Breast Cancer Adjuvant Therapy: Time to Consider Its Time-Dependent Effects

Ismail Jatoi, *University of Texas Health Science Center, San Antonio, TX*
 William F. Anderson, *National Cancer Institute, Bethesda, MD*
 Jong-Hyeon Jeong and Carol K. Redmond, *University of Pittsburgh, Pittsburgh, PA*

Breast cancer is a chronic and heterogeneous disease that may recur many years after initial diagnosis and treatment.¹ This has important implications for the practicing oncologist. For instance, an early effect of adjuvant treatment may diminish over time after cessation of therapy, or, alternatively, there may exist a lag time before some treatment effects become pronounced. Indeed, the risk of breast cancer recurrence and death (hazard rate) varies over time (ie, is nonproportional) according to prognostic and predictive factors (Figs 1 and 2; Table 1).^{6,13} The hazard curve for breast cancer death peaks between 2 and 3 years after initial diagnosis and then declines sharply, suggesting that the biologic mechanisms responsible for early and late cancer-specific events are fundamentally different. Thus the early and late effects of adjuvant therapy may vary accordingly.

For example, Figure 1 shows the annual hazard rates for breast cancer deaths (percent per year) after initial diagnosis among women in the National Cancer Institute's Surveillance, Epidemiology, and End Results 13 Registries database.² The average annual rate of breast cancer deaths is nonproportional overall and by estrogen receptor (ER) expression.¹⁴ Thus the annual hazard rate for all cases peaks near 3% per year between the second and third years after diagnosis and then declines to 1% to 2% per year by the sixth through eighth years. The hazard rates for ER-negative and ER-positive tumors peak at approximately 6.5% and 2% per year, respectively, between the first and third years (ie, > three-fold difference). Notably, ER-negative to ER-positive hazard rates cross between the seventh and eighth years, after which women with ER-negative tumors have a lower rate of breast cancer death than those with ER-positive tumors. Table 1 further shows the fold difference for ER-negative compared with ER-positive tumors over time. ER-negative to ER-positive hazard ratios (HR) were more than 1.0 before the eighth year, after which HRs were less than 1.0.

Similar nonproportional hazard rates are evident for large versus small tumors, positive versus negative lymph nodes, high versus low tumor grade,¹³ the intrinsic molecular breast cancer subtypes,^{6,8} and the molecular prognostic signatures Oncotype DX¹² and Mammaprint⁹⁻¹¹ (Fig 2). Thus hazard rates for relapse among high-risk tumors (eg, nonluminal A, Mammaprint poor signature, and Oncotype high-risk score) show a sharp peak soon after initial diagnosis, similar to ER-negative cancers (Fig 1). Conversely, hazard rates for low-risk tumors (eg, luminal A, Mammaprint good signature, and Oncotype low- and intermediate-risk score) lack a

Fig 1. Annual hazard rates for breast cancer death and ER-negative to ER-positive hazard ratios (Table 1) using the National Cancer Institute's Surveillance, Epidemiology, and End Results 13 Registries Databases (1992 to 2007) for invasive female breast cancer.² Annual hazard rates for breast cancer death overall (all cases combined, $n = 401,693$), estrogen receptor (ER) –negative ($n = 74,567$), and ER-positive ($n = 257,426$) breast cancers. The annual hazard rate for cancer-specific death describes the instantaneous rate of dying from cancer in a specified time interval after initial cancer diagnosis. Hazard rate curves were modeled using cubic splines with join-points selected by Akaike's information criteria^{3,4}; 95% CIs were applied with bootstrap resampling techniques.⁵ Under the null hypothesis of no interaction over time, annual hazard rates for ER-positive and ER-negative breast cancers would be proportional (or similar) with follow-up after initial breast cancer diagnosis. The overall rate of breast cancer death for all cases peaks near 3% per year between the second and third years after initial breast cancer diagnosis and then declines to 1% to 2% per year by the sixth through eighth years. The annual hazard rates for women with ER-negative and ER-positive tumors demonstrate peaks of approximately 6.5% and 2% near the first through third years after initial breast cancer diagnosis, respectively (> three-fold difference). An ER-negative to ER-positive hazard rate cross-over occurs between the seventh and eighth years after breast cancer diagnosis, and then women with ER-negative tumors had a somewhat paradoxically lower rate of breast cancer death than those with ER-positive breast cancers.

sharp peak, similar to ER-positive tumors. These hazard curves suggest that the biologic mechanisms responsible for early and late breast cancer events differ and may therefore respond differently to the same treatment.

Fig 2. Percent relapse-free survival (A, C, E) and annual hazard rates for breast cancer relapse (B, D, F) using data that was provided in supplemental Table 1 of Fan et al.⁶ Relapse-free survival in panels A, C, and E was obtained with the Kaplan-Meier product-limit estimator.⁷ Percent annual hazard rates for breast cancer relapse in panels B, D, and F were modeled as described in Figure 1.^{3,4} Plots for relapse-free survival and annual hazard rates for breast cancer relapse are shown for the intrinsic molecular subtypes (A, B),⁸ MammaPrint prognostic molecular signatures (C, D),⁹⁻¹¹ and Oncotype DX prognostic scores (E, F).¹² All plots demonstrate nonproportional relapse-free survival and nonproportional annual hazard rates for breast cancer recurrence. (A, B) Unsupervised hierarchical molecular clustering demonstrates two main breast cancer classes according to estrogen receptor (ER) expression and/or epithelial cell of origin (luminal or basal). There are two ER-negative (HER2 and basal-like) and two ER-positive (ie, luminal A and B) molecular signatures.⁸ Relapse-free survival is most favorable for luminal A and least favorable for non-luminal A tumors (luminal B, HER2 positive, and basal-like breast cancers). (C, D) MammaPrint is a dichotomous 70-gene prognostic signature.⁹⁻¹¹ Relapse-free survival is most favorable for a good and least favorable for a poor molecular signature. (E, F) Oncotype DX is a 21-gene prognostic score.¹² Relapse-free survival is most favorable for low and least favorable for intermediate and high molecular score.

Although randomized clinical trials compare the outcomes of two treatment groups (and therefore two distinct hazard functions) over time,¹⁵ trial results are often summarized as a single value of the hazard ratio based on the Cox proportional hazards regression model.¹⁶ The underlying assumption is that the effect of an experimental treatment in comparison with that of a standard treatment is proportional over time. If an experimental treatment reduces the risk of recurrence by 25% at 3 years, then it is assumed to have the same proportional benefit throughout all other time points as well. Thus a 25% proportional benefit implies that the risk of recurrence is reduced

from, say, 8% to 6% at year 3, 6% to 4.5% at year 4, 2% to 1.5% at year 6, and so on. Although the absolute benefit varies, the relative benefit (the hazard ratio) remains constant over time. In essence, the Cox model allows for estimation of a treatment effect without regard to any potential time-dependent fluctuations.

Yet there is now compelling evidence that commonly used adjuvant therapy regimens exert their effect primarily on the early hazard peak of recurrences and deaths (reducing the risk of events during the first 4 years after initial diagnosis). Jeong et al¹⁷ demonstrated nonproportional disease-free survival among patients treated with adjuvant

Table 1. ER-Negative to ER-Positive HRs Over Time

Years After Dx	ER-Negative HRs	ER-Positive HRs	HR
0	1.9	1.3	1.4
2	6.2	1.8	3.5
4	4.1	2.0	2.0
6	2.4	1.7	1.4
8	1.3	1.5	0.9
10	0.8	1.4	0.6
12	0.7	1.4	0.5

NOTE. As described in Figure 1A, the ER-negative to ER-positive HRs varied (ie, were nonproportional) over time. For example, HRs were more than 1.0 before the eighth year after initial breast cancer, after which HRs were less than 1.0.

Abbreviations: ER, estrogen receptor; dx, diagnosis; HR, hazard ratio.

tamoxifen versus placebo, with treatment affecting primarily the early hazard peak. Similarly, data from the National Tumor Institute of Milan revealed that nearly all the beneficial effect of adjuvant chemotherapy with cyclophosphamide, methotrexate, and fluorouracil occurred during the initial 4 years after diagnosis.¹⁸ Additionally, analysis of the Cancer and Leukemia Group B and US Breast Cancer Intergroup data revealed that high-dose versus low-dose adjuvant cyclophosphamide, doxorubicin, and fluorouracil reduced the risk of an event (recurrence or death) by 55% in the first year and 30% in the second year, with no advantage after 3 years.¹⁹ Alternatively, there are now data suggesting that the benefit of adjuvant immunotherapy could occur later, if at all.²⁰⁻²² Thus a breast cancer adjuvant immunotherapy trial may have a very different hazard ratio over time than the previous trials of adjuvant chemotherapy and endocrine therapy.

Further evidence of the nonproportional treatment effects for adjuvant therapy comes from the Early Breast Cancer Trialists Cooperative Group worldwide overviews.²³ Table 2 (adapted from Fig 4 in the Early Breast Cancer Trialists Cooperative Group overview²³) shows the average annual recurrence hazard ratio by follow-up intervals for women allocated to adjuvant polychemotherapy versus none. Although hazard ratios are lower (better) for women younger than 50 years than 50+ years for all time periods, the greatest impact of treatment occurs during the first 2 years among both younger and older women (P for time trend $< .001$). Treatment significantly

reduces overall breast cancer mortality through 15 years of observation. However, the average overall effect on mortality ratios is less than that for recurrence, particularly among women in the older age group, and there is no statistically significant trend over time (P for time trend $> .1$).

The popular simple log-rank test, which is identical to the score test from the Cox model when there is a single binary covariate, is optimal only when the proportional hazards assumption holds. If the log-rank test is used inappropriately for nonproportional hazards data, some loss of efficiency of the test is inevitable.²⁴ In this case, an optimal weight function needs to be included in the log-rank test statistic, such as the Harrington-Fleming test.²⁵

Statistical methods for dealing with nonproportional hazards using extensions of the Cox model or other parametric survival models have been available since the 1970s, although they are rarely used in the design or primary analysis of breast cancer trials.²⁶ Several methods exist, including stratification of covariates with nonproportional effects, stratification of the time scale, incorporating time-varying covariates for treatment, or the fitting of parametric models.^{27,28} Also, the concept of frailty might be appropriate to address the issue of nonproportionality.²⁴ In the frailty model, optimal weight functions for the log-rank test can be obtained from the distribution of an omitted covariate in the Cox model that induces the nonproportionality.²⁴ In addition, there is a substantial risk of dying from non-breast cancer causes as the length of follow-up increases, and this should be considered in the development of survival models.^{29,30} To distinguish breast cancer-specific mortality from other causes of death, the concept of competing risks or a cure rate model could be used to estimate the proportion correctly.

The alternative methods in the preceding paragraph can also be problematic. For example, there can be lack of flexibility in using parametric methods, even though they can provide more accurate results as long as the assumed model fits the data well. Models with time-dependent covariates may suffer from unclear interpretation of the results, biases, and/or overfitting of the data set.³¹ Competing risk data are common, but the statistical methods for analysis should be chosen with caution. For example, if one is interested in inferring the proportion of breast cancer deaths, then the method should be based on the cumulative incidence function rather than the Kaplan-Meier survival curve.^{7,29} Also, designing a study with breast cancer mortality as the primary end point would entail larger sample sizes to meet the proposed true error probabilities for a competing risk model.

To fully elucidate the long-term effects of novel therapies, long-term follow-up of patients enrolled in breast cancer clinical trials might be necessary. Indeed, if early- and long-term treatment effects are nonproportional, then the early stoppage of clinical trials might obscure important late effects. Alternatively, some adjuvant treatments may only have early benefits, and longer follow-up may dilute those effects. Thus, when developing time lines for the primary analysis of particular trials, it is essential to consider scientific, statistical, and ethical issues, as well as practical concerns such as increases in cost.

Elucidation of the time-dependent effects of adjuvant therapy will become increasingly important in the years ahead as the number of long-term breast cancer survivors continues to increase. Given that the Cox proportional hazards methods generally ignore the potential time-dependent effects of adjuvant therapy, alternative models that incorporate concepts of breast cancer heterogeneity and competing

Table 2. Worldwide Overview Hazard Ratios for Breast Cancer Recurrence and Mortality for Selected Age Groups and Time Intervals After Diagnosis

Time (years)	Age < 50 Years		Age 50-59 Years		Age < 50 Years		Age 50-59 Years	
	HR	SE	HR	SE	HR	SE	HR	SE
All years	0.63	0.03	0.81	0.02	0.71	0.04	0.88	0.03
0-1	0.52	0.05	0.64	0.03	0.74	0.10	0.82	0.06
2-4	0.69	0.06	0.92	0.04	0.74	0.06	0.87	0.04
4-9	0.79	0.08	0.99	0.05	0.68	0.07	0.91	0.05
10+	0.73	0.15	0.84	0.10	0.57	0.12	0.90	0.08

NOTE. Adapted from Figure 4 in Early Breast Cancer Trialists Cooperative Group overview.²² Estimated breast cancer mortality ratios based on calculating breast cancer mortality indirectly by subtraction of the non-breast cancer mortality rate in a specific year from the all-cause mortality rate for that year.⁵

Abbreviations: HR, hazard ratio; SE, standard error.

risks of mortality should now be considered in the design, reporting, and interpretation of clinical trials.

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

The author(s) indicated no potential conflicts of interest.

AUTHOR CONTRIBUTIONS

Manuscript writing: All authors

Final approval of manuscript: All authors

REFERENCES

1. Jatoi I, Tsimelzon A, Weiss H, et al: Hazard rates of recurrence following diagnosis of primary breast cancer. *Breast Cancer Res Treat* 89:173-178, 2005
2. Surveillance, Epidemiology, and End Results (SEER) Program: SEER*Stat Database: Incidence-SEER 13 Regs Research Data, November 2009 Submission (1992-2007) Katrina/Rita Population Adjustment—Linked To County Attributes. Total U.S., 1969-2007 Counties. National Cancer Institute, Division of Cancer Control and Population Sciences, Surveillance Research Program, Cancer Statistics Branch, released April 2010, based on November 2009 submission. <http://seer.cancer.gov/>
3. Akaike H: Information theory and an extension of the maximum likelihood principle, in Petrov BN, Csaki F (eds): 2nd International Symposium on Information Theory. Budapest, Hungary, Akademiai Kiado, 1973, pp 267-281
4. Rosenberg PS: Hazard function estimation using B-splines. *Biometrics* 51:874-887, 1995
5. Efron B, Tibshirani R: An Introduction to the Bootstrap. Monographs on Statistics and Applied Probability 57. New York, NY, Chapman & Hall, 1993
6. Fan C, Oh DS, Wessels L, et al: Concordance among gene-expression-based predictors for breast cancer. *N Engl J Med* 355:560-569, 2006
7. Kaplan EL, Meier P: Nonparametric estimation from incomplete observations. *J Am Stat Assoc* 53:457-481, 1958
8. Perou CM, Sorlie T, Eisen MB, et al: Molecular portraits of human breast tumours. *Nature* 406:747-752, 2000
9. Sotiropoulos C, Piccart MJ: Taking gene-expression profiling to the clinic: When will molecular signatures become relevant to patient care? *Nat Rev Cancer* 7:545-553, 2007
10. Desmedt C, Piette F, Loi S, et al: Strong time dependence of the 76-gene prognostic signature for node-negative breast cancer patients in the TRANSBIG multicenter independent validation series. *Clin Cancer Res* 13:3207-3214, 2007
11. Dinh P, Sotiropoulos C, Piccart MJ: The evolution of treatment strategies: Aiming at the target. *Breast* 16:S10-S16, 2007 (suppl 2)
12. Paik S, Shak S, Tang G, et al: A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med* 351:2817-2826, 2004
13. Anderson WF, Jatoi I, Devesa SS: Distinct breast cancer incidence and prognostic patterns in the NCI's SEER program: Suggesting a possible link between etiology and outcome. *Breast Cancer Res Treat* 90:127-137, 2005
14. Anderson WF, Chen BE, Jatoi I, et al: Effects of estrogen receptor expression and histopathology on annual hazard rates of death from breast cancer after primary diagnosis. *Breast Cancer Res Treat* 100:121-126, 2006
15. Berry D: Commentary: The hazards of survival comparisons. *Oncologist* 12:510-511, 2007
16. Cox DR: Regression models and life-tables. *J R Stat Soc B* 34:187-220, 1972
17. Jeong JH, Jung SH, Wieand S: A parametric model for long-term follow-up data from phase III breast cancer clinical trials. *Stat Med* 22:339-352, 2003
18. Demicheli R, Miceli R, Moliterni A, et al: Breast cancer recurrence dynamics following adjuvant CMF is consistent with tumor dormancy and mastectomy-driven acceleration of the metastatic process. *Ann Oncol* 16:1449-1457, 2005
19. Berry DA, Cirincione C, Henderson IC, et al: Estrogen-receptor status and outcomes of modern chemotherapy for patients with node-positive breast cancer. *JAMA* 295:1658-1667, 2006
20. Hoos A, Eggermont AM, Janetzki S, et al: Improved endpoints for cancer immunotherapy trials. *J Natl Cancer Inst* 102:1388-1397, 2010
21. Krieger N, Chen JT, Ware JH, et al: Race/ethnicity and breast cancer estrogen receptor status: Impact of class, missing data, and modeling assumptions. *Cancer Causes Control* 19:1305-1318, 2008
22. Berry DA: The hazards of endpoints. *J Natl Cancer Inst* 102:1376-1377, 2010
23. Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: An overview of the randomised trials. *Lancet* 365:1687-1717, 2005
24. Oakes D, Jeong JH: Frailty models and rank tests. *Lifetime Data Anal* 4:209-228, 1998
25. Harrington DP, Fleming TR: A class of rank test procedures for censored survival data. *Biometrika* 69:553-566, 1982
26. Gore SM, Pocock SJ, Kerr GR: Regression models and non-proportional hazards in the analysis of breast cancer survival. *Appl Stat* 33:176-195, 1984
27. Therneau TM, Grambsch PM: Modeling SURVIVAL DATA: Extending the Cox Model. New York, NY, Springer, 2000
28. Perperoglou A, Karamolpoulos A, Houwelingen HC: Approaches in modelling long-term survival: An application to breast cancer. *Stat Med* 26:2666-2685, 2007
29. Kalbfleisch JD, Prentice RL: The Statistical Analysis of Failure Time Data. New York, NY, Wiley & Sons, 1980
30. Gooley TA, Leisenring W, Crowley J, et al: Estimation of failure probabilities in the presence of competing risks: New representations of old estimators. *Stat Med* 18:695-706, 1999
31. Fisher LD, Lin DY: Time-dependent covariates in the Cox proportional-hazards regression model. *Annu Rev Public Health* 20:145-157, 1999

DOI: 10.1200/JCO.2010.32.3550; published online ahead of print at www.jco.org on May 9, 2011