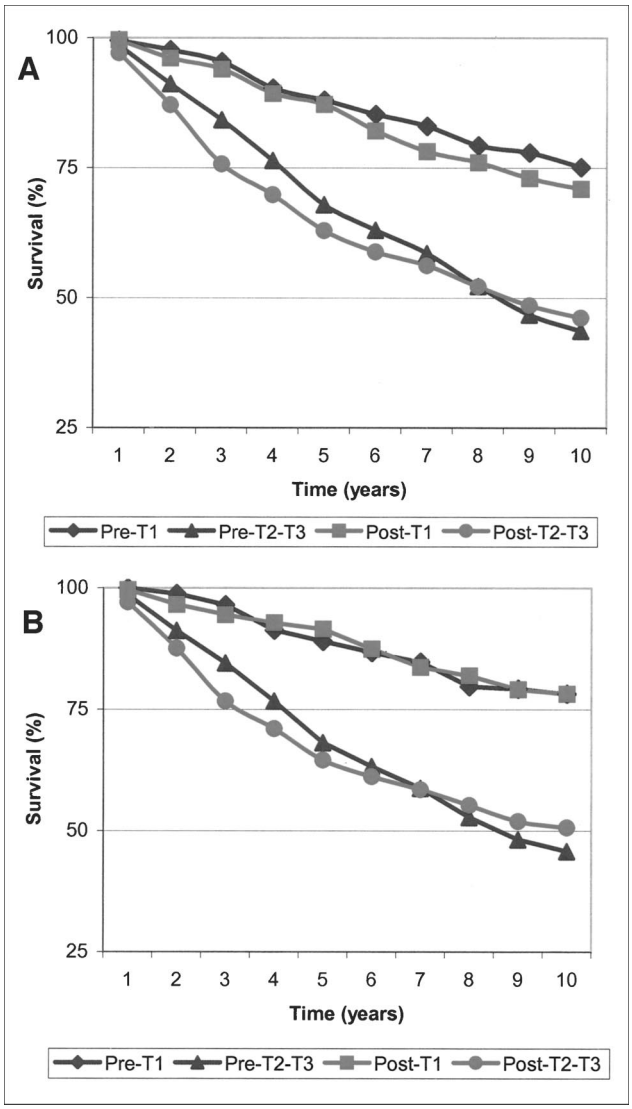

Fig 3. Cumulative survival: survival curves for 1,173 patients undergoing mastectomy without any adjuvant treatment. Deaths from all causes (A) and deaths following recurrence (B) were separately considered. Patients were classified according to menopausal status and tumor size. Pre, premenopausal; Post, postmenopausal.

the time-distribution of deaths was menopause-dependent. In particular, it may be shown that for the T2-T3 subset, the difference between cumulative survival of premenopausal and postmenopausal patients attained statistical significance after 3 years ($P < .025$; log-rank test), whereas it lost its significance afterwards.

DISCUSSION

The utility of screening depends on the hypothesis that, in the clinical context, breast cancer behaves in part as a linearly progressive disease. Clinically evident metastases will only develop in a fraction of patients. This fraction increases with the tumor size at diagnosis. According to this concept, mammography screening, resulting in smaller tumor detection, should significantly and invariably improve the outcome of the disease.

Mammography demonstrated a significant reduction of tumor diameter at discovery for the intervention arm in every trial that has been done. For example, in the Two-County RCT [12], the mean size of invasive tumors was 13.5 mm for women invited to screening and 22.8 mm in controls. The screening effect on tumor diameter was detectable in both women ages 40 to 49 years (tumor diameter range, 21.1 mm to 14.1 mm) and in older patients (tumor diameter range, 23.1 mm to 13.4 mm). Therefore, the invitation to mammography screening resulted in a shifting of the "average patient" with operable invasive breast cancer from T2-T3 to T1 category. This forces the conclusion that, during the trial time, some unscreened women, who had undetected T1 lesions at the same time of a screening detected counterpart, had progression of their cancers from T1 size to T2-T3.

In this study we found that when T-stage changed from T1 to T2-T3 category, the resulting mortality increase showed significantly different patterns across follow-up time in premenopausal compared with postmenopausal patients. It should be noted that screening RCTs have demonstrated different effectiveness for women younger than 50 to 55 years, compared with older women. However, given that menopausal status is the most important biologic factor differentiating the two age groups, the use of menopause-related instead of age-related survival data is biologically well justified.

The hypothesis that the different temporal dynamics of mortality in pre- and postmenopausal patients may have been due to causes other than breast cancer can be reasonably ruled out, because the results were identical whether deaths from all causes or deaths following breast cancer recurrence were analyzed. Therefore, we conclude that these different mortality patterns depend on menopausal status as a trait of breast cancer natural history.

The prevention of mortality increase by tumor progression to more advanced stage becomes a survival gain, from a screening study perspective, because of tumor downstaging that results from earlier diagnosis in the screened group. The different temporal patterns of death associated to menopausal status that we found in this study, however, imply that during early follow-up, the reduction of mortality from tumor downsizing is less important for younger women than for older women, even if the two age groups end up with similar late results. Therefore, diminished survival gain due to screening should be expected in women aged 40 to 49 during the first years after screening is begun.

This is consistent with the low benefit from screening for younger women. Data from both individual RCTs and

meta-analyses, however, showed a paradoxical phenomenon, with actual inversion of mortality rates between the two study arms during the first few years of follow-up, favoring those women *not* offered mammographic screening. The disturbing early mortality inversion is quite evident in the Cox meta-analysis [13], where it reached a statistically significant ratio of 2.4 (95% CI, 1.1% to 5.4%) in favor of the control arm at the third year of follow-up. The delayed trend toward screening benefit is often discussed (eg, in the Kerlikowske study [14]), but the mortality inversion, although consistent through all RCTs, is usually ignored.

Besides the above mentioned downsizing of tumors detected among the invited women, the timing of surgical intervention was the second most remarkable difference between screened and control arms. Because virtually all women diagnosed with breast cancer underwent tumor removal at diagnosis, the surge of breast cancer diagnoses in the screened group during early follow-up resulted in many more operations in women in the screening arm, in comparison with the control arm. This phenomenon is inherent in the screening study design, and we assume it is relevant to the mortality paradox for premenopausal women.

The perturbation of tumor growth kinetics as a result of surgery has been well documented in animal models since the 1950s and1960s, and in particular, it was systematically studied for Lewis-lung carcinoma [15]. Noncurative excision of primary tumor and sham surgery each resulted in an increase of the proliferation index and growth rate of lung metastases, accompanied by a low but consistent decrease in median life span of animals. These findings were further confirmed [16-18] across a variety of tumor types, suggesting that the surgery-related proliferative impulse is likely to be a common phenomenon. A growth-stimulating factor was found in serum of animals given surgery [19] and more recently, direct study of micrometastases proved that primary tumor removal caused in some experimental tumors a switch of micrometastatic foci to the angiogenic phenotype, resulting in growth of metastases, by withdrawing angiogenesis inhibitor factors that maintained distant micrometastases in an avascular steady phase [20].

The effect of primary tumor surgical removal has been considered to cause the double-peaked pattern of the recurrence risk of patients undergoing tumor removal with [21] or without [10] adjuvant chemotherapy (cyclophosphamide, methotrexate, and flourouracil), and of the mortality risk of the latter patients [22]. This phenomenon was the cornerstone for a proposed biologic model for breast cancer metastasis development, incorporating tumor dormancy in specific micrometastatic phases and stochastic transitions between them [23]. A computer simulation of the model generated double peaked relapse histograms reasonably similar to clinical data [24]. According to the model, the early peak of the hazard function for local and distant recurrences in resected breast cancer patients is generated by the sudden acceleration of the metastatic process by means of growth factors and angiogenesis triggering mechanisms due to surgery. We have recently hypothesized that surgery-induced angiogenesis may be a main factor for early recurrence in premenopausal node-positive patients [25].

Taking into account all previous considerations, we suggest a possible explanation for the paradoxical harmful effect of screening on women ages 40 to 49 years. As a result of screening, breast cancers are found at an earlier stage than would be found without screening, but then surgical intervention to remove the primary tumor accelerates metastatic growth, which partially offsets the early detection advantage for a significant subset of patients. In other words, screening and control arms have different surgery timing distributions resulting in a early recruitment of unfavorable events for the screening arm only, which overwhelms, for younger women, the relatively modest early gain from tumor downsizing.

Why was not this effect observed for older women? According to our explanation, the screening effect is the resultant of two opposed trends, one of which (or perhaps both) is time- and age- (menopausal status) dependent. Therefore, it is quite reasonable to propose that for older women there is a greater benefit from early detection and resultant tumor downsizing than harm from unfavorable event synchronization by early surgery. Additionally, the absence of menstrual cycle phases may mean that postmenopausal women are not at risk of being resected at particularly inopportune hormonal milieus.

In conclusion, we suggest the following picture. Primary breast cancer resection results in some acceleration of metastasis development at any age, thus worsening the natural history of the disease in some patients [26]. Mammography screening results in the earlier diagnosis of breast cancer in the invited group, mainly at the first screening round (prevalence screening). This results in earlier surgery-induced recruitment of unfavorable events (recurrence and death) in comparison with the control arm. This harmful trend is modulated by the mortality reduction because of earlier diagnosis, which has different weight in premenopausal and postmenopausal women, during the first few years following surgery. The dynamics of this process results in favorable balance only in the postmenopausal group.

There are other factors that may contribute to the poor results of mammography screening for women younger than 50 years. The reduction in size observed in screen-detected cancers compared with clinically diagnosed cancers may be accompanied by fewer axillary node metastases in women 50 years and older and not in younger women. This difference further suggests different tumor biology (progressive *v* early systemic behavior) in premenopausal and postmenopausal women [27].

Mammography has a poorer signal-to-noise ratio in young women, [28] and this ratio may depend on the menstrual cycle phase during which it is performed [29]. Optimal interscreening interval is also probably age dependent and may be shorter for younger women [12]. It has been shown that, when surgical breast cancer resection occurs in the early luteal phase of the menstrual cycle of premenopausal women, outcomes are better than when the operation is performed within other phases [30]. All these factors, and perhaps others, may have "lowered" the screening effectiveness in the RCTs, but they cannot explain the "increase" of mortality in younger women invited to screening.

In our opinion, the suggested surgery-induced interruption of tumor dormancy in the metastatic foci is the main biologic mechanism underlying the mortality paradox of screening RCTs.

Authors' Disclosures of Potential Conflicts of Interest

The authors indicated no potential conflicts of interest.

REFERENCES

1. Shapiro S, Strax P, Venet L: Periodic breast screening in reducing mortality from breast cancer. JAMA 215:1777-1785, 1971

2. Nystrom L, Andersson Bjurstam N, et al: Long-term effects of mammography screening: Updated overview of the Swedish randomised trials. Lancet 359:909-919, 2002

3. Alexander FE, Anderson TJ, Brown HK, et al.: 14 years of follow-up from the Edimburgh randomised trial of breast cancer screening. Lancet 353:1903-1908, 1999

4. Miller AB, To T, Baines CJ, et al: The Canadian National Breast Screening Study: Update on breast cancer mortality. J Natl Cancer Inst Monogr 22:37-41, 1997

5. National Institutes of Health Consensus Development Conference Statement: Breast Cancer Screening for Women Ages 40-49, January 21-23, 1997—National Institutes of Health Consensus Development Panel. J Natl Cancer Inst Monogr 22:7-18, 1997

6. Gotzsche PC, Olsen O: Is screening for breast cancer with mammography justifiable? Lancet 355:129-134, 2000

7. de Koning HJ: Assesment of nationwide cancer-screening programmes. Lancet 355:80-81, 2000

8. Rozenberg S, Liebens F, Ham H: Screening mammography re-evaluated. Lancet 355:751-752, 2000

9. Duffy SW: Interpretation of the breast screening trials: A commentary on the recent paper by Gotzsche and Olsen. Breast J 10:209-212, 2001

10. Demicheli R, Abbattista A, Miceli R, et al: Time distribution of the recurrence risk for breast cancer patients undergoing mastectomy: Further support about the concept of tumor dormancy. Breast Cancer Res Treat 41:177-185, 1996

11. Ramlau-Hansen H: Smoothing counting process intensities by means of Kernel functions. Ann Stat 11:453-466, 1983

12. Tabar L, Fagerberg G, Chen HH, et al: Efficacy of breast cancer screening by age. Cancer 75:2507-2517, 1995

13. Cox B: Variation in the effectiveness of breast screening by year of follow-up. J Natl Cancer Inst Monogr 22:69-72, 1997

14. Kerlikowske K: Efficacy of screening mammography among women aged 40 to 49 years and 50 to 69 years: Comparison of relative and absolute benefit. J Natl Cancer Inst Monogr 22:79-86, 1997

15. Simpson-Herren L, Sanford AH, Holmquist JP: Effects of surgery on the cell kinetics of residual tumor. Cancer Treat Rep 60:1749-1760, 1976

16. Braunschweiger PG, Schiffer LM, Betancourt S: Tumor cell proliferation and sequential chemotherapy after partial tumor resection in C3H/HeJ mammary tumours. Breast Cancer Res Treat 2:323-329, 1982

17. Gunduz N, Fisher B, Saffer EA: Effect of surgical removal on the growth and kinetics of residual tumor. Cancer Res 39:3861-3865, 1979

18. Bogden AE, Moreau JP, Eden PA: Proliferative response of human and animal tumours to surgical wounding of normal tissues: Onset, duration and inibition. Br J Cancer 75:1021-1027, 1997

19. Fisher B, Saffer EA, Rudok C, et al: Presence of a growth-stimulating factor in serum following primary tumor removal in mice. Cancer Res 49:1996-2001, 1989

20. Holmgren L, O'Reilly MS, Folkman J: Dormancy of micrometastases: Balanced proliferation and apoptosis in the presence of angiogenesis suppression. Nature Med 1:149-153, 1995

21. Demicheli R, Miceli R, Brambilla C, et al: Comparative analysis of breast cancer recurrence risk for patients receiving adjuvant Cyclophosphamide, Methotrexate, Fluorouracil (CMF): Data supporting the occurrence of "cures". Breast Cancer Res Treat 53:209-215, 1999

22. Demicheli R, Valagussa P, Bonadonna G: Double-peaked time distribution of mortality for breast cancer patients undergoing mastectomy. Breast Cancer Res Treat 75:127-134, 2002

23. Demicheli R, Retsky MW, Swartzendruber DE, et al: Proposal for a new model of breast cancer metastatic development. Ann Oncol 8:1075-1080, 1997

24. Retsky MW, Demicheli R, Swartzendruber DE, et al: Computer simulation of a breast cancer metastasis model. Breast Cancer Res Treat 45: 193-202, 1997

25. Retsky M, Demicheli R, Hrushesky W: Premenopausal status accelerates relapse in node positive breast cancer: Hypothesis links angiogenesis, screening controversy. Breast Cancer Res Treat 65:217-224, 2001

26. Demicheli R, Valagussa P, Bonadonna G: Does surgery modify growth kinetics of breast cancer micrometastases?. Br J Cancer 85:490-492, 2001

27. Peer PG, Holland R, Hendriks JH, et al: Age-specific effectiveness of the Nijmegen population-based breast cancer-screening program: Assessment of early indicators of screening effectiveness. J Natl Cancer Inst 86:436-441, 1994

28. Kerlikowske K, Grady D, Barclay J, et al: Effect of age, breast density and family history on the sensitivity of first screening mammography. JAMA 276:33-38, 1996

29. Baines CJ, Vidmar M, McKeown-Eyssen G, et al: Impact of menstrual phase on false-negative mammograms in the Canadian National Breast Screening Study. Cancer 80:720-724, 1997

30. Hrushesky WJ, Bluming AZ, Gruber SA, et al: Menstrual influence on surgical cure of breast cancer. Lancet 2:949-952, 1989