Prognostic and predictive value of the 21-gene recurrence score assay in postmenopausal women with node-positive, oestrogen-receptor-positive breast cancer on chemotherapy: a retrospective analysis of a randomised trial

Kathy S Albain, William E Barlow, Steven Shak, Gabriel N Hortobagyi, Robert B Livingston, I-Tien Yeh, Peter Ravdin, Roberto Bugarini, Frederick L Baehner, Nancy E Davidson, George W Sledge, Eric P Winer, Clifford Hudis, James N Ingle, Edith A Perez, Kathleen I Pritchard, Lois Shepherd, Julie R Gralow, Carl Yoshizawa, D Craig Allred, C Kent Osborne, Daniel F Hayes, for The Breast Cancer Intergroup of North America

SWOG-8814入组了绝经后HR+,LN+患者比较TAM单药和CAF-TAM的效果(后者更优)
本研究仅入组了一部分SWOG-8814的患者,淋巴结阳性数目以及肿瘤大小都没有原人群中多/大

原人群中CT对比TAM，DFS的相对风险HR=0.69 p=0.0003；OS的相对风险HR=0.78 p=0.024
本研究入组人群中,CT对比TAM，DFS相对风险HR=0.72 p=0.048；OS的相对风险HR=0.77 p=0.19

在TAM单药组，RS是DFS的强烈预测因子，且可以根据阳性淋巴结数目分层。
低风险/中风险/高风险组10年DFS分别为：60% / 49% / 43%
将RS最为连续变量，风险比例模型分析中出现50差分时DFS的风险比HR=2.64 p=0.006
，但HR并不是一成不变的p=0.0016.在前五年HR=5.55 p=0.0002，五年之后RS不再有预测效应(HR=0.86)，但由于前五年RS的强烈作用其预测作用在整个研究时间内都存在。
在TAM单药组，RS是OS的强烈预测因子，且可以根据淋巴结阳性数目进行分层。
低风险/中风险/高风险组10年OS分别为：77% / 68% / 51%
再去除淋巴结状态影响后，风险比例模型中出现50差分时OS风险比HR=4.42 p=0.0006,同样RS也不是在整个时长内有预测作用的p=0.0005

RS是10年DFS从化疗中获益的强烈预测因素。
在本研究的总体人群中，化疗带来的10年DFS获益 74/219 vs 63/148 p=0.054.而从化疗中的获益是可以根据RS风险进行分层的。
在低风险组, CAR-T vs TAM 10年DFS：64% vs 60%; HR=1.02 p=0.97(患者从化疗中无显著获益)
在中风险组 HR=0.72 p=0.48(患者从化疗中获益无统计学意义)；
在高风险组 10年DFS: 55% vs 43%; HR=0.59 p=0.033 患者从化疗中有显著的获益。

同样的，RS对于10年OS的预测效应差异在RS不同风险分组中也存在。在低风险组HR=1.18 p=0.63 中风险组HR=0.84 p=0.85.在高风险组 CAR-T vs TAM 10年OS: 68% vs 51% HR=0.56 p=0.027

相似的结果还体现在乳腺癌乳腺癌特异性生存中

在SWOG-8814的总体人群中 和 本研究总体人群中化疗带来的DFS获益在整个的研究时长内都是持续存在的，并持续到五年以上(HR始终偏向于化疗组)
但根据RS进行分层的人群中，仅仅在高风险组体现出化疗的获益并持续在五年以上，低或中风险组HR随时间推移是不一致的。低风险组不能从化疗中获益，而中风险组可以从化疗中获益，但不是在前五年内。

在对化疗获益与RS作为连续变量进行分析，并根据淋巴结阳性数目进行分层后发现，随着时间的推移，RS的影响并不是恒定的:RS可以预测化疗前5年的DFS(交互作用p=0·029)，但不能预测5年后的DFS(p=0.58)。同样RS与OS之间也存在这种关系。

在综合淋巴结受累数目、治疗方式评估RS对10年DFS的预测关系分析时，淋巴结受累数目增多时DFS的独立预测因子,在RS大于10左右治疗方式对DFS开始产生影响。
由于RS对短期的DFS预测比长期更为显著，在分析5年DFS时，RS大于20左右治疗方式才会对DFS产生影响

Patients with a high recurrence score seemed to benefit greatly from the addition of chemotherapy to tamoxifen, whereas those with a low recurrence score did not.

Recent studies have shown the value of the recurrence score when used with the standard pathology report,[6–8] which improved physician and patient decision making in lymph-node-negative scenarios. Use of the recurrence

These patients are routinely treated with chemotherapy and endocrine adjuvant therapy.[10] However, exploratory data suggest that those with higher concentrations of tumour oestrogen receptors might not derive benefit from chemotherapy, even if they are at high risk of recurrence because of positive nodes.[11–13] Some studies have shown less benefit of chemotherapy when the

(Prof J N Ingle MD); **Mayo Clinic, Jacksonville, FL, USA** (Prof E A Perez MD); **Sunnybrook Cancer Center, University of Toronto, Toronto, ON, Canada** (Prof K I Pritchard MD); **National Cancer Institute of Canada**

Clinical Trials Group, Queen's University, Kingston, ON, Canada (Prof L Shepherd MD); **Seattle Cancer Care Alliance, Seattle, WA, USA** (Prof J R Gralow MD); **Washington University School of Medicine, St Louis, MO, USA** (Prof D C Allred MD); **Baylor College of Medicine Cancer Center, Houston, TX, USA** (Prof C K Osborne MD); **and University of Michigan Comprehensive Cancer Center, Ann Arbor, MI, USA** (Prof D F Hayes MD)

Correspondence to: Prof Kathy S Albain, Loyola University Chicago Stritch School of Medicine, 2160 South First Ave, Maywood, IL 60153-5589, USA
kalbain@lumc.edu

node-positive disease was both oestrogen-receptor positive and HER2 negative (ERBB2 negative).[11,14,15]

Consequently, we analysed the 21-gene recurrence score assay in a phase 3 node-positive trial that contained a tamoxifen-only control group.[16] Our two co-primary objectives were to establish whether the assay provides prognostic information for women with node-positive disease treated only with tamoxifen, and whether the assay allows prediction of a node-positive group that does not benefit from anthracycline-based chemotherapy.

Methods

Patients and procedures

The Southwest Oncology Group (SWOG)-8814, INT-0100 study was a phase 3, open-label, parallel-group, randomised controlled trial.[16] The study design and main results of the trial have been reported elsewhere.[16] Briefly, postmenopausal women with axillary node-positive breast cancer were eligible for inclusion if they had oestrogen-receptor-positive or progesterone-receptor-positive tumours, or both, classified by local institutional standards. Enrolled patients were randomly assigned in a 2:3:3 ratio to one of three drug regimens: (1) tamoxifen alone (20 mg per day orally) for 5 years; (2) six cycles of CAF (cyclophosphamide 100 mg/m² orally on days 1–14, doxorubicin 30 mg/m² intravenously on days 1 and 8, and fluorouracil 500 mg/m² intravenously on days 1 and 8)[17] followed by tamoxifen (CAF-T); or (3) CAF with concurrent tamoxifen (CAFT).

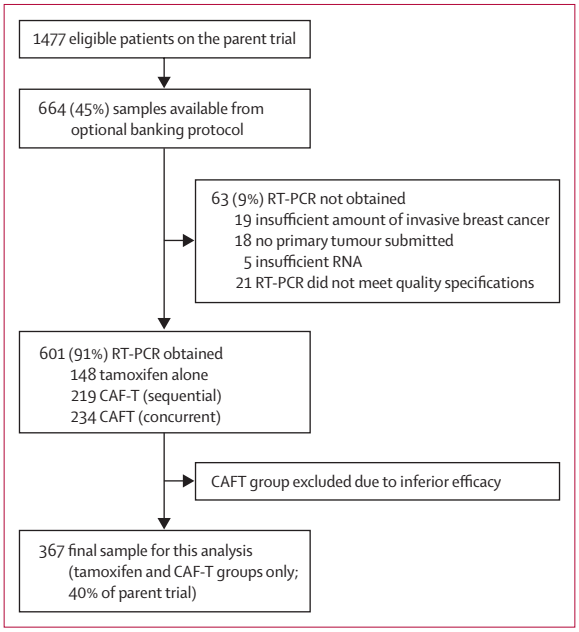

Figure 1: **Modified REMARK diagram**
Profile shows the specimen acquisition, distribution, and processing for the RT-PCR analyses, resulting in the final sample size of 367 patients. REMARK=Reporting Recommendations for Tumour Marker Prognostic Studies. CAF-T=cyclophosphamide, doxorubicin, and fluorouracil followed by tamoxifen. CAFT=cyclophosphamide, doxorubicin, and fluorouracil with concurrent tamoxifen.

CAF cycles were repeated every 28 days. Randomisation was done by computer-generated sequence and stratified by number of positive nodes (1–3 *vs* ≥4), progesterone-receptor status (positive *vs* negative), and interval from surgery (≤6 weeks *vs* >6 weeks).

The primary endpoint of SWOG-8814 was disease-free survival, defined as time from registration to breast-cancer relapse (local or distant), new primary breast cancer, or death due to any cause, whichever came first. Overall survival was calculated from registration to death due to any cause. Patients without an event were censored at the last follow-up visit. After mature 10-year survival data were obtained, follow-up for recurrence ended because of financial constraints, but known deaths are still recorded.

The combined chemotherapy groups (CAF-T and CAFT) showed superior disease-free survival and overall survival over 10 years compared with the tamoxifen group.[16] The addition of chemotherapy sequentially (CAF-T) was better than simultaneous treatment (CAFT). Patients with four or more positive nodes derived more benefit from chemotherapy than did those with one to three positive nodes, but CAF benefit remained after adjustment for nodal status and other variables.

Translational study design

This translational study, approved by the National Cancer Institute (NCI #8814A-ICSC), was led by SWOG for The Breast Cancer Intergroup of North America and is reported according to the Reporting Recommendations for Tumour Marker Prognostic Studies (REMARK).[18] When enrolled on SWOG-8814, we asked patients for permission for central banking of their paraffin-embedded tumours for future studies. Consenting patients signed a separate informed consent document for assessment of biomarkers measured in tumour tissue in relation to outcome (protocol SWOG-9445).

Laboratory personnel at Genomic Health (Redwood City, CA, USA), who were masked to patient clinical data including outcome, undertook the 21-gene recurrence score assay (Onco*type* DX). The design and statistical plan were finalised before merging the assay results and clinical data and analysing the data at the SWOG Statistical Center. The study was approved by an independent central institutional review board.

Because of the inferior efficacy of the CAFT regimen compared with the CAF-T regimen in the parent trial, we excluded CAFT from this analysis, and therefore only compared the sequential CAF-T group with the tamoxifen control group.

The RT-PCR assay was done on the 21 prespecified genes (16 cancer-related genes—including groups related to oestrogen receptor, progesterone receptor, proliferation, HER2, and invasion—and five reference genes) by use of isolated RNA from fixed, paraffin-embedded tissue, in accordance with standardised methods.[1] The recurrence score was derived from

reference-normalised gene-expression measurements, and ranged from 0 to 100.

Tumour grade was assessed centrally (by FLB) by use of the modified Bloom-Richardson score from haemotoxylin-eosin-stained tissue sections. In a previous exploratory biomarker study,[11] we undertook central immunohistochemistry (scored by DCA) for oestrogen receptor by the Allred score,[19] HER2 by TAB250, and P53 on most samples available in the current study.

Statistical analysis

The primary, prespecified outcome of the translational study was disease-free survival, with overall survival as a secondary endpoint, as in the parent trial. Since the distant recurrence-free interval was not available, we undertook an exploratory analysis of breast-cancer-specific survival. In this exploratory analysis, only deaths due to breast cancer were events, censoring all deaths not due to breast cancer at time of death and alive patients at the last follow-up visit. We used two-sided $\alpha=0\cdot05$ significance levels. The primary analysis specified modelling continuous recurrence score as a linear term in a Cox regression model. Although analyses used recurrence score as a continuous variable, secondary analyses used the clinical recurrence score categories of low (<18), intermediate (18–30), and high (≥31).[1]

For the first co-primary objective, the prognostic effect of recurrence score, we examined whether higher recurrence score was associated with shorter disease-free survival in the tamoxifen-alone group. The second co-primary objective of the predictive effect of the recurrence score was tested by including an interaction term of continuous recurrence score and chemotherapy in the model. This model tested whether the difference in outcome from randomised treatment depended on increasing recurrence score.

Cox regression models were adjusted for number of positive nodes (1–3 *vs* ≥4), a stratifying, highly prognostic factor from the parent trial. The assumption of

	This study			Parent trial: tamoxifen alone and CAF-T groups (n=927)
	Tamoxifen alone (n=148)	CAF-T (n=219)	Overall (n=367)	
Age (years)				
Mean (SD; range)	60·8 (7·8; 45–79)	60·1 (7·4; 42–81)	60·4 (7·5; 42–81)	61·1 (7·2; 37–81)
30–54	35 (23·6)	55 (25·1)	90 (24·5)	205 (22·1)
55–64	62 (41·9)	107 (48·9)	169 (46·0)	443 (47·8)
≥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
Ethnic origin (black)	12 (8·1)	15 (6·8)	27 (7·4)	83 (8·9)
Tumour 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 (20·5)	75 (20·4)	210 (22·7)
HER2-positive by RT-PCR assay	13 (8·8)	30 (13·7)	43 (11·7)	NA
Tumour grade				
1	55 (37·2)	76 (34·7)	131 (35·7)	NA
2	82 (55·4)	112 (51·1)	194 (52·9)	NA
3	11 (7·4)	31 (14·2)	42 (11·4)	NA
Recurrence score				
Mean (SD; range)	26·1 (17·0; 0–85)	27·0 (19·9; 0–93)	26·6 (18·8; 0–93)	NA
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)	NA
High risk (≥31)	47 (31·8)	71 (32·4)	118 (32·2)	NA
Mean follow-up for disease-free survival (censored only; years)	9·1	9·0	9·0	9·2
Disease-free survival event	66 (44·6)	77 (35·2)	143 (39·0)	395 (42·6)
Deaths	47 (31·8)	55 (25·1)	102 (27·8)	321 (34·6)

Data are n (%) unless otherwise indicated. CAF-T=cyclophosphamide, doxorubicin, and fluorouracil followed by tamoxifen. ER=oestrogen receptor. NA=not available. PgR=progesterone receptor.

Table 1: **Patient and tumour characteristics in this study compared with the parent trial**

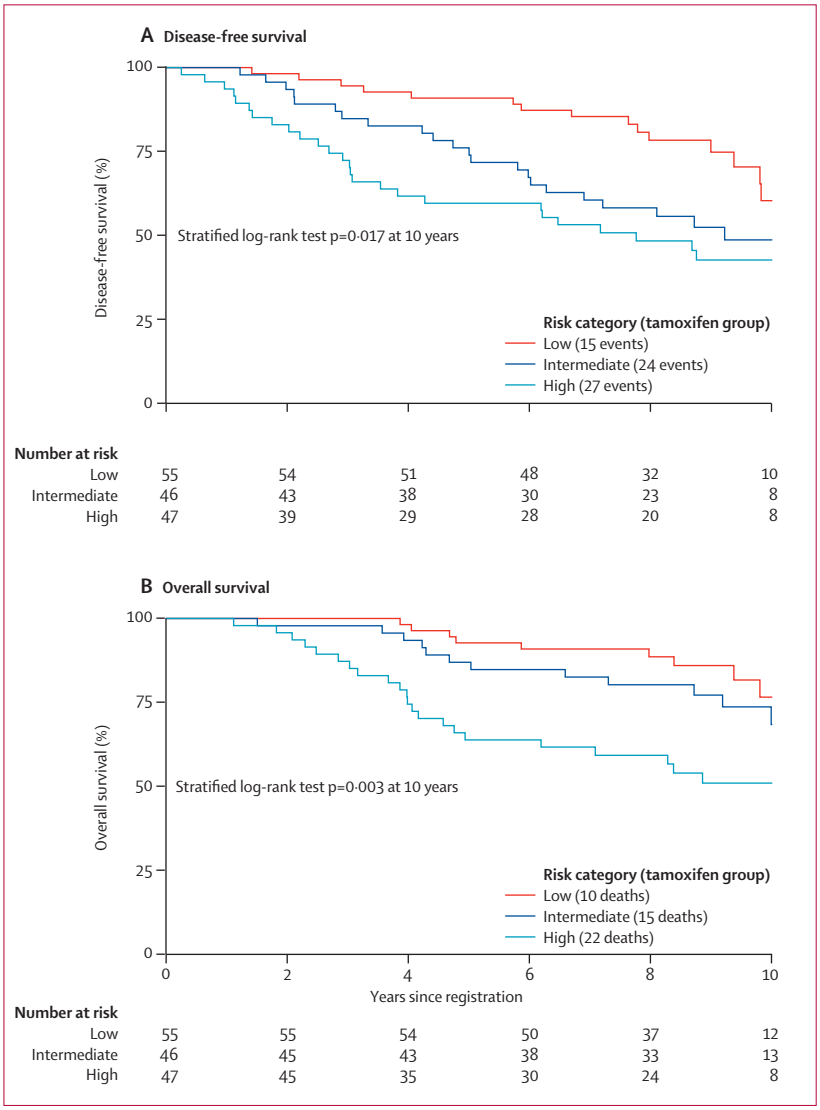

***Figure 2*: Prognostic disease-free survival and overall survival analyses by recurrence score group in patients assigned to tamoxifen alone**
The log-rank tests are stratified by number of positive nodes.

proportional hazards, tested in each model, was satisfied apart from when recurrence score was included in the model, suggesting that the effect of recurrence score was not constant over the entire time period. Thus, the time axis was divided into less than 5 years and 5 years or more (at the end of tamoxifen therapy and midway in follow-up), allowing estimation and testing of different hazard ratios (HRs) for each time period. Cox models for each period showed no violation of proportional hazards. We constructed Kaplan-Meier survival plots and used log-rank tests (stratified by number of positive nodes) of survival truncated at 10 years (because of low numbers at risk after 10 years) to test differences between survival curves. Cox models included 13 years of follow-up to reflect the entire follow-up period and since the models are less affected by the low numbers at risk after 10 years. To determine the estimated probability of a disease-free survival event by 5 or 10 years, the linear recurrence score was allowed to have time-varying effects by means of a flexible proportional odds approach[20] that included number of positive nodes (1–3 *vs* ≥4) as a covariate. For graphs showing risk of disease-free survival event by recurrence score and treatment, we presented prediction of CAF benefit for recurrence score of 50 or lower because of high uncertainty at greater recurrence score levels. In these graphs, we also showed risks separately for the prognostic strata of one to three positive nodes and four or more positive nodes.[21] Statistical analyses were done with Stata version 10.1.

Role of the funding sources

The design of the study was approved by The Breast Cancer Intergroup of North America, SWOG, and Genomic Health, and subsequently approved by independent peer review by the National Cancer Institute. Tumour assays were undertaken by laboratory personnel at Genomic Health who had no 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 in the study. Analytical results were confirmed by Genomic Health statisticians (CY, RB) by visiting the SWOG Statistical Center. Four authors (SS, RB, FLB, CY) are employees of a sponsor and contributed to the interpretation and writing of the report. The report was drafted in its entirety by the authors without benefit of paid assistance. Content of the final report was not subject to approval from the National Cancer Institute or the corporate sponsor. The corresponding author had final responsibility for the decision to submit for publication.

Results

Tumour samples were available for 664 (45%) of the 1477 patients in SWOG-8814 (figure 1), including 413 (45%) of the 927 patients in the CAF-T and tamoxifen groups. RT-PCR analysis was feasible in 367 (40%) specimens from the tamoxifen and CAF-T groups (tamoxifen, 148 [89%] of 166 samples; CAF-T, 219 [89%] of 247), suggesting no bias by group in sample availability. Analyses were not done for the remaining 46 (11%) of samples because of exhaustion of invasive tumour in the block, no submission of primary tumour, or technical issues.

The subset of patients analysed in this study were representative of those in the parent trial by age, ethnic origin, progesterone-receptor status, and duration of follow-up (table 1). However, patients in this subset had a slightly lower number of positive nodes and a smaller tumour size than did those in the parent trial. 11·7% of patients in this analysis were HER2-positive based on

the 21-gene assay. The recurrence score was distributed over the three risk levels and balanced between treatment groups.

The benefit in disease-free survival for CAF-T versus tamoxifen alone in the parent trial was similar to that seen in this subset of patients after adjustment for number of positive nodes. The HR for disease-free survival for chemotherapy versus tamoxifen was 0·69 (95% CI 0·56–0·84; p=0·0003) in the parent trial and 0·72 (0·51–1·00; p=0·048) for this subset. The HR for overall survival was 0·78 (0·63–0·97; p=0·024) in the parent trial and 0·77 (0·52–1·14; p=0·19) in this subset, adjusted for number of positive nodes.

The recurrence score was highly prognostic for disease-free survival within the tamoxifen-alone group (figure 2A), stratified by number of positive nodes (p=0·017). The 10-year disease-free survival estimates were 60%, 49%, and 43% for low, intermediate, and high-risk categories, respectively. In a Cox regression model, the continuous recurrence score was highly significant (p=0·006), with HR 2·64 (95% CI 1·33–5·27) for a 50-point difference. The HR for recurrence score was not constant over time by the test for proportional hazards (p=0·0016). In the first 5 years, the HR was 5·55 (2·32–3·28; p=0·0002). For those patients who survived beyond 5 years, the recurrence score was no longer prognostic (HR 0·86, 0·27–2·74; p=0·80), but the initial strong effect persisted over the entire period.

The recurrence score risk category was prognostic for overall survival over 10 years (stratified log-rank p=0·003) in the tamoxifen-alone group (figure 2B). The 10-year overall survival estimates for patients with low, intermediate, and high recurrence scores were 77%, 68%, and 51%, respectively. After adjustment for number of positive nodes, the HR for overall survival was 4·42 (95% CI 1·96–9·97; p=0·0006) for a 50-point

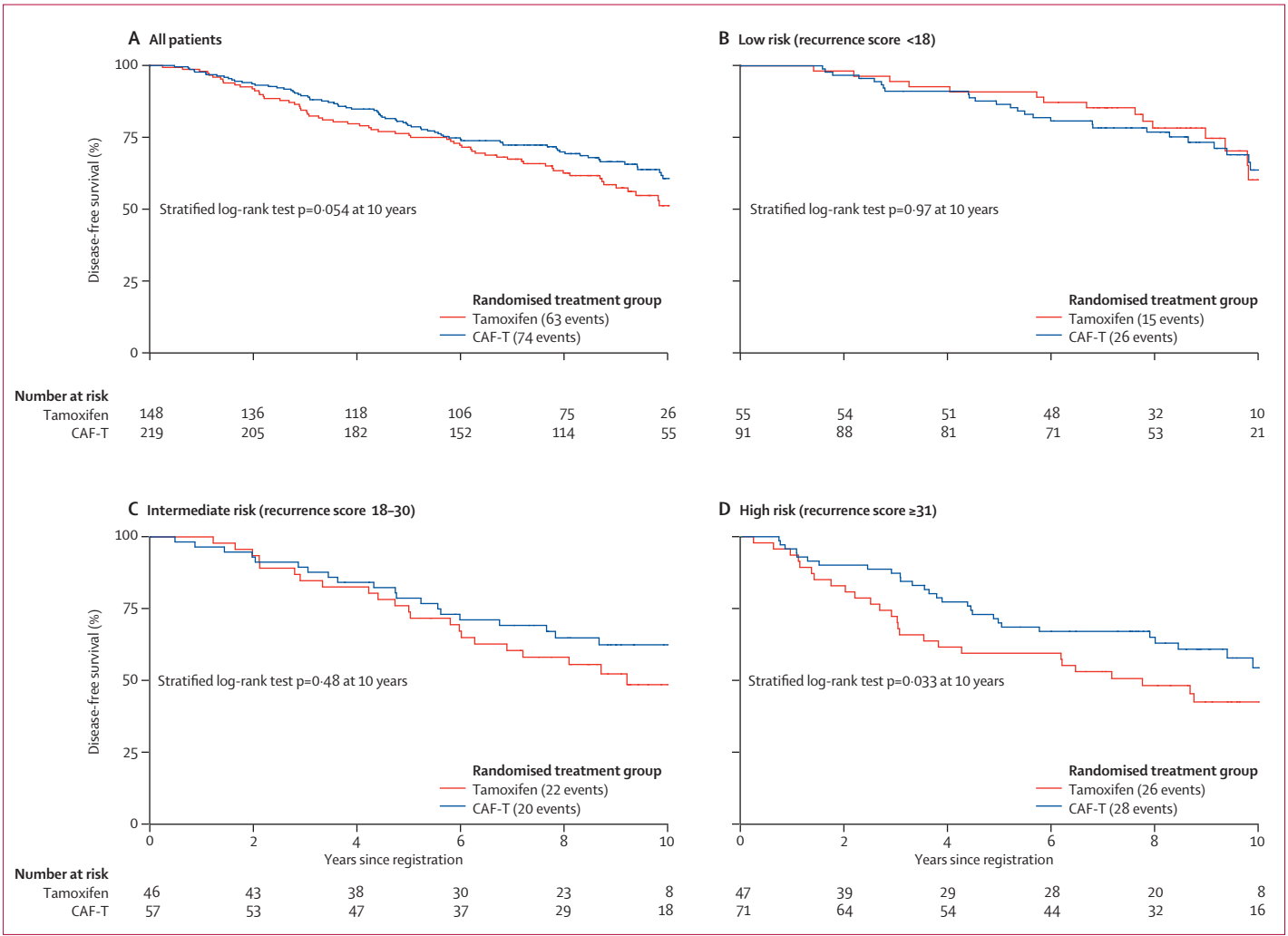

Figure 3: **Primary disease-free survival endpoint by treatment and recurrence score groups**
Disease-free survival by treatment (CAF-T vs tamoxifen alone) overall (A), and outcomes within each recurrence score risk group of low (B), intermediate (C), and high (D). The log-rank tests are stratified by number of positive nodes. CAF-T=cyclophosphamide, doxorubicin, and fluorouracil followed by tamoxifen.

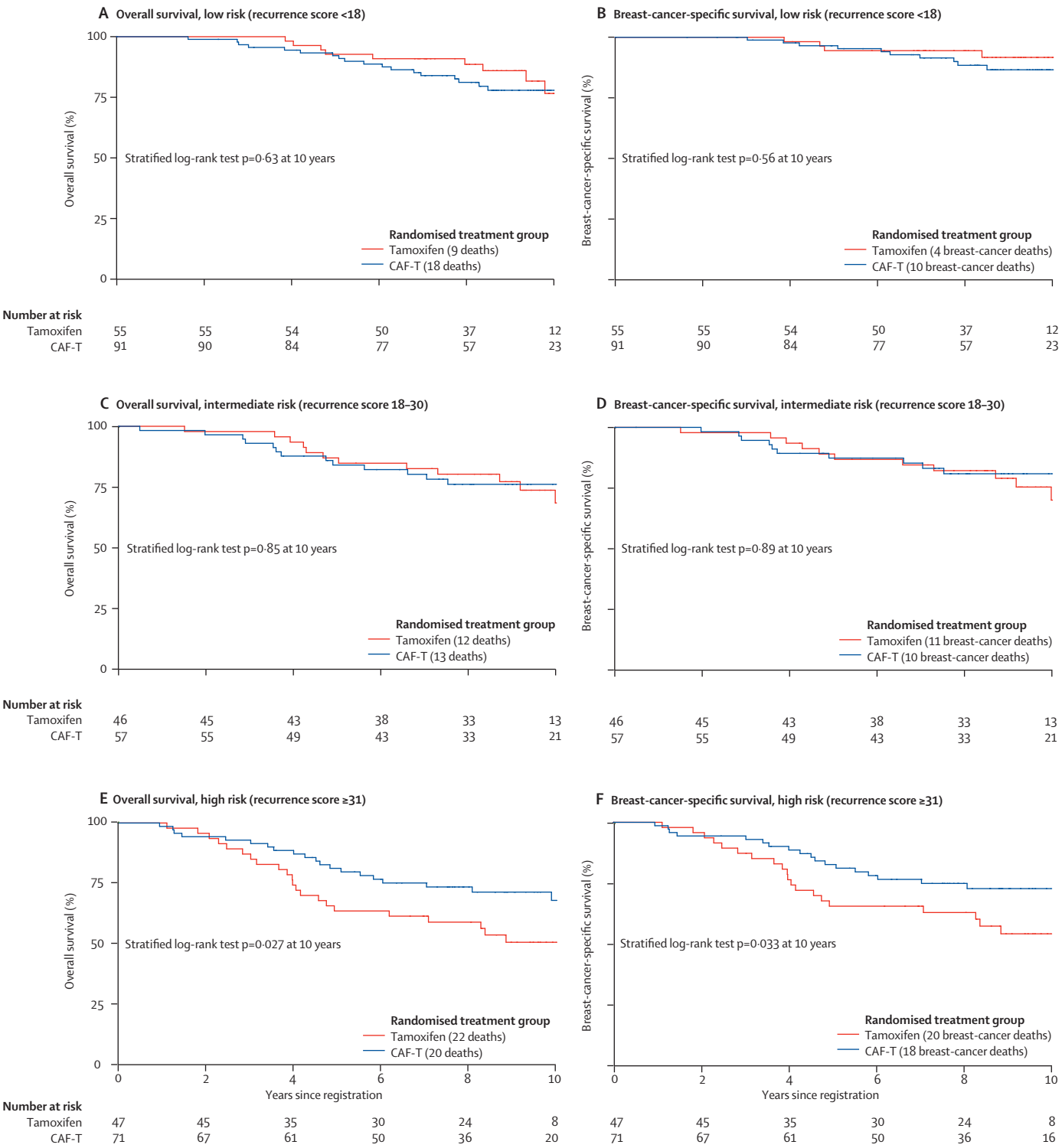

***Figure 4*: Secondary endpoint of overall survival and exploratory endpoint of breast-cancer-specific survival by recurrence score group**
Overall survival by recurrence score group (A, C, and E) and breast-cancer-specific survival by recurrence score group (B, D, and F), all adjusted for number of positive nodes. CAF-T=cyclophosphamide, doxorubicin, and fluorouracil followed by tamoxifen.

difference, with similar failure of proportional hazards assumption over time (p=0·0005).

The recurrence score was a strong predictive factor of benefit from CAF for disease-free survival. Figure 3A shows improved disease-free survival over 10 years for CAF-T versus tamoxifen alone in the entire recurrence score sample (stratified log-rank p=0·054, adjusted for number of positive nodes), but degree of CAF benefit depended on the recurrence score. There was no apparent benefit for scores of less than 18 (figure 3B, stratified log-rank p=0·97; HR 1·02, 95% CI 0·54–1·93) or between 18 and 30 (figure 3C, stratified log-rank p=0·48; HR 0·72, 0·39–1·31). However, there was a significant advantage of treatment with CAF-T compared with tamoxifen for patients with a recurrence score of 31 or more (figure 3D, stratified log-rank p=0·033; HR 0·59, 0·35–1·01). 10-year disease-free survival estimates in patients with a low recurrence score were 64% for the CAF-T group versus 60% for the tamoxifen group and, for those with high recurrence score, 55% versus 43%, respectively.

Similar differences in the predictive value of the recurrence score were seen for overall survival over 10 years. There was no significant benefit from CAF for patients with a low (stratified log-rank p=0·63, figure 4A) or intermediate (stratified log-rank p=0·85, figure 4C) recurrence score. However, there was a significant benefit from CAF in patients with a high recurrence score (stratified log-rank p=0·027, figure 4E), which did not vary by age (data not shown). 10-year estimates for overall survival in patients with a high recurrence score were 68% for the CAF-T group and 51% for the tamoxifen group. Corresponding HRs for overall survival for chemotherapy versus no chemotherapy after adjustment for number of positive nodes were 1·18 (95% CI 0·55–2·54; p=0·68) for patients with a low recurrence score, 0·84 (0·40–1·78; p=0·65) for intermediate recurrence score, and 0·56 (0·31–1·02; p=0·057) for high recurrence score. Similar outcomes were seen for breast-cancer-specific survival (figure 4B, D, F), with 10-year estimates for patients with a high recurrence score of 73% for the CAF-T group and 54% for the tamoxifen group (stratified log-rank p=0·033).

Figure 5 shows HRs for disease-free survival for benefit from CAF for the parent trial, the entire recurrence score subset, and then by categorised recurrence score. HRs in the parent trial and the entire recurrence score subset show a consistent benefit over time (ie, proportional hazards), with an effect of chemotherapy lasting beyond 5 years. The data for the high recurrence score subset are suggestive of 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 5 years. In the intermediate group, there might be slight benefit overall, but not in the first 5 years. Confidence intervals are wide because of small numbers of later events.

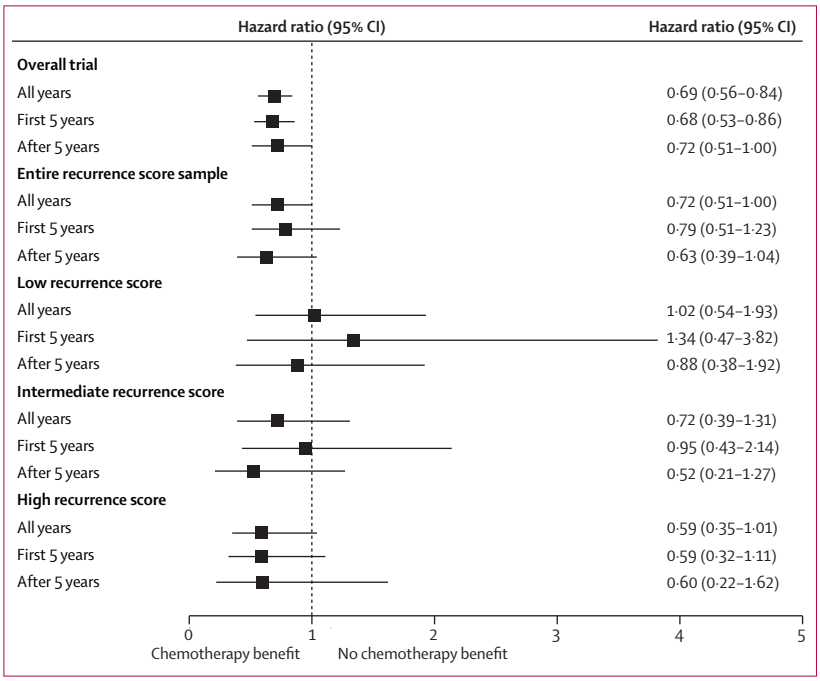

***Figure 5*: Disease-free survival hazard ratios for tamoxifen alone versus CAF-T**
Hazard ratios (95% CI) for the overall parent trial, the entire recurrence score sample, and recurrence score groups of low, intermediate, and high. CAF-T=cyclophosphamide, doxorubicin, and fluorouracil followed by tamoxifen.

	All years	First 5 years	After 5 years
Modelled 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)
Chemotherapy* 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*			
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			
10	0·95 (0·59–1·52)	1·24 (0·62–2·48)	0·72 (0·38–1·36)
18	0·83 (0·56–1·22)	1·03 (0·58–1·81)	0·67 (0·40–1·14)
25	0·74 (0·53–1·04)	0·87 (0·53–1·42)	0·64 (0·39–1·05)
31	0·67 (0·48–0·93)	0·75 (0·48–1·18)	0·61 (0·35–1·04)
40	0·57 (0·39–0·83)	0·61 (0·38–0·96)	0·56 (0·28–1·11)

Data are hazard ratio (HR; 95% CI). RS=recurrence score. *Benefit in disease-free survival of cyclophosphamide, doxorubicin, and fluorouracil followed by tamoxifen (CAF-T) versus tamoxifen alone.

***Table 2*: Disease-free survival hazard ratios adjusted for number of positive nodes for chemotherapy benefit by recurrence score over time**

The primary analysis of prediction was to test increasing chemotherapy benefit as the linear recurrence score increased. We analysed the interaction of treatment effect and the linear recurrence score, adjusting for number of positive nodes (1–3 *vs* ≥4). Table 2 shows the model, calibrated to recurrence score 0 as the referent and recurrence score/50 (ie, corresponds to a 50-point difference). Over the entire period, the recurrence score

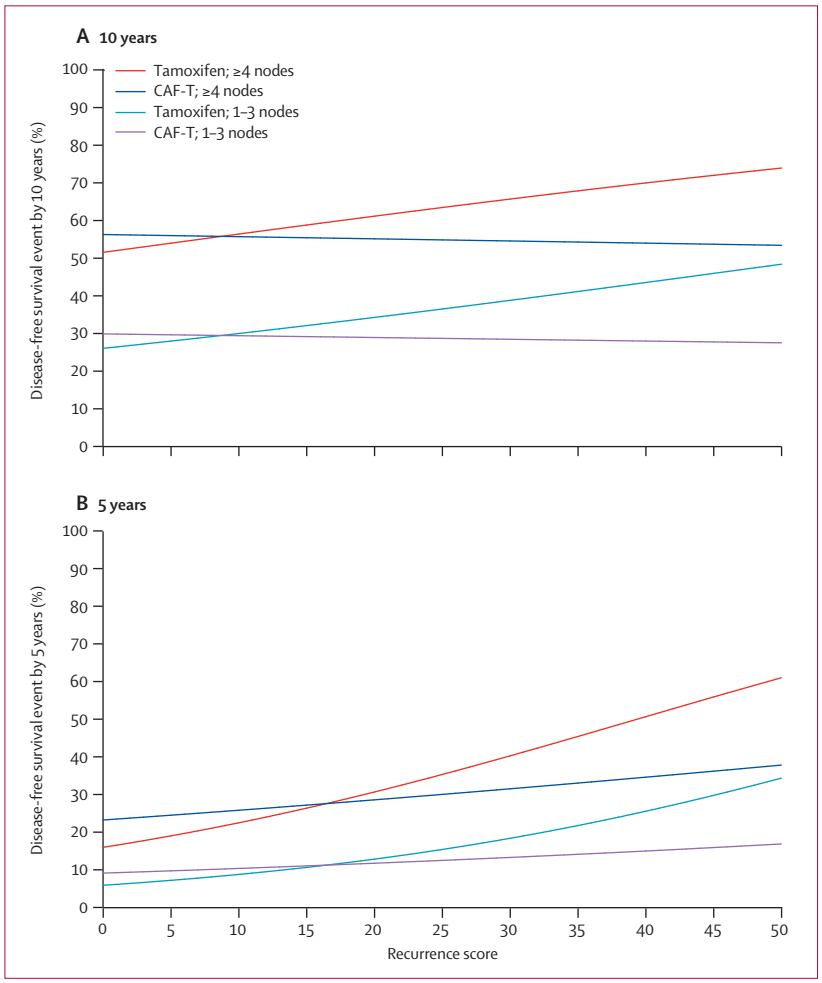

***Figure 6*: Risks of a disease-free survival event by linear recurrence score, treatment, and number of positive nodes, for 10-year and 5-year timepoints**
CAF-T=cyclophosphamide, doxorubicin, and fluorouracil followed by tamoxifen.

See Online for webappendix

by treatment interaction is p=0·053 for disease-free survival. However, the effect of the recurrence score on treatment is not constant over time: recurrence score predicts chemotherapy benefit in the first 5 years (interaction p=0·029), but not after 5 years (interaction p=0·58). Nevertheless, the cumulative benefit of CAF persists up to 10 years. In the analysis of overall survival, there was a significant interaction of recurrence score and treatment over the entire period (p=0·026) and in the first 5 years (p=0·016), but not after 5 years (p=0·87). Therefore, recurrence score 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 those with a higher recurrence score). The strong initial effects carry forward sufficiently so that overall differences are still seen at late timepoints.

Figure 6A shows prediction of any disease-free survival event within 10 years by number of positive nodes, treatment, and recurrence score. Increasing involvement of axillary lymph nodes was prognostic for disease-free survival. The treatments start diverging at a recurrence score of approximately 10, although any clinically significant benefit from CAF is not evident until a much higher recurrence score. Because the recurrence score has better short-term than long-term prediction, estimates at 5 years are shown in figure 6B. The treatments are equivalent up to a recurrence score of approximately 20, but diverge at higher recurrence score values. The 95% prediction intervals around the estimates are shown in webappendix p 2–3. These bounds are specific to a particular recurrence score value so cannot be used to test the significance of chemotherapy benefit, which depends on a range of recurrence score values.

We assessed whether other markers measured by central pathological review could predict degree of chemotherapy benefit as effectively as the recurrence score risk categories. Tumour grade was prognostic for disease-free survival overall (p=0·008), but did not interact with prediction of chemotherapy benefit (p=0·26). 316 (86%) samples with a recurrence score had data for oestrogen receptor expression by Allred scoring,[19] 352 (96%) samples had HER2 by TAB250, and 312 (85%) had both. The best cut-off point for clinical use of Allred-scored oestrogen receptor was 0–6 (n=147, 47%) versus 7–8 (n=169, 53%) with a marginal predictive effect (p=0·16). There might be a benefit from CAF if the disease was HER2-positive or oestrogen-receptor score was six or lower (n=170, p=0·06, stratified log-rank test at 10 years). However, there was no benefit in disease-free survival if oestrogen-receptor score was high (7 or 8) and the disease was HER2 negative (n=142, p=0·81). In this latter group, 58% of patients had a low recurrence score, 24% had an intermediate score, and 18% had a high score.

The interaction of treatment benefit and recurrence score remained significant after adjustment for age, ethnic origin, tumour size, progesterone status, grade, P53, and HER2 by TAB250. Because oestrogen-receptor expression is a part of the recurrence score, adjustment for Allred-scored oestrogen-receptor expression made the interaction non-significant (p=0·15). There was a moderate negative (–0·38) correlation of Allred-scored oestrogen receptor with recurrence score, although some tumours with high oestrogen-receptor expression (by Allred score or by RT-PCR from the recurrence score assay) had a high recurrence score (webappendix p 4–5). Thus, the predictive capability of the recurrence score might not be completely captured by consideration of known markers measured by immunohistochemistry.

Discussion

Our study suggests that patients with involved axillary lymph nodes, but a low recurrence score, do not seem to benefit from anthracycline-based chemotherapy, whereas those with a higher recurrence score have major benefit, independent of the number of positive nodes. TRANSBIG collaborators presented analyses of a non-randomised

cohort of 106 patients with one to three positive nodes. In a subset of patients who were identified as low risk by the 70-gene profile,[3] patients given chemotherapy had similar survival to those who were not.[22] Taken together, these data suggest that there could be subgroups of patients within the oestrogen-receptor-positive, node-positive breast cancer population that do not respond in the expected way to chemotherapy, and that these subgroups can be identified by use of multigene assays.[23,24]

This study challenges the current treatment standard of adjuvant chemotherapy for all women with positive axillary nodes and oestrogen-receptor-positive breast cancer.[25] This standard is based on several decades of phase 3 clinical trials that showed a survival benefit from chemotherapy when added to endocrine therapy alone in premenopausal women and more recently, post-menopausal 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] Avoidance of the toxic effects and other costs of adjuvant chemotherapy when it might not be needed is an important goal in breast cancer treatment.

There is a continuing debate about the role of the recurrence score and other multigene assays in addition to standard pathology variables for prognosis and prediction. Whereas the 21-gene recurrence score assay provides a reproducible method to classify the biology of a given patient's tumour for prediction of chemotherapy benefit, standard pathology testing might provide another means of determining chemotherapy benefit. In exploratory, post-hoc analyses, high levels of oestrogen-receptor protein expression ("endocrine responsiveness") measured centrally predicted no benefit from chemotherapy.[11–13,26] Additionally, St Gallen guidelines endorse the use of degree of endocrine responsiveness in chemotherapy decision making.[26,28] In our study, a subset of patients with a high concentration of oestrogen-receptor protein and HER2-negative disease did not seem to benefit from CAF added to tamoxifen.

A much larger study would be needed to show a significant increase in prediction using a multigene assay after accounting for standard pathological assays. In part, this increase in sample size is attributable to measuring the same pathways by both methods, so one method must have much less measurement error to show improvement. However, our exploratory analysis and those of others have been consistent in showing that a significant interaction between multigene assay and chemotherapy benefit is maintained after adjustment for standard factors.[1,24,29,30] The recurrence score assay provided better discrimination of individual tumour behaviour and a more reliable prediction of patients who would benefit versus those who would not than did 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] Continuing prospective trials should answer how to best select therapy when this discordance exists. For now, the most recent St Gallen guidelines allow the use of multigene assays to select adjuvant therapy.[28]

It remains to be shown 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 is usually to avoid chemotherapy when it was initially thought to be needed before the assay.[6–8,24]

There are limitations to our results. This study included a population of postmenopausal women with oestrogen-receptor-positive, node-positive breast cancer, so whether the findings translate to premenopausal patients is unclear. 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 National Surgical Adjuvant Breast and Bowel Project (NSABP) study with cyclophosphamide, methotrexate, and fluorouracil[5] are based on older standards of chemotherapy, so the predictive value of the recurrence score assay might differ in current practice when other types of chemotherapy or dosing schedules are used. Although high recurrence scores are associated with more pathological complete remissions from taxanes given in the neoadjuvant setting,[31] recurrence score prediction of taxane efficacy from phase 3 trials is not available. Nonetheless, this analysis and others with different gene profiles suggest that certain biological subtypes of breast cancer might be inherently sensitive or resistant to chemotherapy in general.

Our retrospective analysis included a subset of patients from SWOG-8814, although overall treatment effect and demographics were similar to those in the parent trial. In view of the low endpoint event rate, especially in the low recurrence score group, CIs were broad; therefore, estimated benefit of CAF at specific recurrence score values should be interpreted with caution. Whereas there was no apparent benefit from CAF in patients with a low recurrence score for all endpoints, the possibility of benefit cannot be completely ruled out. The lack of proportional hazards seen in our study is substantiated by previous reports about the major effect 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 tumours with a high recurrence score.[1] Finally, our study used disease-free survival as the primary endpoint, since unlike the NSABP analysis,[1] we did not prospectively collect data for distant recurrence-free interval. Thus, the prognostic and predictive effects of the recurrence score might differ because of the inclusion of disease-free survival events such as second primary cancers and breast recurrences. However, the results for breast-cancer-specific survival were also consistent.

Thus, our study provides further data on the value of a multigene assay for prognosis in patients with oestrogen-receptor-positive, node-positive breast cancer treated with adjuvant tamoxifen. Moreover, our results suggest that the 21-gene recurrence score assay might predict which of these patients derive benefit from an anthracycline-based chemotherapy regimen and those who may not, despite higher risk because of positive nodes. Current treatment guidelines generally recommend chemotherapy for high-risk breast cancer.[25] Prospective studies with larger sample sizes are essential to establish who benefits most from modern endocrine therapy plus chemotherapy, and whether use of multigene assays affects survival.

Contributors

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 preparation of the report. All authors participated in data interpretation. The study biostatisticians (WEB, CY, RB) undertook all analyses. All authors reviewed the contents of the report and approved the submitted version.

Conflicts of interest

KSA, RBL, and PR declared occasional speaker's bureau, continuing medical education lecture, or advisory board honoraria for Genomic Health. KSA, WEB, and DFH declared research funding from Genomic Health to their institutions but with no direct payments to themselves. GWS and DCA have served as paid consultants to Genomic Health. DFH has collaborated with Genomic Health on other unfunded research endeavours. PR has ownership interest in Adjuvant Online. CH had equity interest in Genomic Health until June, 2008, when it was divested 100%. SS, RB, FLB, and CY are full-time employees and stockholders in Genomic Health. GNH, I-TY, NED, EPW, JNI, EAP, KIP, LS, JRG, and CKO declared that they have no relevant conflicts of interest.

Acknowledgments

The SWOG-8814 trial was funded entirely by the US National Cancer Institute. This retrospective assessment of banked tumour tissue was supported in part by the following Public Health Service Cooperative Agreement grant numbers awarded by the National Cancer Institute, Department of Health and Human Services, 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, and CA52772; by the National Cancer Institute of Canada and Canadian Cancer Society; and supported in part by Genomic Health.

References

1 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* 2004; **351:** 2817–26.

2 Habel LA, Shak S, Jacobs MK, et al. A population-based study of tumor gene expression and risk of breast cancer death among lymph node-negative patients. *Breast Cancer Res* 2006; **8:** R25.

3 Van de Vijver MJ, He YD, van't Veer LJ, et al. A gene-expression signature as a predictor of survival in breast cancer. *N Engl J Med* 2002; **247:** 1999–2009.

4 Van't Veer LJ, Paik S, Hayes DF. Gene expression profiling of breast cancer: a new tumor marker. *J Clin Oncol* 2005; **23:** 1631–35.

5 Paik S, Tang G, Shak S, et al. Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor-positive breast cancer. *J Clin Oncol* 2006; **24:** 3726–34.

6 Oratz R, Paul D, Cohn AL, Sedlacek SM. Impact of a commercial reference laboratory test recurrence score on decision making in early-stage breast cancer. *J Oncol Pract* 2007; **3:** 182–86.

7 Kamal AH, Loprinzi CL, Reynolds C, et al. How well do standard prognostic criteria predict Oncotype Dx scores? 2007 ASCO Annual Meeting; Chicago, IL, USA; June 1–5, 2007. Abstract 76.

8 Lo SS, Mumby PB, Rychlik K, et al. 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* (in press).

9 Harris L, Fritsche H, Mennel R, 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:** 5287–312.

10 Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. *Lancet* 2005; **365:** 1687–717.

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 randomised clinical trials. *Lancet* 2004; **364:** 858–68.

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; **994:** 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.

15 Berry DA, Cirrincione C, Henderson IC, et al. Estrogen-receptor status and outcomes of modern chemotherapy for patients with node-positive breast cancer. *JAMA* 2006; **295:** 1658–67.

16 Albain KS, Barlow WE, Ravdin PM, et al, for The Breast Cancer Intergroup of North America. Adjuvant chemotherapy and timing of tamoxifen in postmenopausal patients with endocrine-responsive, node-positive breast cancer: a phase 3, open-label, randomised controlled trial. *Lancet* 2009; published online Dec 10, DOI:10.1016/S0140-6736(09)61523-3.

17 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.

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). *Br J Cancer* 2005; **93:** 387–91.

19 Allred DC, Harvey JM, Berardo M, Clark GM. Prognostic and predictive factors in breast cancer by immunohistochemical analysis. *Mod Pathol* 1998; **11:** 155–68.

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; **21:** 2175–97.

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:** 3628–36.

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. DOI:10.1007/s10549-008-0130-2.

23 Sotiriou C, Pusztai L. Gene-expression signatures in breast cancer. *N Engl J Med* 2009; **360:** 790–800.

24 Albain KS, Paik S, van't Veer L. Prediction of adjuvant chemotherapy benefit in endocrine-responsive early breast cancer using multigene assays. *Breast* (in press).

25 National Comprehensive Cancer Network. NCCN practice guidelines in oncology. Invasive breast cancer v.2.2008. http://www.nccn.org (accessed Oct 15, 2009).

26 Goldhirsch A, Wood WC, Gelber RD, et al. Progress and promise: highlights of the international expert consensus on the primary therapy of early breast cancer 2007. *Ann Oncol* 2007; **18:** 1133–44.

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:** R81.

28 Goldhirsch A, Ingle JN, Gelber RD, et al. Thresholds for therapies: highlights of the St Gallen international expert consensus on the primary therapy of early breast cancer 2009. *Ann Oncol* 2009; **20:** 1319–29.

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:** 4063–71.

30 Knauer M, Straver M, Rutgers E, et al. The 70-gene MammaPrint signature is predictive for chemotherapy benefit in early breast cancer. *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:** 7265–77.

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:** 252–60.

33 Saphner T, Tormey DC, Gray R. Annual hazard rates of recurrence for breast cancer after primary therapy. *J Clin Oncol* 1996; **13:** 2738–46.