

Published in final edited form as:

Lancet Oncol. 2010 January ; 11(1): 55–65. doi:10.1016/S1470-2045(09)70314-6.

Prognostic and Predictive Value of the 21-Gene Recurrence Score Assay in a Randomized Trial of Chemotherapy for Postmenopausal, Node-Positive, Estrogen Receptor-Positive Breast Cancer

Kathy S. Albain, MD¹[Professor], William E. Barlow, MD²[Professor], Steven Shak, MD³[no academic appointment], Gabriel N. Hortobagyi, MD⁴[Professor], Robert B. Livingston, MD⁵[Professor], I-Tien Yeh, MD⁶[Professor], Peter Ravdin, MD, PhD⁶[Professor], Roberto Bugarini, DStat³[no academic appointment], Frederick L. Baehner, MD³[no academic appointment], Nancy E. Davidson, MD⁷[Professor], George W. Sledge, MD⁸[Professor], Eric P. Winer, MD⁹[Professor], Clifford Hudis, MD¹⁰[Professor], James N. Ingle, MD¹¹[Professor], Edith A. Perez, MD¹²[Professor], Kathleen I. Pritchard, MD¹³[Professor], Lois Shepherd, MD¹⁴[Professor], Julie R. Gralow, MD¹⁵[Professor], Carl Yoshizawa, PhD³[no academic appointment], D. Craig Allred, MD¹⁶[Professor], C. Kent Osborne, MD¹⁷[Professor], and Daniel F. Hayes, MD¹⁸[Professor] for The Breast Cancer Intergroup of North America

¹ Loyola University Chicago Stritch School of Medicine, Cardinal Bernardin Cancer Center, Maywood, IL

² Southwest Oncology Group Statistical Center, Seattle, WA

³ Genomic Health, Inc., Redwood City, CA

⁴ University of Texas MD Anderson Cancer Center, Houston, TX

⁵ University of Arizona Cancer Center, Tucson, AZ

⁶ University of Texas Health Science Center at San Antonio, San Antonio, TX

⁷ University of Pittsburgh Cancer Institute, Pittsburgh, PA

⁸ Indiana University Medical Center, Indianapolis, IN

Address for editorial correspondence: Kathy S. Albain, M.D., Professor of Medicine, Loyola University Chicago Stritch School of Medicine, 2160 South First Ave, Maywood, IL 60153-5589. Phone: (708) 327-3102, Fax: (708) 327-2210, kalbain@lumc.edu.
Address for other inquiries: Group Chair's Office (SWOG-8814/21-gene), Southwest Oncology Group, 24 Frank Lloyd Wright Drive, PO Box 483, Ann Arbor, MI 48106-0483

AUTHOR CONTRIBUTIONS AND CONFLICT OF INTEREST

The corresponding author (KSA), as the principal investigator, participated in all phases of this study, including design and writing of the ancillary protocol, submission for NCI approval, analysis, interpretation and manuscript preparation. All co-authors participated in data interpretation. The study biostatisticians conducted all analyses. Coauthors reviewed the manuscript contents and approved the submission version.

The following co-authors declared occasional speaker's bureau, CME lecture and/or advisory board honoraria for Genomic Health, Inc: Kathy S. Albain, Robert B. Livingston, and Peter M. Ravdin. These authors declared research funding from Genomic Health, Inc. to their institutions but with no direct payments to themselves: Kathy S. Albain, William E. Barlow, and Daniel F. Hayes. D. Craig Allred and George W. Sledge have served as paid consultants to Genomic Health, Inc. Daniel F. Hayes has collaborated with Genomic Health, Inc. on other unfunded research endeavors. Peter M. Ravdin has ownership interest in Adjuvant Inc. Clifford Hudis had equity interest in Genomic Health, Inc. until June, 2008 when it was divested 100%. Steven Shak, Roberto Bugarini, Frederick L. Baehner, and Carl Yoshizawa are full-time employees and stockholders in Genomic Health, Inc. The following co-authors declared no relevant conflicts of interest: Nancy E. Davidson, Julie R. Gralow, Gabriel N. Hortobagyi, James N. Ingle, C. Kent Osborne, Edith A. Perez, Kathleen I. Pritchard, Lois Shepherd, Eric P. Winer, and I-Tien Yeh.

⁹ Dana-Farber Cancer Institute, Boston, MA

¹⁰ Memorial Sloan-Kettering Cancer Center, New York, NY

¹¹ Mayo Clinic, Rochester, MN

¹² Mayo Clinic, Jacksonville, FL

¹³ Sunnybrook Cancer Center, University of Toronto, Toronto, Ontario, Canada

¹⁴ National Cancer Institute of Canada Clinical Trials Group, Queen's University, Kingston, Ontario, Canada

¹⁵ Seattle Cancer Care Alliance, Seattle, WA

¹⁶ Washington University School of Medicine, St. Louis, MO

¹⁷ Baylor College of Medicine Cancer Center, Houston, TX

¹⁸ University of Michigan Comprehensive Cancer Center, Ann Arbor, MI

SUMMARY

Background—The 21-gene Recurrence Score assay (RS) is prognostic for women with node-negative, estrogen receptor (ER)-positive breast cancer (BC) treated with tamoxifen. A low RS predicts little benefit of chemotherapy. For node-positive BC, we investigated whether RS was prognostic in women treated with tamoxifen alone and whether it identified those who might not benefit from anthracycline-based chemotherapy, despite higher recurrence risks.

Methods—The phase III trial S8814 for postmenopausal women with node-positive, ER-positive BC showed that CAF chemotherapy prior to tamoxifen (CAF-T) added survival benefit to tamoxifen alone. Optional tumor banking yielded specimens for RS determination by RT-PCR. We evaluated the effect of RS on disease-free survival (DFS) by treatment group (tamoxifen versus CAF-T) using Cox regression adjusting for number of positive nodes.

Findings—There were 367 specimens (40% of parent trial) with sufficient RNA (tamoxifen, 148; CAF-T, 219). The RS was prognostic in the tamoxifen arm ($p=0.006$). There was no CAF benefit in the low RS group (logrank $p=0.97$; HR=1.02, 95% CI (0.54,1.93)), but major DFS improvement for the high RS subset (logrank $p=.03$; HR=0.59, 95% CI (0.35, 1.01)), adjusting for number of positive nodes. The RS-by-treatment interaction was significant in the first 5 years ($p=0.029$), with no additional prediction beyond 5 years ($p=0.58$), though the cumulative benefit remained at 10 years. Results were similar for overall survival and BC-specific survival.

Interpretation—In this retrospective analysis, the RS is prognostic for tamoxifen-treated patients with positive nodes and predicts significant CAF benefit in tumors with a high RS. A low RS identifies women who may not benefit from anthracycline-based chemotherapy despite positive nodes.

Multi-gene tumor assays report useful prognostic information for women with axillary node-negative breast cancer (BC).(1-4) Of these, the 21-gene Recurrence Score assay (RS) provides a prognosis for patients with estrogen receptor (ER)-positive disease treated with tamoxifen alone.(1) In one study, the RS also predicts chemotherapy benefit from standard chemotherapy (CMF).(5) Patients with high RS appeared to benefit greatly from the addition of chemotherapy to tamoxifen, whereas those with low RS did not.

Recent studies showed the value of the RS in addition to the standard pathology report(6-8), improving physician and patient decision-making in lymph node negative scenarios. Use of the RS as a prognostic and predictive tool in ER-positive, lymph node-negative BC was recommended by the American Society of Clinical Oncology.(9)

There have been no assessments of the value of the RS in patients with ER-positive disease and involved axillary nodes from a study that contains a tamoxifen-alone control group. These patients are routinely treated with chemotherapy and endocrine adjuvant therapy.(10) But exploratory data suggest that those with higher tumor ER levels may not derive chemotherapy benefit, even if at high recurrence risk due to positive nodes.(11–13) Some studies have demonstrated less chemotherapy benefit when the node-positive disease was both ER-positive and HER2-negative.(11,14,15)

Consequently, we analyzed the 21-gene RS assay in a phase III node-positive trial that contains a tamoxifen-alone control. Our two co-primary objectives were to determine whether the assay 1) provides prognostic information for women with node-positive disease treated only with tamoxifen, and 2) allows prediction of a node-positive group that does not benefit from anthracycline-based chemotherapy.

PATIENTS AND METHODS

Summary of the Parent Trial (SWOG 8814/TBCI 0100)

Postmenopausal women with axillary node-positive BC classified as ER and/or progesterone receptor (PgR)-positive by local institutional standards were eligible. Stratification variables included number of positive nodes (1–3, ≥ 4), PgR status and interval from surgery. Randomization was to 5 years of tamoxifen alone, 6 cycles of CAF (16) followed by tamoxifen (CAF-T), or CAF with concurrent tamoxifen (CAFT) in a 1.0:1.5:1.5 ratio. Disease-free survival (DFS) was the time from registration to BC relapse (local or distant), new primary BC, or death due to any cause, whichever came first. Overall survival (OS) was calculated from registration to death due to any cause. Patients without an event were censored at the last follow-up visit. Follow-up for recurrence ended after mature 10-year survival data were available due to financial constraints, but known deaths are still recorded.

The comparison of the combined chemotherapy groups (CAF-T and CAFT) to tamoxifen alone showed superior DFS and OS over 10 years.(17) The addition of chemotherapy sequentially (CAF-T) was better than simultaneous administration (CAFT). Number of involved nodes (1–3 versus ≥ 4) was highly prognostic, but CAF benefit remained after adjustment for nodes and other variables.

Design of this Translational Study

This National Cancer Institute-approved translational study (NCI #8814A-ICSC) was led by the Southwest Oncology Group (SWOG) for The Breast Cancer Intergroup of North America and is reported according to REMARK recommendations.(18) When enrolled on the parent trial, patients were asked to allow central banking of their paraffin-embedded BC for future studies and signed a separate informed consent document for evaluation of biomarkers measured on tumor tissue in relation to outcome (protocol S9445). Genomic Health Inc. (GHI) performed the 21-gene RS assay (Oncotype DX[®]) blinded to patient clinical data including outcome. The design and statistical plan were finalized before merging the assay results and clinical data and analyzing the data at the SWOG Statistical Center. The study was approved by an independent central Institutional Review Board.

Because of the inferior efficacy of concurrent tamoxifen and CAF in the parent trial, that arm was excluded, so this analysis compares the sequential CAF-T group to the tamoxifen control. Tumor samples were available for 664 (45%) of the 1,477 patients in the parent study (Figure 1), including 413 (45%) of the 927 on the two arms included here. The RT-PCR analysis was feasible in 367 specimens (40%) of the 2 relevant arms of S8814: 148 of 166 (89%) samples for tamoxifen alone and in 219 of 247 (89%) for CAF-T, suggesting no bias by arm in sample availability. Analyses were not performed in the remainder (11%) due

to exhaustion of invasive tumor in the block, no submission of primary tumor, or technical issues.

The GHI RT-PCR assay on the pre-specified 21 genes (16 cancer-related genes including groups related to ER/PgR, proliferation, HER2 and invasion; and, 5 reference genes) isolated RNA from fixed, paraffin-embedded tissue, according to standardized methodology(1). The RS was derived from reference-normalized gene expression measurements, and ranged from 0 to 100.

Tumor grade was evaluated centrally (by FLB) using modified Bloom-Richardson score from H&E-stained tissue sections. In a previous exploratory biomarker study(11), central immunohistochemistry for ER by Allred score(19), HER2 by TAB250, and P53 were scored (by DCA) on most samples available in the current study.

Statistical Considerations

The primary, pre-specified outcome was DFS, with OS as a secondary endpoint as in the parent trial. Distant recurrence-free interval (DRFI) was not available, so we conducted an exploratory analysis of breast cancer specific survival (BCSS). In BCSS, only deaths due to BC were events, censoring all deaths not due to BC at time of death and alive patients at the last follow-up visit. Two-sided $\alpha = 0.05$ significance levels were used. The primary analysis specified modeling continuous RS as a linear term in a Cox regression model. Although analyses utilized RS as a continuous variable, secondary analyses used the clinical RS categories of low (<18), intermediate ($18-30$), and high (≥ 31).(1)

For the first co-primary objective, the prognostic effect of RS, we examined whether higher RS was associated with shorter DFS in the tamoxifen-alone group. The second co-primary objective of the predictive effect of the RS was tested by including an interaction term of continuous RS and chemotherapy in the model. This tested whether the difference in randomized treatment depended on increasing RS.

Cox regression models were adjusted for number of positive nodes (1-3 versus 4+), a stratifying, highly prognostic factor from the parent trial. The assumption of proportional hazards, tested in each model, was satisfied except when RS was included in the model, indicating the effect of RS was not constant over the entire time period. Thus the time axis was divided into <5 years and ≥ 5 years (at the end of tamoxifen therapy and midway in follow-up) allowing estimation and testing of different hazard ratios for each time period. Cox models for each period separately showed no violation of proportional hazards. Kaplan-Meier survival plots were employed and log-rank tests (stratified by number of positive nodes) of survival truncated at 10 years (due to low numbers at risk after 10 years) were used to test differences between survival curves. Cox models include 13 years of follow-up since they reflect the entire follow-up period and are less influenced by the low numbers at risk after 10 years. To determine the estimated probability of a DFS event by 5 or 10 years the linear RS was allowed to have time-varying effects using a flexible proportional odds approach(20) that included number of positive nodes (1-3 versus 4+) as a covariate. For Figures 6 and 7 prediction of CAF benefit was presented only for $RS \leq 50$ due to high uncertainty at greater RS levels and shown separately for the prognostic strata of 1-3 and ≥ 4 positive nodes.(21)

Role of the Funding Source

The parent randomized trial was funded entirely by the U.S. National Cancer Institute. This retrospective evaluation of banked tumor tissue was funded jointly by the National Cancer Institute and by Genomic Health Incorporated (GHI). The design of the study was approved by the North American Breast Cancer Intergroup, the Southwest Oncology Group (SWOG),

and GHI and subsequently approved by independent peer review by the NCI Tumor assays were performed by GHI without knowledge of treatment assignment or clinical outcome. These data were then merged with clinical data at the SWOG Statistical Center. The study biostatistician (WEB) had the only direct access to all data. Analytic results were confirmed by GHI statisticians (CY, RB) by visiting the SWOG Statistical Center. Four co-authors (SS, RB, FB, CY) are employees of a sponsor (GHI) and contributed to the interpretation and writing of the manuscript. The manuscript was drafted in its entirety by the co-authors without benefit of paid assistance. Content of the final manuscript was not subject to approval from the National Cancer Institute or the corporate sponsor. The corresponding author (KSA) had full and final responsibility to submit for publication.

RESULTS

Patient Demographics and Study Sample Descriptors

Patients in this study were representative of the parent trial by age, race, PgR status and duration of follow-up (Table 1). This sample had slightly lower number of positive nodes and smaller tumor size and 12% were HER2-positive based on the 21-gene assay. The RS was distributed over the three risk levels and balanced between treatment groups. The overall DFS benefit for CAF-T versus tamoxifen alone in the parent trial was comparable in this study subset after adjustment for number of positive nodes. The DFS hazard ratio (HR) for chemotherapy versus not was 0.69 (95% CI 0.56,0.84; $p<0.001$) in the parent trial and 0.72 (95% CI 0.51,1.00; $p=0.048$) in this subset. In the parent trial, the HR for OS was 0.78 (95% CI 0.63,0.97; $p=0.024$) and in this subset was 0.77 (95% CI 0.52,1.14; $p=0.19$), adjusted for number of positive nodes.

Prognosis of Patients in Tamoxifen-Along Group by RS

The RS was highly prognostic for DFS within the tamoxifen-alone group (Figure 2, Panel A), stratified by number of positive nodes ($p=0.017$). The 10-year DFS estimates were 60%, 49% and 43% for low, intermediate and high risk categories, respectively. In a Cox regression model, the continuous RS was highly significant ($p=0.006$), with HR=2.64 (95% CI 1.33,5.27; $p=0.006$) for a 50 point difference. The HR for RS was not constant over time by the test for proportional hazards ($p=0.0016$). In the first 5 years, the HR was 5.55 (95% CI 2.32,3.28; $p<0.001$). For those surviving beyond 5 years, the RS was no longer prognostic (HR=0.86; $p=0.80$), but the initial strong effect persisted over the entire period.

The RS risk category was prognostic for OS over 10 years (stratified log-rank $p=0.003$) in the tamoxifen-alone group (Figure 2, Panel B). The 10-year OS estimates for low, intermediate and high RS were 77%, 68%, and 51%, respectively. Adjusting for nodes, the OS HR was 4.42 (95% CI 1.96,9.97; $p<0.001$) for a 50 point difference, with similar failure of proportional hazards assumption.

Prediction of CAF Benefit by RS

The RS was a strong predictive factor of CAF benefit for DFS (Figure 3). Panel A shows improved DFS over 10 years for CAF-T versus tamoxifen alone in the entire RS sample (stratified log-rank $p=0.054$, adjusted for number of positive nodes), but degree of CAF benefit depended on the RS. There was no apparent benefit for scores <18 (Panel B, logrank $p=0.97$; HR=1.02, 95% CI (0.54–1.93)) or 18–30 (Panel C, logrank $p=0.48$; HR=0.72, 95% CI (0.39–1.31)). However, there was a significant advantage for CAF-T compared to tamoxifen alone for patients with RS ≥ 31 (Panel D, logrank $p=0.033$; HR=0.59, 95% CI (0.35–1.01)). For patients with low RS, the 10-year DFS estimates for CAF-T versus tamoxifen were 64% versus 60% and for those with high RS, were 55% versus 43%, respectively.

Similar differences in the RS predictive utility were observed for OS over 10 years (Figure 4). There was no statistically significant benefit to CAF for the low ($p=0.63$, Panel A) or intermediate ($p=0.85$, Panel B) RS groups. However, there was a significant CAF benefit in the high RS group ($p=0.0271$, Panel C), which did not vary by age. Ten-year estimates for OS in high RS for CAF-T versus tamoxifen were 68% and 51%, respectively. Corresponding OS hazard ratios for chemotherapy versus no chemotherapy adjusting for nodes were 1.18 (95% CI 0.55, 2.54; $p=0.68$) for low RS, 0.84 (95% CI 0.40, 1.78; $p=0.65$) for intermediate RS, and 0.56 (95% CI 0.31, 1.02; $p=0.057$) for high RS. Similar outcomes were observed for BCSS (Panels D–F), with 10-year estimates for the high RS group of 73% for CAF-T and 54% for tamoxifen ($p=0.033$).

Figure 5 depicts DFS HRs for CAF benefit for the parent trial, the entire RS subset, and then by categorized RS. Hazard ratios in the parent trial and the entire RS subset show a consistent benefit over time (i.e., proportional hazards), with chemotherapy effect lasting beyond 5 years. The high RS subset shows an even stronger benefit that also persists over time. Failure of the proportional hazards assumption is seen for the low and intermediate risk groups which have inconsistent effects over time. There is no suggestion of benefit in the low risk group overall or in the first five years. In the intermediate group there may be slight benefit overall, but not in the first five years. Confidence intervals are wide due to small numbers of later events. Only in the low risk group is there no evidence of a benefit cumulatively over the entire period.

The primary analysis of prediction was to test increasing chemotherapy benefit as the linear RS increased. We analyzed the interaction of treatment effect and the linear RS, adjusting for number of positive nodes (1–3 versus 4+). Table 2 shows the model, calibrated to RS=0 as the referent and RS/50 (i.e. corresponds to a 50 point difference); all estimated HRs; the interaction p-values for this test; and examples of HRs for CAF-T versus tamoxifen at selected RS values. Over the entire time period, the significance of the RS-treatment interaction is $p=0.053$ for DFS. However, the effect of the RS on treatment is not constant over time. In the first five years, RS predicts chemotherapy benefit (interaction $p=0.029$; Table 2), but not after five years ($p=0.58$). Nevertheless, the cumulative CAF benefit persists out to 10 years. In the OS analysis, there was a significant interaction of RS and treatment overall ($p=0.026$), in the first 5 years ($p=0.016$), but not after 5 years ($p=0.87$). Therefore, RS has both strong prognostic and predictive effects on survival in the first 5 years, but limited *additional* effects in women surviving beyond 5 years (except in higher RS). The strong initial effects carry forward sufficiently so overall differences are still seen at late time points.

Prediction of CAF Benefit by Recurrence Score and Number of Positive Nodes

Prediction of any DFS event within 10 years is displayed by number of positive nodes, treatment, and RS in Figure 6A. Increasing involvement of axillary lymph nodes was prognostic. The treatments start diverging at approximately RS=10, though any clinically significant CAF benefit is not evident until much higher RS. Because the RS has better short- than long-term prediction, estimates at 5 years are presented in Figure 6B. The treatments are equivalent up to approximately RS=20, but diverge at higher RS values. The 95% prediction intervals around the estimates are depicted in Webfigure 1. These bounds are specific to a particular RS value so cannot be used to test the significance of chemotherapy benefit which depends on a range of RS values.

Single or Combination Markers and Prediction of Outcomes

We assessed whether other markers measured by central pathologic review could predict degree of chemotherapy benefit as well as the RS risk categories. Tumor grade was

prognostic for DFS overall ($p=0.008$) but did not interact with prediction of chemotherapy benefit ($p=0.26$). ER by Allred scoring (19) was available for 316 (86%) of the sample with RS and HER2 by TAB250 was available for 352 (96%), with 312 (85%) having both. The best cut-point for clinical use of Allred-scored ER was 0–6 ($n=147$, 46.5%) versus 7–8 ($n=169$, 53.5%) with a marginal predictive effect ($p=0.16$). There may be a CAF benefit if the disease was either HER2-positive and/or ER score ≥ 6 ($n=170$, $p=0.06$, stratified log rank test at 10 years). However, there was no DFS benefit if ER score was high (7 or 8) and the disease was HER2-negative ($n=142$, $p=0.81$). In this latter group, the RS distribution was low (58%), intermediate (24%), and high (18%).

The interaction of treatment benefit and RS remained significant adjusting for age, race, tumor size, PgR status, grade, p53 and HER2 by TAB250. Because ER is a part of the RS, adjusting for Allred-scored ER level did make the interaction non-significant ($p=.15$). There was a moderate negative (-0.38) correlation of Allred-scored ER with RS, though some tumors with high ER (by Allred score or by RT-PCR from the RS assay) had a high RS (Webfigure 2). Thus, the predictive capability of the RS may not be completely captured by consideration of known markers measured by immunohistochemistry.

DISCUSSION

Our study suggests patients with involved axillary lymph nodes but a low RS do not appear to benefit from anthracycline-based chemotherapy, whereas those with a higher RS have major benefit, independent of number of positive nodes. TRANSBIG collaborators presented analyses of a non-randomized cohort of 106 patients with 1–3 positive nodes. A subset identified as low-risk by the 70-gene profile (3) had identical survival whether chemotherapy was given or not. (22) Taken together, these data suggest that there may be subgroups within the ER-positive, node-positive BC population that behave differently with respect to the otherwise-accepted role for chemotherapy, and that these subgroups can be identified using multigene assays. (23,24) {paragraphs 2 and 1 switched}

This study challenges the current treatment standard of adjuvant chemotherapy for all women with positive axillary nodes and ER-positive BC. (25) This standard is based on several decades of phase III clinical trials that demonstrated a survival benefit to chemotherapy when added to endocrine therapy alone in premenopausal and more recently, postmenopausal women. (10,11,26) In a recent international survey, identification of a molecular signature to select patients who could be spared chemotherapy was voted the highest translational research priority in breast cancer worldwide. (27) Avoiding the toxicity and other costs of adjuvant chemotherapy when it may not be needed is an important goal in BC treatment.

There is an ongoing debate regarding the independent added role of the RS and other multigene assays to standard pathology variables for prognosis and prediction. While the 21-gene RS assay provides a reproducible method to classify the biology of a given patient's tumor for prediction of chemotherapy benefit, standard pathology testing may provide another means of determining chemotherapy benefit. In exploratory, post-hoc analyses, high levels of ER protein expression ("endocrine responsiveness") measured centrally predicted lack of chemotherapy benefit. (11–13,26), and St. Gallen guidelines endorse the use of degree of endocrine responsiveness in chemotherapy decision-making. (26,28) In our study, a subset defined by both high ER protein level and HER2-negative disease appeared to have no benefit from CAF added to tamoxifen.

It would take a much larger study than ours to demonstrate a statistically significant increase in prediction using a multigene assay after accounting for standard pathological assays. In

part, this is attributable to measuring the same pathways by both methods, so one method must have much less measurement error to demonstrate improvement. However, our exploratory analysis and those of others have been consistent that a significant multigene assay-chemotherapy benefit interaction was maintained after adjustment for standard factors.(1·24·29·30). Specifically for the RS assay, it provided better discrimination of individual tumor behavior and a more reliable prediction of those who would benefit versus not compared to the traditional assays in these studies. Furthermore, there is a 25–30% discordance rate between risk levels predicted by standard variables and multigene assays(24). Ongoing prospective trials should answer how to best select therapy when this discordance exists. For now, the new St. Gallen guidelines allow the use of multigene assays to select adjuvant therapy(28).

It remains to be demonstrated that less costly and more available assays would actually lead to different clinical decisions about treatment. That said, in decision-making studies the use of multigene assays result in a change in treatment plan about a third of the time, and this change usually is to avoid chemotherapy when it was initially thought to be needed pre-assay. (6·8·24).

There are limitations regarding our results. This was a population of postmenopausal women with ER-positive, node-positive BC, so it is unclear whether the findings translate to *premenopausal* patients. However, the performance of the assay in node-negative disease was the same across all ages.(1·5) Our results with anthracycline-based chemotherapy and those of the NSABP with CMF(5) are based on older standards of chemotherapy, so the predictive utility of the RS assay may differ in current practice using other types of chemotherapy or dosing schedules. While high RSs are associated with more pathologic complete remissions from taxanes given in the neoadjuvant setting(31), RS prediction of taxane efficacy from phase III trials is not available. Nonetheless, these analyses and others with different gene profiles suggest that certain biologic subtypes of breast cancer may be inherently sensitive or resistant to chemotherapy in general.

Our retrospective analyses involved a subset of the parent trial, although overall treatment effect and demographics were similar to the parent trial. Given the low event rate, particularly in the low RS group, confidence intervals were broad so that estimated CAF benefit at specific RS values should be interpreted with caution. Whereas there was no apparent CAF benefit in low RS for all endpoints, the possibility of benefit cannot be completely ruled out. The lack of proportional hazards observed in our study is supported by previous reports regarding the major impact of adjuvant chemotherapy in the first years of follow-up(10), the indolent nature of luminal A biology over time(32·33), and early-onset recurrence in high RS tumors(1). Finally, our study employed DFS as the primary endpoint, since unlike the NSABP analysis(1), we did not prospectively collect DRFI. Thus, the prognostic and predictive effects of the RS may differ due to the inclusion of DFS events such as second primary cancers and breast recurrences. However, the results for BCSS are also consistent.

In conclusion, our study provides further data on the value of a multi-gene assay for prognosis in patients with ER-positive, node-positive BC treated with adjuvant tamoxifen. Moreover, our results suggest that the 21-gene RS assay may predict which of these patients derive benefit from an anthracycline-based chemotherapy regimen and those who may not, despite higher risk due to positive nodes. Current treatment guidelines generally recommend chemotherapy for high risk BC.(25) Prospective studies with larger samples are critical to determine who optimally benefits from modern endocrine therapy plus chemotherapy, and whether use of these assays impact survival.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This investigation was supported in part by the following PHS Cooperative Agreement grant numbers awarded by the National Cancer Institute, DHHS, CA32102, CA38926, CA21115, CA02599, CA60138, CA25224, CA77202-06, CA04920, CA58658, CA13612, CA37981, CA76447, CA22433, CA58416, CA20319, CA58686, CA04919, CA46441, CA58861, CA27057, CA32734, CA35281, CA12644, CA16385, CA45560, CA58882, CA14028, CA35176, CA46282, CA46113, CA52650, CA03096, CA28862, CA35090, CA58723, CA35283, CA45807, CA35200, CA35119, CA45450, CA46136, CA42777, CA35261, CA45466, CA35117, CA46368, CA58348, CA12213, CA52654, CA35128, CA58415, CA52623, CA35192, CA45377, CA35996, CA52757, CA76132, CA35431, CA76462, CA45461, CA35084, CA76429, CA35178, CA67663, CA63844, CA52772; by the National Cancer Institute of Canada and Canadian Cancer Society; and supported in part by Genomic Health, Inc.

Funding: This work was funded by the National Cancer Institute and by Genomic Health Incorporated.

References

1. Paik S, Shak S, Tang G, Kim C, Baker J, Cronin M, Baehner FL, Walker MG, Watson D, Park T, Hiller W, Fisher ER, Wickerham DL, Bryant J, Wolmark N. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med.* 2004;351:2817–26. [PubMed: 15591335]
2. Habel LA, Shak S, Jacobs MK, Capra A, Alexander C, Pho M, Baker J, Walker M, Watson D, Hackett J, Blick NT, Greenberg D, Fehrenbacher L, Langholz B, Quesenberry CP. A population-based study of tumor gene expression and risk of breast cancer death among lymph node-negative patients. *Breast Cancer Research.* 2006;8(3):R25. [PubMed: 16737553]
3. Van de Vijver MJ, He YD, van't Veer LJ, Dai H, Hart AAM, Voskuil DW, et al. A gene-expression signature as a predictor of survival in breast cancer. *N Engl J Med.* 2002;247:1999–2009. [PubMed: 12490681]
4. Van't Veer LJ, Paik S, Hayes DF. Gene expression profiling of breast cancer: a new tumor marker. *J Clin Oncol.* 2005;23(8):1631–5. [PubMed: 15755970]
5. Paik S, Tang G, Shak S, Kim C, Baker J, Kim W, Cronin W, Baehner BL, Watson D, Bryant J, Costantino J, Geyer CE Jr, Wickerham DL, Wolmark N. Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor-positive breast cancer. *J Clin Oncol.* 2006;24:3726–3734. [PubMed: 16720680]
6. Oratz R, Paul D, Cohn AL, Sedlacek SM. Impact of a commercial reference laboratory test reference laboratory test recurrence score on decision making in early-stage breast cancer. *J Oncol Practice.* 2007;3:182–6.
7. Kamal AH, Loprinzi CL, Reynolds C, Dueck AC, Geiger XJ, Ingle LN, Carlson RW, Hobday TJ, Winer EP, Perez EA, Goetz MP. How well do standard prognostic criteria predict Oncotype Dx scores? *J Clin Oncol.* 2007;25(18S):576.
8. Lo SS, Mumby PB, Rychlik K, Smerage J, Kash J, Chew HK, Hayes D, Albain KS. Extended follow-up results from a prospective multi-center study of the impact of the 21-gene recurrence score assay on medical oncologist and patient adjuvant breast cancer treatment selection, satisfaction and anxiety. *J Clin Oncol.* 2009 In Press [as senior author KSA attests galleys have been returned to the JCO and publication is pending the editorial, will be published this month we are told].
9. Harris L, Fritsche H, Mennel R, Norton L, Ravdin P, Taube S, et al. American Society of Clinical Oncology 2007 update of recommendations for the use of tumor markers in breast cancer. *J Clin Oncol.* 2007;25(33):5287–312. [PubMed: 17954709]
10. Early Breast Cancer Trialists' Collaborative Group. Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomized trials. *Lancet.* 2005;365:1687–1717. [PubMed: 15894097]
11. Albain K, Barlow W, O'Malley F, et al. Concurrent versus sequential chemohormonal therapy versus tamoxifen alone for postmenopausal, node-positive, ER and/or PgR-positive breast cancer:

- mature outcomes and new biologic correlates on phase III intergroup trial 0100 (S8814). *Breast Cancer Res Treat.* 2005;90:95.
12. Fisher B, Jeong J-H, Bryant J, et al. Treatment of lymph-node-negative, oestrogen-receptor-positive breast cancer: long-term findings from National Surgical Adjuvant Breast and Bowel Project randomized clinical trials. *Lancet.* 2004;364:858–868. [PubMed: 15351193]
 13. Castiglione-Gertsch M, Price KN, Goldhirsch A, et al. Endocrine responsiveness and tailoring adjuvant therapy for postmenopausal lymph node-negative breast cancer: A randomized trial of the International Breast Cancer Study Group (Trial IX). *J Natl Cancer Inst.* 2002;94:1054–65.
 14. Hayes DF, Thor AD, Dressler LG, et al. HER2 and response to paclitaxel in node-positive breast cancer. *N Engl J Med.* 2007;257:1496–506. [PubMed: 17928597]
 15. Berry DA, Cirincione C, Henderson IC, et al. Estrogen-receptor status and outcomes of modern chemotherapy for patients with node-positive breast cancer. *JAMA.* 2006;295:1658–1667. [PubMed: 16609087]
 16. Bull JM, Tormey DC, Li SH, et al. A randomized comparative trial of Adriamycin versus methotrexate in combination drug therapy. *Cancer.* 1978;41:1649–57. [PubMed: 348293]
 17. Albain KS, Barlow WE, Ravdin PM, et al. A randomized trial of adjuvant chemotherapy and tamoxifen timing in postmenopausal, endocrine-responsive, node-positive breast cancer. *The Lancet.* 2009 In Press.
 18. McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, Clark GM. for the Statistics Subcommittee of the NCI-EORTC Working Group on Cancer Diagnostics. Reporting recommendations for tumour MARKer prognostic studies (REMARK). *British J Cancer.* 2005;93:387–391.
 19. Allred DC, Harvey JM, Berardo M, Clark GM. Prognostic and predictive factors in breast cancer by immunohistochemical analysis. *Modern Pathol.* 1998;11:155–168.
 20. Royston P, Parmar MK. Flexible parametric proportional-hazards and proportional-odds models for censored survival data, with application to prognostic modelling and estimation of treatment effects. *Stat Med.* 2002 Aug 15; 21(15):2175–97. [PubMed: 12210632]
 21. Singletary SE, Allred C, Ashley P, et al. Revision of the American Joint Committee on Cancer staging system for breast cancer. *J Clin Oncol.* 2002;20(17):3628–36. [PubMed: 12202663]
 22. Mook S, Schmidt MK, Viale G, et al. The 70-gene prognosis-signature predicts disease outcome in breast cancer patients with 1–3 positive lymph nodes in an independent validation study. *Breast Cancer Res Treat.* 2008 (published online July 27).
 23. Sotiriou C, Pusztai L. Gene-expression signatures in breast cancer. *New Engl J Med.* 2009;360:790–800. [PubMed: 19228622]
 24. Albain KS, Paik S, van't Veer L. Prediction of adjuvant chemotherapy benefit in endocrine-responsive early breast cancer using multigene assays. *The Breast.* 2009 In Press [KSA attests that galleys have been returned to journal and publication is any time now].
 25. NCCN Practice Guidelines in Oncology. Invasive Breast Cancer. www.NCCN.org, v.2.2008
 26. Goldhirsch A, Wood WC, Gelber RD, Coates AS, Thurlimann B, Senn H-J. Panel members. Progress and promise: highlights of the international expert consensus on the primary therapy of early breast cancer 2007. *Annals of Oncology.* 2007;18:1133–44. [PubMed: 17675394]
 27. Dowsett M, Goldhirsch A, Hayes DF, Senn HJ, Wood W, Viale G. International web-based consultation on priorities for translational breast cancer research. *Breast Cancer Res.* 2007;9(6):R81. [PubMed: 18034879]
 28. Goldhirsch A, Ingle JN, Gelber RD, Coates AS, Thurlimann B, Senn H-J. Panel members. Thresholds for therapies: highlights of the St. Gallen international expert consensus on the primary therapy of early breast cancer 2009. *Annals of Oncology.* 2009;20:1319–1329. [PubMed: 19535820]
 29. Goldstein LJ, Gray R, Badve S, et al. Prognostic utility of the 21-gene assay in hormone receptor-positive operable breast cancer compared with classical clinicopathologic features. *J Clin Oncol.* 2008;26(25):4063–71. [PubMed: 18678838]
 30. Knauer M, Straver M, Rutgers E, et al. The 70-gene MammaPrint signature is predictive for chemotherapy benefit in early breast cancer. *The Breast.* 2009;18(suppl 1):S36. (abstract 73).

31. Gianni L, Zambetti M, Clark K, et al. Gene expression profiles in paraffin-embedded core biopsy tissue predict response to chemotherapy in women with locally advanced breast cancer. *J Clin Oncol.* 2005;23:7265–7277. [PubMed: 16145055]
32. Chapman J-A, Meng D, Shepherd L, et al. Competing causes of death from a randomized trial of extended adjuvant endocrine therapy for breast cancer. *J Natl Cancer Inst.* 2008;100:252–260. [PubMed: 18270335]
33. Saphner T, Tormey DC, Gray R. Annual hazard rates of recurrence for breast cancer after primary therapy. *J Clin Oncol.* 1996;13:2738–2746. [PubMed: 8874335]

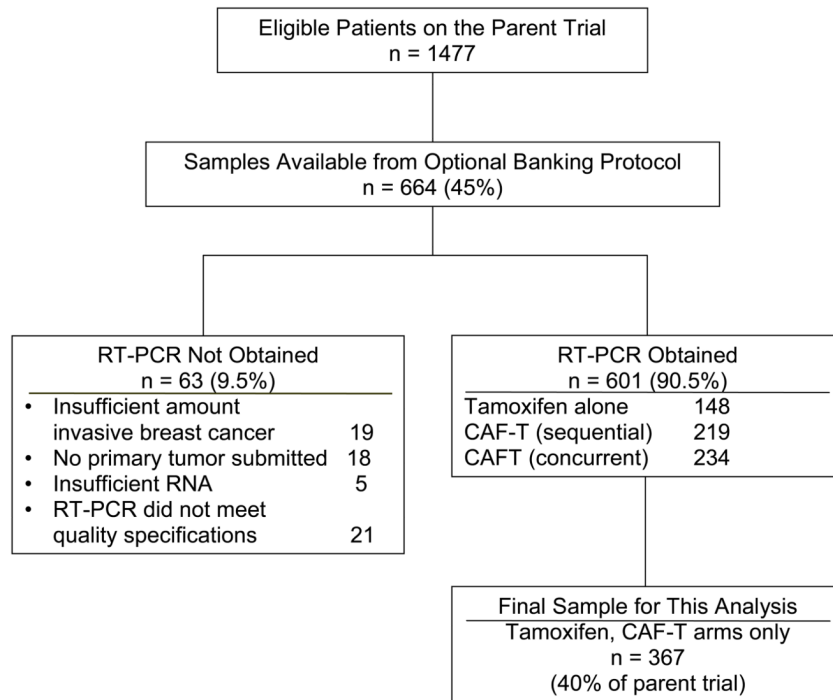


Figure 1. Modified REMARK diagram, showing the specimen acquisition, distribution and processing for the RT-PCR analyses, resulting in the final sample size of 367 patients.

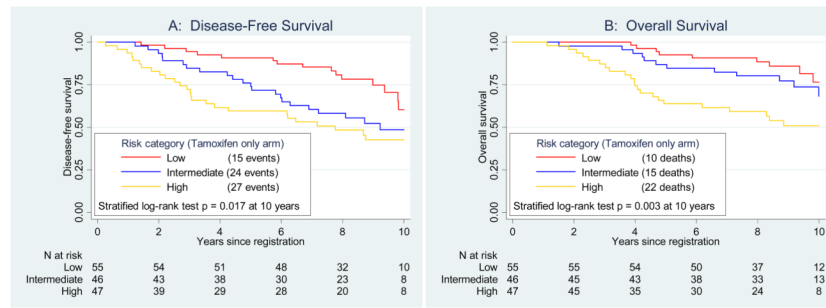


Figure 2. Prognostic disease-free survival and overall survival analyses by Recurrence Score (RS) group in patients treated with tamoxifen alone. Panel A depicts disease-free survival and Panel B, overall survival. The log-rank tests are stratified by number of positive nodes.

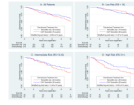


Figure 3. Primary disease-free survival endpoint by treatment and Recurrence Score (RS) groups. The log-rank tests are stratified by number of positive nodes. Panel A shows disease-free survival by treatment (CAF followed by tamoxifen versus tamoxifen alone) overall, and Panels B–D depicts the outcomes within each RS risk group of low, intermediate, and high, respectively.

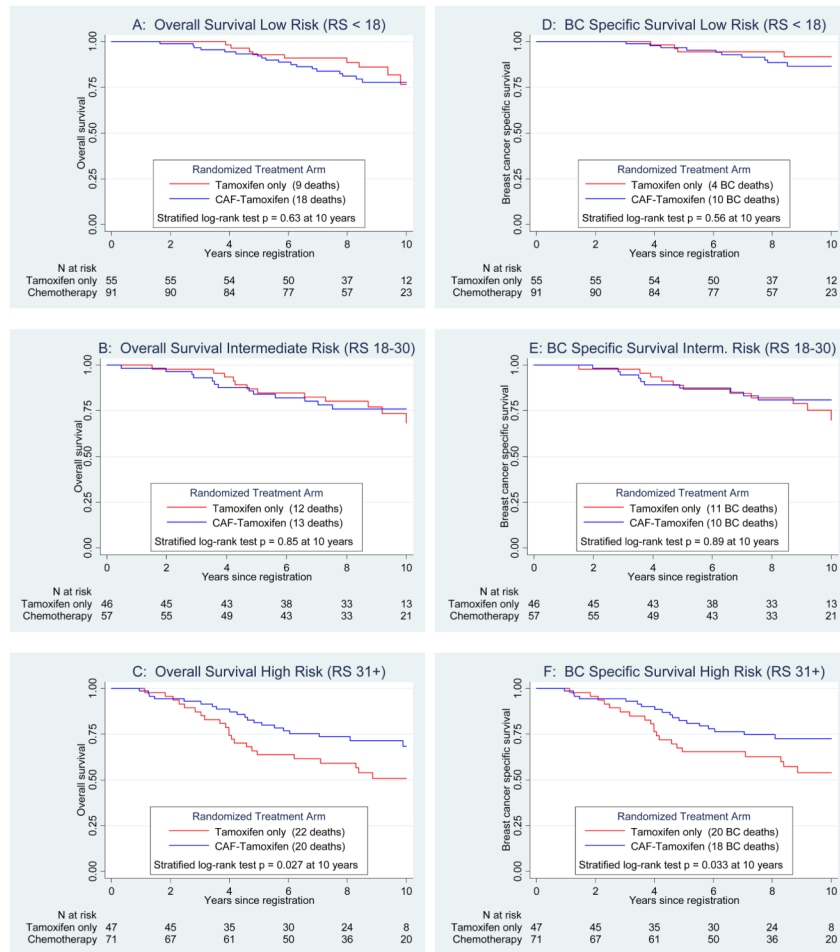


Figure 4. Secondary endpoints of overall survival by Recurrence Score (RS) groups (Panels A, B and C) and the exploratory endpoint of breast cancer specific survival by RS groups (Panels D, E and F), all adjusted for number of positive nodes.

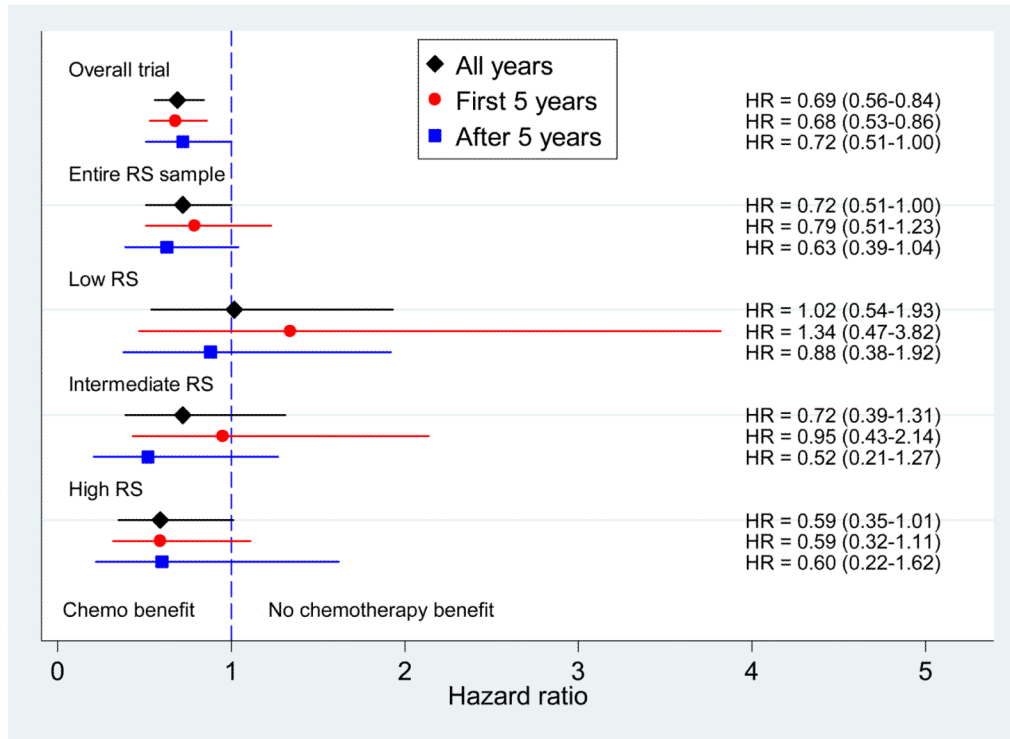


Figure 5. Disease-free survival hazard ratios with 95% confidence intervals for tamoxifen alone versus CAF followed by tamoxifen displayed (from top to bottom) for the overall parent trial, the entire Recurrence Score (RS) sample, and RS groups of low, intermediate, and high. The diamond indicates all years, the circle the first 5 years and the square, after 5 years.

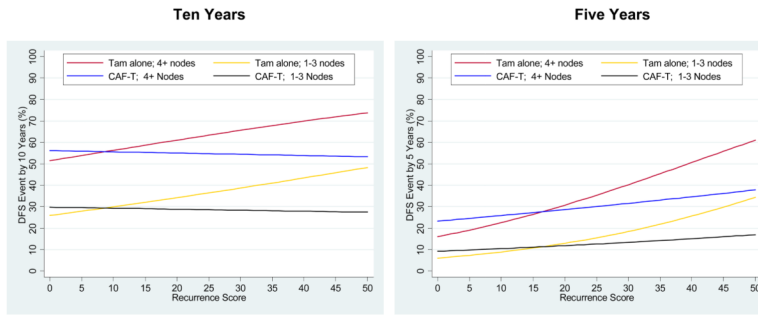


Figure 6. Risks of a disease-free survival (DFS) event by linear Recurrence Score, by treatment (tamoxifen alone and CAF followed by tamoxifen) and number of positive nodes, depicted for 10-year (left panel) and 5-year (right panel) landmarks. In each panel, the top two lines are represent the subset with 4 or more positive axillary nodes and the bottom 2 lines, 1–3 positive nodes.

Table 1

Patient and Tumor Characteristics in This Study versus the Parent Trial

Characteristic	This Study			Parent Trial: Tamoxifen only and CAF-T Arms (n=927)
	Tamoxifen only (n=148)	CAF then Tamoxifen (n=219)	Overall (n=367)	
Age Range	45–79 years	42–81 years	42–81 years	37–81 years
Mean (SD)	60.8 years (7.8)	60.1 years (7.4)	60.4 years (7.5)	61.1 years (7.2)
Age 30–54	35 (23.7%)	55 (25.1%)	90 (24.5%)	205 (22.1%)
Age 55–64	62 (41.9%)	107 (48.9%)	169 (46.1%)	443 (47.8%)
Age 65+	51 (34.5%)	57 (26.0%)	108 (29.4%)	279 (30.1%)
1–3 positive nodes	94 (63.5%)	133 (60.7%)	227 (61.9%)	541 (58.4%)
ER positive by RT-PCR assay	145 (98.0%)	210 (95.9%)	355 (96.7%)	NA
Black race	12 (8.1%)	15 (6.8%)	27 (7.4%)	83 (8.9%)
Tumor size				
< 2 cm	46 (31.1%)	74 (33.8%)	120 (32.7%)	292 (31.5%)
2–5 cm	94 (63.5%)	136 (62.1%)	230 (62.7%)	568 (61.3%)
> 5 cm	8 (5.4%)	9 (4.1%)	17 (4.6%)	67 (7.2%)
PgR negative by RT-PCR assay	27 (18.2%)	49 (22.4%)	76 (20.7%)	NA
PgR negative by local institution	30 (20.3%)	45 (20.6%)	75 (20.4%)	210 (22.6%)
HER2 positive by RT-PCR assay	13 (8.8%)	30 (13.7%)	43 (11.7%)	NA
Tumor grade				
1	55 (37.2%)	76 (34.7%)	131 (35.7%)	NA
2	82 (55.4%)	112 (51.1%)	194 (52.9%)	
3	11 (7.4%)	31 (14.2%)	42 (11.4%)	
Recurrence Score				
Range	0 – 85	0 – 93	0 – 93	
Mean (SD)	26.1 (17.0)	27.0 (19.9)	26.6 (18.8)	
Low risk (< 18)	55 (37.2%)	91 (41.6%)	146 (39.8%)	NA
Intermediate risk (18–30)	46 (31.1%)	57 (26.0%)	103 (28.1%)	
High risk (≥31)	47 (31.8%)	71 (32.4%)	118 (32.2%)	
Mean follow-up for DFS (censored only)	9.1 years	9.0 years	9.0 years	9.2 years
DFS event	66 (44.6%)	77 (35.2%)	143 (39.0%)	395 (42.6%)
Deaths	47 (31.8%)	55 (25.1%)	102 (27.8%)	321 (34.7%)

Table 2
Disease-Free Survival Hazard Ratios Adjusted for Number of positive nodes for Chemotherapy Benefit by Recurrence Score over Time

	All Years		First 5 years		After 5 years	
	HR	95% CI	HR	95% CI	HR	95% CI
Modeled HR estimates						
Nodes (4+)	2.44	1.75-3.42	2.49	1.58-3.92	2.37	1.44-3.91
Chemotherapy at RS=0	1.12	0.61-2.06	1.58	0.66-3.76	0.78	0.34-1.83
RS/50 (50 point difference)	2.71	1.37-5.36	5.77	2.42-13.79	0.92	0.30-2.83
Chemo * RS/50	0.43	0.18-1.01	0.30	0.10-0.89	0.66	0.16-2.82
Interaction p-value	(0.053)	-	(0.029)	-	(0.58)	-
Treatment effect overall*	HR	95% CI	HR	95% CI	HR	95% CI
Entire RS sample	0.72	0.51 - 1.00	0.79	0.51-1.23	0.63	0.39-1.04
At selected RS values						
RS = 10	0.95	0.59-1.52	1.24	0.62-2.48	0.72	0.38-1.36
RS = 18	0.83	0.56-1.22	1.03	0.58-1.81	0.67	0.40-1.14
RS = 25	0.74	0.53-1.04	0.87	0.53-1.42	0.64	0.39-1.05
RS = 31	0.67	0.48-0.93	0.75	0.48-1.18	0.61	0.35-1.04
RS = 40	0.57	0.39-0.83	0.61	0.38-0.96	0.56	0.28-1.11

* Benefit in disease-free survival of CAF added to tamoxifen versus tamoxifen alone